EasyAsk: An In-App Contextual Tutorial Search Assistant for Older Adults with Voice and Touch Inputs

WEIWEI GAO* and KEXIN DU*, Department of Computer Science and Technology, Tsinghua University, China
YUJIA LUO, Department of Computer Science and Technology, Tsinghua University, China
WEINAN SHI† and CHUN YU, Department of Computer Science and Technology, Tsinghua University, China and Key Laboratory of Pervasive Computing, Ministry of Education, China
YUANCHUN SHI, Department of Computer Science and Technology, Tsinghua University, China, Key Laboratory of Pervasive Computing, Ministry of Education, China, and Qinghai University, China

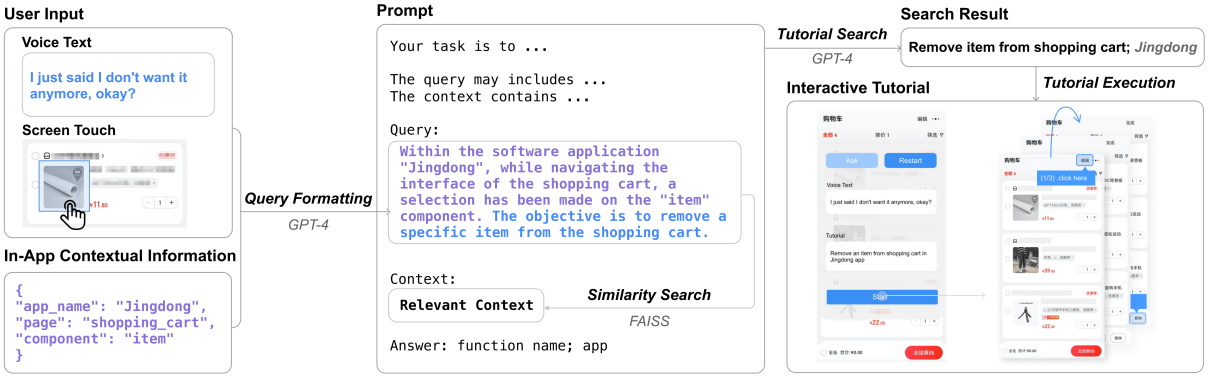

Fig. 1. EasyAsk helps older adults obtain interactive tutorials with voice and touch. Users can invoke EasyAsk to ask a question and touch UI components on the interface to provide necessary information. By recognizing users' questions and utilizing in-app contextual information, EasyAsk formats clear and contextually rich queries based on the verbal questioning patterns of older adults. EasyAsk then searches for a tutorial corresponding to users' intentions. Finally, users can follow the interactive tutorial displayed on their smartphones to complete tasks.

*Both authors contributed equally to this research.
†Corresponding author.

Authors' addresses: Weiwei Gao, gww20@mails.tsinghua.edu.cn; Kexin Du, dkx22@mails.tsinghua.edu.cn, Department of Computer Science and Technology, Tsinghua University, Beijing, China; Yujia Luo, luo-yj23@mails.tsinghua.edu.cn, Department of Computer Science and Technology, Tsinghua University, Beijing, China; Weinan Shi, swn@tsinghua.edu.cn; Chun Yu, chunyu@tsinghua.edu.cn, Department of Computer Science and Technology, Tsinghua University, Beijing, China and Key Laboratory of Pervasive Computing, Ministry of Education, Beijing, China; Yuanchun Shi, shiyc@tsinghua.edu.cn, Department of Computer Science and Technology, Tsinghua University, Beijing, China and Key Laboratory of Pervasive Computing, Ministry of Education, Beijing, China and Qinghai University, Xining, Qinghai, China.

ACM 2474-9567/2024/9-ART102
https://doi.org/10.1145/3678516

An easily accessible tutorial is crucial for older adults to use mobile applications (apps) on smartphones. However, older adults often struggle to search for tutorials independently and efficiently. Through a formative study, we investigated the demands of older adults in seeking assistance and identified patterns of older adults' behaviors and verbal questions when seeking help for smartphone-related issues.

Informed by the findings from the formative study, we designed EasyAsk, an app-independent method to make tutorial search accessible for older adults. This method was implemented as an Android app. Using EasyAsk, older adults can obtain interactive tutorials through voice and touch whenever they encounter problems using smartphones. To power the method, EasyAsk uses a large language model to process the voice text and contextual information provided by older adults, and another large language model to search for the tutorial. Our user experiment, involving 18 older participants, demonstrated that EasyAsk helped users obtain tutorials correctly in 98.94% of cases, making tutorial search accessible and natural.

CCS Concepts: • **Human-centered computing** → **Accessibility systems and tools**; *Interaction design*; *Empirical studies in HCI*.

Additional Key Words and Phrases: Older Adults, Tutorial, Multi-modal Inputs, Context Awareness

ACM Reference Format:
Weiwei Gao, Kexin Du, Yujia Luo, Weinan Shi, Chun Yu, and Yuanchun Shi. 2024. EasyAsk: An In-App Contextual Tutorial Search Assistant for Older Adults with Voice and Touch Inputs. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 102 (September 2024), 27 pages. https://doi.org/10.1145/3678516

1 INTRODUCTION

Many older adults face challenges in navigating smartphones and frequently find it difficult to use various applications [10, 72], and the inability to utilize technologies results in feelings of alienation and being out-of-touch [55]. When encountering problems related to smartphones, older adults generally prefer seeking assistance from a real person [4, 42], but this option may not always be available [3], especially given the growing number of older adults living alone [50]. Indeed, some older adults use smartphones only as classic phones (e.g., making phone calls, texting) [31]. Thus, it is necessary to design assistive tools that can enable older adults to independently and efficiently search for tutorials on their smartphones, which can also significantly increase older adults' independent living [38].

Some older adults can use existing tools to access tutorials. Some turn to general search engines like Google. However, the results from such search engines are often overwhelming, with low-quality tutorials, that are unfriendly for older adults who often have declining cognitive abilities. There are tutorial search tools specifically designed for older adults, but they mainly focus on making the interface age-friendly, providing only visual or video tutorials that bring context-switching costs. Compared to visual and video tutorials, previous studies have shown that interactive tutorials which can provide task-specific step-by-step instructions are the most effective way for older adults to learn smartphones [43, 63]. To help older adults access interactive tutorials, prior work has explored creating multi-modal interactive tutorials by help-givers [30] and automatically generating tutorials based on text and app metadata [88]. However, older adults' inquiries in real life can have many problems (e.g., incomplete speech information, redundant speech information, and ambiguous descriptions). While such work helps older users acquire tutorials efficiently, prior work only focuses on the form and efficiency of generating tutorials but has not systematically studied the patterns of verbal questions and behaviors when older adults ask for help with smartphone functions.

To better understand the demands and patterns of older adults when asking for smartphone assistance, we conducted a formative study including in-depth interviews with 16 older participants over two sessions. In session 1, we identified their current status of seeking assistance. We summarized the main methods that older participants used to seek assistance and the problems they encountered with each method and explored their ideal way of seeking help. In session 2, participants were asked to simulate help-seeking behaviors, and we

summarized the verbal questioning patterns and behavioral patterns of older adults when seeking help for smartphone problems. Findings from the formative study guided the design of our proposed method.

To help older adults obtain easily accessible smartphone tutorials without worrying about organizing the voice commands, we present EasyAsk, a method that incorporates in-app contextual information to help older adults search for interactive tutorials through voice and touch (Figure 1). We implemented EasyAsk as an app on Android devices. Using EasyAsk, older users can directly ask questions on the current interface when they encounter problems using smartphones. If necessary, they can touch UI components on the current interface to supplement information about their problem. EasyAsk automatically supplements in-app contextual information to better understand users' intentions (Figure 1, left). Once the users' voice inputs and contextual information are obtained, EasyAsk uses a large language model to generate formatted queries based on the patterns of older adults' inquiries identified from the previous formative study. Then the formatted queries and relevant information from our tutorial knowledge base are incorporated into a prompt, which is sent to another large language model, searching for the tutorial that best matches the users' intentions (Figure 1, center). Finally, users can follow interactive tutorials with task-specific step-by-step instructions displayed on their smartphones to complete tasks (Figure 1, right). Additionally, EasyAsk has included 578 tutorials from 14 applications.

We conducted a user experiment with 18 older participants to evaluate the accuracy and usability of EasyAsk, compared with the other two baseline methods (Baidu[1] and Mailing[2]). In the user experiment, each participant was first instructed on how to use three methods. Secondly, they used the three methods to complete tutorial search tasks sequentially. Finally, participants rated the three methods on five dimensions (mental demand, physical demand, performance, effort and frustration level) and provided comments on them. In the experiment, we collected data and verbal feedback for both objective evaluation and subjective evaluation. The objective evaluation revealed that EasyAsk achieved an overall accuracy of 98.94% during the experiment. Further ablation study was conducted based on user data, confirming the necessity of supplementing contextual information and formatting older users' queries. Results from the subjective evaluation demonstrated that EasyAsk significantly outperformed two baseline methods in the dimensions of mental and physical demand. Users also gave higher ratings to EasyAsk's performance, expressing a desire to use EasyAsk in their daily lives for tutorial search without requiring human assistance. Participants all reported that EasyAsk allowed them to naturally and effortlessly ask questions without deliberately organizing their language, significantly reducing the burden of inquiry.

Our work contributes:

- Verbal questioning and behavioral patterns of older adults when seeking help for smartphone tutorials, summarized from a formative study
- EasyAsk, a method that incorporates in-app contextual information to help older adults search for tutorials through voice and touch
- User experiment to understand older adults' experience and evaluation of EasyAsk

2 RELATED WORK

In this section, we first introduce the challenges and solutions for older adults in using smartphones (Section 2.1). Then we illustrate two interaction methods suitable for older adults: voice-assistant (Section 2.2) and multi-modal interaction (Section 2.3).

2.1 Research for Smartphone Learning of Older Adults

2.1.1 Challenges for Older Adults Using Smartphones. With the advancement of technology and the exacerbation of global aging, there has been a persistent digital divide between older adults and intelligent devices [15, 17, 48, 52].

[1]https://www.baidu.com
[2]https://mailing.coldlake1.com

Quan-Haase et al. indicated that older adults who are above 65 have lower technological literacy compared to young people [60]. From a cognitive perspective, aging is associated with cognitive declines in spatial working memory and renders learning computer skills increasingly challenging [12, 34, 48]. In terms of physical capabilities, Vaportzi et al. proposed that the decline in physical functions, such as decreased vision, similarly hinders older adults from effectively using smart devices [78]. Psychologically, older adults often experience feelings of helplessness when confronted with smart devices [36], and perceive they have inadequate technological understanding [17, 24].

When encountering difficulties with smartphones, older adults are willing to seek help from others rather than trying by themselves [3]. They exhibit a higher propensity to seek help compared to middle-aged and younger individuals [3]. Research [14] indicated that older adults prefer utilizing social networks to solve problems, including family members [8, 51], friends and other close acquaintances and professional resources [2]. Nevertheless, the resources for older adults to seek help are constrained [3, 9]. Research [3] found that some older adults lack family members or friends and less than half of older adults could obtain responses through social networks.

In addition, few studies focused on older adults seeking help with technology. Existing research [28] indicated that older adults encounter significant difficulties when utilizing text-based input methods during information retrieval. Moreover, due to the challenge of accurately articulating their issues, older adults effectively utilizing search engines for information retrieval poses a significant obstacle [42]. In summary, older adults exhibit a pressing demand for seeking help with smartphones, but they lack good conditions for seeking help.

2.1.2 Tutorial Design for Older Adults. The method of teaching and learning smartphones for older adults is an ongoing concern [42], and there are many studies on tutorial teaching methods [19, 21, 32, 62, 63], such as video tutorial [5, 56] and interactive tutorial [30, 82, 89]. Zhou et al. proposed a method to reduce tutorial steps to improve learning [89]. Hagiya et al. proposed a tutorial to guide text entry [23]. Study [80] investigated the optimal amount of guidance. Previous research has found that the most effective tutorial format for older adults is interactive tutorial [19, 63], compared with step-by-step instructional methods and online help files. Jorge et al. compared video instruction with step-by-step instruction for learning effectiveness [62].

Moreover, some researchers studied how to improve the effectiveness of tutorials. For example, the addition of multi-media cues will make the tutorial clear [84]. Leung et al. proposed a method to increase icon comprehension for older adults [44]. In addition, some studies presented design principles for older adults [86], such as Facebook's design guidelines [6] and Mione Community's design guidelines [16]. Chen et al. summarized design principles for step-by-step tutorials [4]. These studies provide design guidelines for our study, and we will find appropriate ways to follow the guidelines in our tutorials.

As mentioned above, prior work has focused on innovation in the learning method and tutorials for older adults, but few studies have developed methods for tutorial search among older adults. Chen et al. designed a fixed FAB (Floating Action Button) button as a search entry, allowing older adults to search for tutorials anytime during the use of their mobile phones through text input [4]. However, this method still faces issues of unnatural and indirect search methods and presents significant challenges in understanding text inputs provided by older users. In summary, enabling older adults to quickly and naturally search for tutorials is still unresolved. In this paper, we propose a natural and efficient tutorial search method, that alleviates the difficulties older adults face when seeking assistance.

2.2 Voice Assistants for Older Adults

Voice assistants have been widely used in recent years [26]. Using speech is feasible and convenient for older adults [33, 64, 73], and they generally express a willingness to interact with their phones using voice [64, 68]. Moreover, older adults have more positive attitudes toward voice assistants than keyboard interactions [35, 90].

It not only avoids the problem of older adults' understanding graphic metaphors in interface interaction [65], but also solves the problem that older adults interact with interfaces as their vision declines [57]. Tubin et al. found that voice interaction can improve the efficiency of older adults in using devices [75].

The current research on voice assistants for older adults still requires further exploration [65] as voice assistants have a positive impact in many areas of life [1, 7, 58, 65, 69, 85]. However, only a few studies have demonstrated the solutions for voice assistants for older adults. Some research showed preferences of older adults using voice assistants [20, 33, 70], while the special way of expression of older adults poses a great challenge to speech recognition. David Gollasch et al. found older adults tend to use more unusual words and have unclear pronunciation compared to young people [20]. Moreover, older adults often have verbose speaking [33, 49], long breaks between individual words [20, 22, 68], high word error rate [39, 79], unclear sentence structures, and difficulty formulating structured command sentences [33], which lead to mismatches between the content of their speech and the understanding of their speech. Some research focused on older adults' interaction with voice assistants in certain scenes. When cooking in the kitchen, older adults may find it difficult to engage in conversation with assistants, and they often experience challenges with hearing and understanding [37].

The above studies revealed the general voice preferences of older adults. To date, no research has investigated the demands of older adults using voice assistants in help-seeking scenarios, and whether they face other difficulties remains to be explored. Additionally, few studies have provided available solutions to the above problems. Hagiya et al. proposed training models based on acoustic and ASR results to detect speech input errors [22]. Therefore, this paper will focus on exploring the problems and supporting solutions of older adults using voice assistants in help-seeking scenarios where older adults face difficulties in using smart devices.

2.3 Multi-modal Interaction for Older Adults

It has been demonstrated that multi-modal input methods can improve interaction performance [11, 40, 40, 53, 71, 77] and user experience [53, 54, 76]. Users can interact with devices using voice and touch [40, 71, 77]. A speech-assisted error correction system, designed to aid text input on watches, is presented in [40]. Sim et al. designed a touch-assisted error correction system to solve the problem of inaccurate voice input in noisy environments [71]. In addition, Degbelo et al. [11] proposed a combination of speech and position for map editing which can adjust the speech recognition range of commands on an aircraft according to the condition of the physical space. Moreover, Zhao et al. have combined speech and eye position for interaction [87]. Lee et al. proposed a system that combined speech and image [41], which uses a combination of captured images and speech to solve the ambiguity problem of "this" and "that" in utterances.

Gianluca Schiavo et al. proved that older adults are willing to use multi-modal interaction to control devices [67]. However, few studies have focused on multi-modal interaction design for older adults. Some of them explored the preferences of the elderly for multi-modal interactions in different scenarios. Schiavo et al. showed that older adults accept mid-air one-hand gestures and voice commands in daily activities, and seldom use both interactions simultaneously [66]. Kamali et al. found that older adults prefer to interact with assistants by voice and hand gestures in health teaching scenarios [13]. Older adults tend to interact with assistants by voice in cooking scenarios [37]. Other studies focused on improving the efficiency of interaction between older adults and smart devices through multi-modal interaction. Hagiya et al. [22] proposed a method that provides text input to correct speech error rate. Hu et al. [27] proposed a method, using facial recognition and voice, to reduce the hassle of multiple conversations.

In summary, although research on the needs of older adults using multi-modal interaction in seeking-help scenarios is required, the benefits of using multi-modal interaction to improve interaction performance and user experience have been proven. Therefore, our work will also leverage multi-modal capabilities to design our method.

3 FORMATIVE STUDY

To explore the opportunities and challenges of helping older adults search for smartphone tutorials, we conducted a formative study with 16 older adults who have smartphone-related problems. The formative study included in-depth semi-structured interviews to verify the demands of older adults in seeking assistance, and to investigate their behavioral and verbal patterns when seeking help for smartphone-related problems.

3.1 Participants

We recruited 16 older adults who were required to: 1) be aged 55 or over; 2) own a touch-enabled smartphone; 3) have basic knowledge of smartphone use but frequently encounter problems; 4) have experience seeking assistance with smartphone-related problems. Participants were recruited through social media in the local community and by word-of-mouth referrals. IRB approval was obtained from the local community. Each participant received $15 for their participation. Table 1 shows their demographic information. The participants, all Chinese, included 9 females and 7 males, aged between 56 and 76 (*Mean* = 66, *SD* = 5.46). They reported an average smartphone usage duration of 6 years (*SD* = 2.53).

Table 1. Demographics of participants for formative study.

PID	Gender	Age	Educational Attainment	Job	Work Status	Smartphone Usage Years
P1	F	60	Less than high school	Homemaker	Retired	2.5
P2	F	73	High school graduate	Business Manager	Retired	7.5
P3	M	63	Less than high school	Boiler Operator	Retired	5.5
P4	M	63	Associate's degree	Business Manager	Retired	8
P5	M	68	Less than high school	Laborer	Retired	7
P6	M	68	High school graduate	Laborer	Retired	3
P7	M	63	High school graduate	Electrician	Retired	5
P8	F	61	High school graduate	Accountant	Retired	6.5
P9	F	75	Less than high school	Laborer	Retired	2
P10	M	56	High school graduate	Business Manager	Working	10
P11	M	66	Less than high school	Driver	Retired	5
P12	F	69	Less than high school	Laborer	Retired	2
P13	F	68	Less than high school	Laborer	Retired	5
P14	F	67	High school graduate	Laborer	Retired	6
P15	F	76	High school graduate	Teacher	Retired	8
P16	F	63	High school graduate	Laborer	Retired	10

3.2 Procedure

We conducted a one-on-one semi-structured interview with each participant. The interview was divided into two sessions. In session 1, participants were asked about their overall experiences while seeking assistance with smartphone-related problems. In session 2, participants were asked to simulate help-seeking behaviors, and data such as queries and behaviors were recorded.

3.2.1 Session 1: Identifying the Status of Seeking Assistance. To identify the status of older adults when seeking help with smartphone-related problems, we designed 7 questions as follows:

(1) Have you encountered any problems that you did not know how to operate?
(2) When faced with such situations, what actions did you take, and at what frequency did these actions occur?
(3) Why did you take such actions, and what was the psychological motivation behind them?
(4) Are there any negative impacts of such actions, such as lower efficiency or inconvenience in daily life?
(5) What behaviors do you exhibit when seeking help from others? (Ask the older adults to describe their specific behaviors regarding the above-mentioned problem-solving methods.)
(6) What is the most ideal way of seeking help?
(7) Under what circumstances do you proactively seek mobile phone teaching? Is it during the use of the phone, or do you suddenly think of a problem, or are there other circumstances?

3.2.2 Session 2: Simulating Help-seeking Behaviors. To gain a deeper understanding of the behavior of older adults when seeking help and collecting the query data, we implemented the contextual inquiry method [61] and used the passive inquiry approach. We first asked participants to simulate seeking help for smartphone problems, observing and recording their query data and behavior. Secondly, we inquired about their issues and motivations when seeking help, once the participants stopped their simulations.

To improve the authenticity of the data while reducing the cognitive burden on the participants, we wanted each participant to ask questions they had encountered while using mobile applications but had not previously solved. Therefore, we first collected information about the functions they were unfamiliar with when using mobile applications:

(1) What are the apps that are commonly used in daily life?
(2) What are some memorable functions that you would like to perform but do not know how to?

Furthermore, some participants forgot certain functions in the applications or may not know how to describe them. To address this issue, we identified a set of functions in commonly used apps, which older adults may not know how to perform. We then asked them colloquially if they had encountered any problems referencing the set of functions (for example, "Have you ever encountered a situation where you did not want to receive messages from your friends?"). Through the above-mentioned questions, we identified 14 apps commonly used by the participants. These apps contain 240 functions for which the participants often needed help, which can help us analyze the types of problems and design the next tasks for older adults to accomplish.

Afterward, we conducted contextual inquiries with 16 participants. Throughout the contextual inquiries, we collected 240 sets of data from the participants.

3.3 Findings

3.3.1 Current Practice of Seeking Assistance. All the participants reported that they had encountered problems using their smartphones, and they had experiences seeking help. Table 2 shows seven common ways for the participants to seek help, and each participant used various ways.

15 participants, when faced with problems, would ask for help from people around them, especially seeking assistance from younger people. P8 noted "*The problems should be easy for young people to solve.*" P7 also sought peer support, mentioning "*If young people do not have time, I will ask people my age, such as my friends, to help me.*" One-quarter of the participants have taken screenshots of problems and sent them to friends. However, P7 also noted that in many cases, peers of the same age were unable to solve the problem, leading to a situation where he gave up.

But asking people around can also bring problems, especially psychological pressure. 6 participants reported that they had some concerns when asking people around them (P1, P2, P3, P4, P6, P9). P1 was concerned about facing rejection: "*Not all young people will answer my questions.*" P2 was worried about "*wasting other people's time*". Sometimes the help-givers are not very patient: "*My family members might not patiently help me when they are busy, just leaving me to learn to use it on my own, and they won't provide more guidance*" (P9).

Table 2. Common methods used by participants when seeking assistance and the percentage of using each method.

Index	Method to Ask for Help	Use Percentage
M1	Asking for help from people around you	93.75%
M2	Ask for help when children are around	62.5%
M3	Wait for children to return and ask for help	43.75%
M4	Screenshot and ask friends on social media	25%
M5	Internet search	6.25%
M6	Explore on your own	12.5%
M7	Give up	31.25%

43.75% of the participants waited for children to return and ask for help. P3 mentioned that if family members are not around, he tends to remember the problems and asks when they return home, but declining memory becomes an issue: "*I often forget those questions, and waiting for someone to come back home to ask is too troublesome.*"

Only 12.5% of participants have tried to solve problems on their own. P1 and P4 reported concern about making mistakes and causing issues with their smartphones through self-operation ("*If there's no one else around, I won't operate it myself, because I am afraid of making mistakes.*" – P1), which indicates participants' lacking confidence in their own abilities makes them hesitant to explore answers with smartphones by themselves. Only one participant (P10) reported experience searching for tutorials online.

When recalling seeking help behaviors, all participants expressed a preference for directly asking when facing problems using their smartphones. P8 explained that "*If there is a problem, I want to solve it directly, or I'll forget*". They would show the phone to people around them, explaining the problems they encountered, as P2 described: "*I will show people my phone, describing the issue while pointing to the interface.*"

3.3.2 Ideal Way of Seeking Assistance. In session 1, we asked participants about their ideal way of seeking assistance. The key terms primarily mentioned by them included direct interaction, instant answers, and step-by-step instruction. 4 participants expressed the desire to ask questions verbally and directly when having problems using smartphones, as P1 mentioned "*I hope to be able to ask directly when I have a question and get an immediate answer.*" Instant answers are mentioned by five participants. P2 noted "*I want to get answers directly without waiting too long.*" Participants also expressed expectations for tutorial formats. Three participants mentioned that they would like step-by-step tutorials which "*can help the elderly on how to operate step by step*" (P8). Furthermore, participants reported that it would be beneficial to have tools to assist them in searching for tutorials: "*I hope to have a platform where I can ask questions anytime and get immediate answers*" (P6).

3.3.3 Behavioral Patterns. Through tasks in help-seeking simulations, we observed that participants seeking assistance from help-givers primarily engaged in three behaviors: firstly, describing the problems they encountered; secondly, using their phones to supplement necessary information; and thirdly, completing tasks under guided instructions.

The first two behaviors do not have a specific order, and the behaviors of different participants are not entirely consistent across various tasks. In most tasks, participants would describe their problems while showing the corresponding phone interface to the help-givers, especially for tasks that involve specific contextual information (e.g., P7 asked how to send multiple photos at once with showing the phone's photo album). If necessary, participants may simultaneously touch or operate the phone to supplement information (e.g., P10 was charged an extra 20 yuan for a product and he added the product to the shopping cart, showing the price to the help-giver). Sometimes, participants simply point to a specific element on the phone screen (e.g., P6 asked a help-giver to

help him add a person on WeChat by pointing to a phone number). When the problems were relatively simple, some participants asked questions first and then handed the phone to the help-giver for a demonstration. For example, P9 said "*A stranger wants to add me on WeChat. I don't know how to reject. Can you help me?*", then she handed her phone to the help-giver. When the problems are not related to a specific interface, participants directly asked questions without displaying the phone interface (e.g., P5 accidentally uninstalled an app and wanted to restore the uninstalled app. He only asked the question without showing the phone interface).

Regarding task completion, most participants prefer having the help-giver use their phones to demonstrate step-by-step operations. P9, however, prefers holding the phone herself and completing tasks step by step under the verbal guidance of the help-giver. P9 explained "*Because this way can help me learn this function better.*"

3.3.4 Verbal Questioning Patterns. We analyzed the 240 queries collected in help-seeking simulations, summarizing the patterns of queries of older adults when asking smartphone questions (shown in Table 8).

Queries are primarily characterized by incomplete, redundant and ambiguous information. To address the inaccuracies caused by these characteristics, supplementing contextual information and recognizing redundant information is essential, which can help narrow down the search scope. Furthermore, if older participants can indicate objects through manipulation like touch, it can help avoid some confusion.

In the contextual inquiries, we found that although the older participants expressed redundant information when verbalizing problems, all participants indicated that describing a large amount of content would be exhausting for them and they would prefer to state their needs more simply. P5 and P6 stated that they did not want to say so much because it was tiring, but they were afraid that the help-givers would not understand without much information.

The vocabulary of queries was also presented with too specific and inaccurate descriptions, mainly due to older participants' misperception of the function of apps and the components on the interface. To tackle the problems caused by these two characteristics, it is equally important to supplement contextual information and, at the same time, to adjust or generalize their queries based on the patterns of older participants' verbal questioning.

4 METHOD

Inspired by the findings of behavioral and verbal questioning patterns from our formative study, we present EasyAsk, a method that utilizes in-app contextual information to help older adults search for interactive tutorials through voice and touch (Figure 1). The goal of our method is to enable older users to describe problems naturally, directly, and easily, and receive accurate recommendations for interactive tutorials that help solve their problems. We implemented the method as an Android app. The interactive usage process of the app is designed primarily based on user behavioral patterns (section 3.3.3), while the specific technical implementation is mainly inspired by verbal questioning patterns of older adults (section 3.3.4).

From the user's usage process, we take a user Wei who is unable to close the floating window on Wechat as an example, and the EasyAsk usage process (shown in Figure 2) is as follows.

a) Wei is using Wechat to chat with her mom, and she notices there is a floating window on the right side that she does not know how to remove. To remove it, Wei invokes EasyAsk by clicking on the "Ask" button (EasyAsk exists in the form of a floating action button with the label "Ask" on top, which can also be freely dragged).
b) Wei asks "How to delete it" with clicking on the floating window that she wants to remove. EasyAsk automatically monitors when the user's inquiries come to a halt.
c) Wei confirms the voice recognition text is accurate, and clicks on the "Search" button to search for the tutorial with the highest likelihood. EasyAsk displays the name of the tutorial found, which is "Close the floating window on Wechat".

d) After confirming the tutorial is what she needs, Wei clicks on the “Start" button to follow the task-specific step-by-step tutorial to remove the annoying floating window.

The 4-step usage process aligns with the user behavioral patterns (as detailed in section 3.3.3) when seeking help for smartphone problems, enhancing the naturalness, friendliness, and learnability of our method for older adults. The initial steps align with the practices of “describing the problems they encountered" and “using their phones to supplement necessary information", and the latter steps are fundamentally equivalent to “completing tasks under guided instructions".

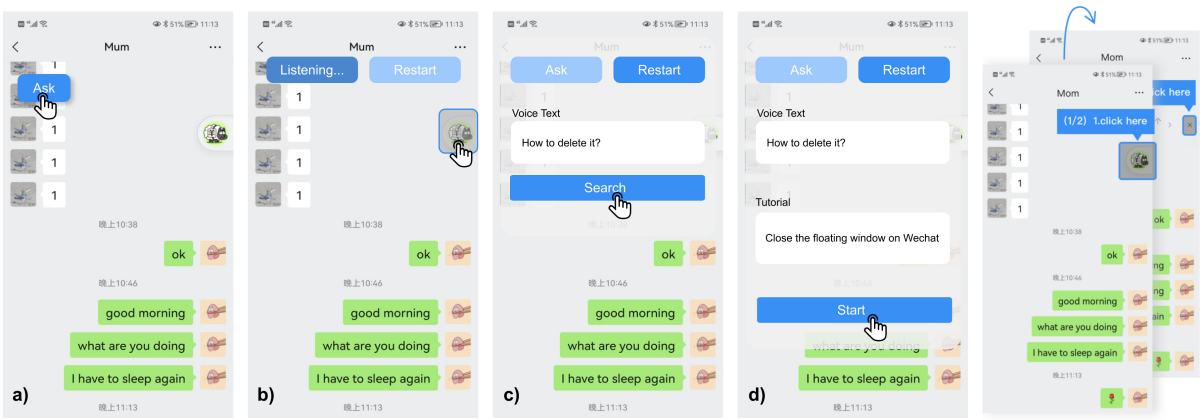

Fig. 2. The user’s process of using EasyAsk to obtain a tutorial about how to close a floating window on Wechat.

To systematically address the problems brought from the verbal questioning patterns identified in the formative study (as detailed in section 3.3.4 and Table 8), we propose three schemes in our technical method: supplementing contextual information, query formatting, and tutorial search.

4.1 Supplementing Contextual Information

Observing the pattern of “incomplete information" (Table 8) such as missing app and interface information, we believe supplementing contextual information for users’ queries can increase the accuracy of the tutorial search. Additionally, to address the fatigue and frustration caused by the “redundant information" as P5 and P6 stated (section 3.3.4), supplementing contextual information can be effective in decreasing the pressure of inquiring. Inspired by the example query of “ambiguous descriptions" (Table 8) as the user could touch on the unfamiliar icon, we found users’ touch interactions contain key information of their problem and we decided to utilize them. Thus, we first use ASR function from the iFlytek API to recognize the user’s voice and convert it into text. Then we use Android accessibility services to capture user touch interactions and gather additional contextual information, such as the current app and specific interface.

4.2 Query Formatting

While the issues of “incomplete information" and “ambiguous descriptions" can be addressed by supplementing contextual information, the problems related to “redundant information", “overly specific vocabulary" (Table 8) in the older users’ queries still require further resolution. To address the issues, we use a language model to modify and optimize user’s query. Using GPT-4’s chat completion API, we applied the technique of few-shot prompting. Our designed prompt (Figure 5) contains contextual information from the users, allowing for more

precise optimization. We collected specific terms associated with each app, emphasizing that the language model should not modify these terms. Within the prompt, we highlighted the verbal questioning characteristics of older users. We also instruct the model to generalize specific vocabulary, enhancing the readability of the users' queries. Additionally, we provide some typical examples to GPT-4 (e.g., "I want to recommend this dragon fruit to my friend" should be rewritten as "Recommend a product to a friend"). We used queries collected in the formative study to verify the effectiveness of the prompt to address the verbal questioning problems of "redundant information" (e.g., "*I saw a video and I liked it, but I swiped past it, so I'm not sure how to find it again*" is modified as "*I want to search for a liked video*") and "overly specific vocabulary" (e.g., "*I want to return these shoes because I don't like them*" is modified as "*I want to return this item*"). Subsequently, the optimized voice text, along with contextual information, is transformed into a formatted query for tutorial search.

4.3 Tutorial Search

Based on the functions unfamiliar to older users collected in the formative study, we built a tutorial database, containing a total of 578 tutorials for 14 commonly used apps. Each tutorial is named according to its features and the corresponding app. The name of each tutorial is unique and can serve as an identifier. We recorded the operation sequences of each function (including opening, clicking events, sliding, etc.). Following that, we recorded the coordinates or IDs of the involved controls and generated interactive tutorial scripts as JSON files. To make step-by-step interactive tutorials, we used scripts to create masked guidance on the interface, guiding users to click on the components to complete operations.

Following that, we built our tutorial knowledge base, as a lengthy document containing the names and explanations of all tutorials. We used GPT-4 to generate explanations for each tutorial based on their names, followed by manual verification from experts. Throughout this process, we validated the GPT-4's understanding of tutorials. We also formatted the document to help GPT-4 understand the structure of the knowledge base, which was indicated in the prompt (Figure 6).

Given the extensive length of our document, we leverage LangChain [3], a framework providing modular abstractions for the components necessary to work with large language models (LLMs), enabling the language model to connect with external data sources. We use tools provided by Langchain to load and split the document into smaller chunks, convert text chunks into embeddings, and use FAISS to perform a similarity search on the chunks with the formatted query. Then, we fill the formatted query ("query" in the prompt) and the most similar chunks ("context" in the prompt) in the prompt (Figure 6), passing it to GPT-4 to generate the answer to the formatted query.

With the limitation of the prompt (Figure 6), the result returned by GPT-4 is limited to the tutorial names in the knowledge base. As we require the most likely tutorial to be returned, GPT-4 will always output a tutorial name. Because of the uniqueness of the tutorial name, the tutorial names in the knowledge base are one-to-one mapped to the names in the tutorial database, so the corresponding tutorial will be matched and returned to the frontend, allowing users to complete the operations.

4.4 Implementation

We implemented EasyAsk as an Android app, which can work on Android phones up to Android 13. Root access is not required in this app. The tutorials and users' interaction logs were saved in a dedicated server.

[3]https://github.com/langchain-ai/langchain

5 USER EXPERIMENT

We conducted a user experiment with two objectives included: one is to collect the query texts from the older users to measure the accuracy of EasyAsk, and the other is to compare EasyAsk with the other two basline methods quantitatively and qualitatively.

5.1 Baseline Methods

We define our baseline methods as how older adults independently seek smartphone tutorials online, excluding the possibility of seeking assistance from others.

In our formative study, we found that participants who use mobile phones to search for tutorials often use Baidu, a well-known and popular search engine. Therefore, to verify if EasyAsk is superior to Baidu, the method commonly used by older adults, Baidu serves as our first baseline method. However, EasyAsk is specifically designed as an assistant for older adults, while Baidu, being a general search engine, is not aimed at older adults in terms of functionality and provides search results not limited to tutorials. Thus, to make a more precise comparison between EasyAsk and a method that also aims to help older adults search for smartphone tutorials, we use Mailing as the other baseline method. Mailing is also a mobile tutorial search program designed for older adults to acquire smartphone tutorials, with a total of 138,872 users [46], despite not being as well-known as Baidu. As a mini-program within WeChat, Mailing has an age-friendly interface and provides over 4,000 step-by-step text and image tutorials, as well as video tutorials. Both baselines support voice and text inputs and can search for tutorials supported by EasyAsk. Additionally, they both display search results in a waterfall layout, with 10 results loaded by default. Therefore, EasyAsk has two baseline methods for comparison, allowing for a more comprehensive evaluation of EasyAsk's usability and providing additional insights. The comparison of the three methods is summarized in Table 3.

Table 3. Comparison of EasyAsk and baseline methods.

	Feature	Tutorial Quantity	Tutorial Format	Age-friendly Design	Inputs	Search Results Format
EasyAsk	Help older adults search smartphone tutorials	500+	Interactive (in-app guidance)	✓	Voice & Touch	Only one result
Mailing	Help older adults search smartphone tutorials	4000+	Non-interactive (text & image, video)	✓	Voice & Text	10 results loaded by default
Baidu	Help people search information	-	Non-interactive (text & image, video)	×	Voice & Text	10 results loaded by default

5.2 Participants

We recruited 18 older adults (shown in Table 4, P15 and P16 also participated in the formative study) with the same requirements and in the same way as the formative study. The participants consist of 9 females and 9 males, with an average age of 67 (SD = 4.22) and an average smartphone usage of 6 years (SD = 5.87). Only P17 and P28 were using iPhones in their daily lives, but they had used Android phones. The rest of the participants were all using Android phones. Each participant was compensated $20 for their time.

We investigated the smartphone-tutorial searching status of the participants. First, we asked their commonly used methods (Table 2) to ask for help when encountering problems using the phone. 5 participants would use "internet search" (M5) method to solve their problems, while most participants preferred asking for help from children and people around them (M1, M2). To explore the understanding and usage of the two baseline methods

among the participants, we separately asked participants about whether they used these two methods in daily life, and if used, what was the frequency. Results showed that 9 participants used Baidu (P16, P20 and P31 noted that they had used Baidu to search for tutorials). However, none of the participants ever used Mailing and only 2 participants (P12 and P15) heard about it, which is mainly related to the inadequate promotion of these elderly assistance apps. Many of these kinds of apps often rely on young people to promote them to the older adults around them. Regarding the favorite inputting methods, 10 participants tended to use the voice input method.

Table 4. Demographics, smartphone usage status, the order of using three methods to complete the task and the quantity of the tasks of participants in user experiment.

PID	Gender	Age	Educational Attainment	Smartphone Usage Years	Methods to Ask for Help	Baidu Usage	Task Order	Task Quantity
P15	Female	63	High school graduate	10	M1	No	MEB	5
P16	Female	76	High school graduate	8	M2,M3,M1,M5	No	EMB	6
P17	Male	66	High school graduate	9	M1,M2,M4,M5	Yes; 3-4x/month	BME	3
P18	Male	71	Less than high school	6.5	M1,M2	Yes; 2x/week	BME	6
P19	Female	60	High school graduate	5	M1,M4	Yes; 4x/week	MBE	4
P20	Male	60	High school graduate	5.5	M1,M2,M4,M5	Yes; freq varies	MBE	7
P21	Male	65	High school graduate	2	M1	Yes; low freq	MBE	7
P22	Female	61	High school graduate	4	M1,M6,M5	No	EBM	6
P23	Female	69	Less than high school	7.5	M2	No	BME	2
P24	Male	64	High school graduate	8	M1	No	EBM	7
P25	Male	67	High school graduate	7	M2,M3,M6	Yes; low freq	EBM	5
P26	Male	67	High school graduate	3	M1,M2	No	MEB	4
P27	Female	70	Associate's degree	7.5	M2,M3	Yes; 1x/month	MEB	8
P28	Female	67	Less than high school	5.5	M1,M2,M3	Yes; low freq	EMB	5
P29	Male	66	High school graduate	5	M4	No	EMB	3
P30	Male	70	Less than high school	8	M1	No	BEM	3
P31	Female	69	High school graduate	3.5	M2,M3,M5	Yes; daily use	BEM	7
P32	Female	71	High school graduate	4.5	M2,M3	No	BEM	6

5.3 Procedure

We first collected the demographic and mobile phone usage information of participants. Then each participant was asked to identify 2 to 10 problems that they frequently encounter when using mobile apps. After thorough communication with the participants to ensure mutual understanding of the problems, we recorded these problems and assigned each participant a task set based on their problems. The tasks in the task set for each participant were to search tutorials for these problems using the three methods (Our tutorial database was extensive enough to support tutorials for the problems identified by each participant). Creating task sets based on individual smartphone usage ensured that the queries posed by participants reflected real-life situations. All the participants used the same Android phone installed with EasyAsk. We had affirmed that 100% of the participants have experience using Android phones. We allowed sufficient time for participants to become familiar with the phone. Following this, we instructed each participant on using three different methods. Once proficiency was achieved in each method, participants used each method sequentially to accomplish their assigned tutorial search tasks. Finally, we conducted a post-stimulus survey, where each participant evaluated the three methods.

5.3.1 Method Instructions. As shown in Table 4, different participants present varying levels of proficiency with smartphones and the Baidu app. Therefore, we provided instructions on using Baidu, Mailing, and EasyAsk to search for smartphone tutorials.

We conducted the instruction in the order of Baidu, Mailing, and EasyAsk. Our instruction of each method can be divided into three steps: case-based teaching, task practice, and independent exploration:

(1) **Case-based Teaching.** Firstly, to maintain consistency in teaching, for each method, we used the same case ("How to close WeChat floating windows") for teaching and explained the case to participants in the beginning. As our core objective is to teach participants how to search for tutorials, we only taught them how to use the three methods for tutorial searches. Although each method has its unique features, the main steps for searching for tutorials are consistent: opening the app, describing the problem, and searching for tutorials. Therefore, we sequentially guided participants through these three steps for each method while providing explanations. If participants raised questions during the process, we addressed them immediately.
   (a) We first showed the location of the app and opened it.
   (b) Then we taught how to describe the problem. Since the interface and interaction of each method differ, the teaching approach varies slightly.
   When teaching Baidu and Mailing apps, we advised participants to supplement app information and ask questions as accurately as possible to improve search result accuracy. We then demonstrated and executed the process of describing the problem, noting the interface operation steps for Baidu and Mailing are basically the same. First, we clicked on the search box, then pressed the "voice search" button, and asked, "How to close WeChat floating windows?".
   When teaching EasyAsk, we informed participants that EasyAsk supports describing problems with voice and touch, and natural inquiries are supported without supplementing app information. Then, we demonstrated and executed the inquiry process. First, we clicked the "ask" button, then clicked on the floating window, and asked, "How to close this?"
   (c) Finally, we displayed how to search tutorials. All three methods automatically initiate a search after the voice inquiries. Thus, after the previous step of describing the problem, the app displayed the search results and we showed them to the participants.
(2) **Task Practice.** After previous teaching, we asked participants to use the method to search for a problem they commonly encounter while using their phones, following the steps taught in the first step. If participants encountered difficulties or forgot how to operate during the practice, we would only provide verbal prompts instead of completing the operations for them.
(3) **Independent Exploration.** Finally, we allowed participants to independently explore the app until they felt confident in using it to search for tutorials and were ready to move on to learning the next method.

We recorded the time taken from the beginning of the instruction to when the participants finished their independent search operations.

5.3.2 Tutorial Search Tasks. Acquiring the basic knowledge of three methods, each participant was asked to sequentially use each method to search for tutorials corresponding to their assigned tasks.

It is noted that participants only need to search and confirm whether they have successfully found the correct tutorial once with each method, without necessarily following the steps to complete the corresponding operations step by step.

We define an accurate search in each task as finding at least one correct tutorial that satisfies the search intent of the participant, enabling them to solve their problem. After the participant completed their search, a researcher examined the search results displayed on the current interface to determine whether the search was accurate or not. Since EasyAsk only returns a single search result, the researcher could assess its correctness.

However, as Baidu and Mailing display multiple results, the researcher identified the first search result meeting the participant's needs to solve their problem among the search results and recorded its order.

To avoid potential interference from the order of using the three methods, we employed a counterbalancing technique. There were a total of six possible sequences, shown in the Table 4 (e.g., "EBM" stands for using EasyAsk first, followed by Baidu, and finally Mailing to complete the task), among which each set of three participants followed the same order, and this assignment was entirely random.

Given the decline in cognitive abilities among older adults, despite prior instruction on the three methods, there was a possibility of forgetting during the experiment. Therefore, we occasionally reminded participants about the usage of different methods during the task if they needed to.

In addition to search accuracy and the order of the first correct search result, we also recorded the task completion time (including search time and time spent browsing search results), whether the participants needed prompts during the task completion, and whether the participants could find the correct search result.

5.3.3 Post-stimulus Survey. After participants completed all the tasks, we conducted a post-stimulus survey which included 5 ratings on a 7-point Likert scale: Mental Demand, Physical Demand, Performance, Effort and Frustration Level. These ratings, selected from NASA-TLX (Task Load Index) questionnaire [25], could help us quantify participants' perceived cognitive workloads and subjective evaluation of the three methods. According to the questions in the NASA-TLX questionnaire, we asked participants to rate the three methods on these five rating scales and provide reasons for there rating scores. We recorded their rating scores, reasons behind the scoring and any spontaneous comments during the survey.

6 RESULTS

In the user experiment, we collected experimental data and ratings from 18 participants. Each participant completed an average of 5 tasks (Table 4), searching for tutorials using three different methods for each task. We present our objective evaluation using data collected in the tutorial search tasks, and then demonstrate subjective evaluation with participants' ratings of the three methods in the post-stimulus survey.

6.1 Objective Evaluation

Utilizing data gathered from tutorial search tasks, we compared user skilled level in three methods, efficiency in completing tasks using these methods, and their respective effectiveness.

Table 5. Objective Evaluation.

	Learning Time (s)		Prompted Participants	Task Completion Time (s)		Order of Correct Result			Participants Finding Correct Search Result
	Mean	SD		Searching	Browsing	Mean	SD	Range	
EasyAsk	198	63.2	28.70%	11	28	-	-	-	97.85%
Mailing	170	69.6	53.20%	22	37	3.77	3.12	1-12	73.53%
Baidu	181	71.9	29.80%	22	32	1.76	1.21	1-6	83.93%

6.1.1 Skilled Level Comparison. The average learning time (Table 5) for participants and the percentage of participants who need prompts (Table 5) when using different methods to complete tasks can indicate their skilled level with these methods. We performed the Friedman test on the average learning time of the three methods and found no significant difference. Although EasyAsk has the longest learning time, it is mainly because we need to teach participants how to use our input method with voice and touch, which is novel for them. Despite being a new input method, participants found this method convenient (P17, P19, P23, P32).

As mentioned in the formative study, we provided occasional prompts about the usage of different methods during the task because of the decline in cognitive abilities among older adults. We prompted participants including how to access different methods and how to ask questions. When using EasyAsk, only 28.70% of the participants need prompts. While Mailing has the highest percentage. This is mainly because the entry path for Mailing is too long (e.g., P17 said "*Finding Mailing is a bit of a headache*"). Additionally, participants lack experience with Mailing, leading to a higher need for prompts. Although participants had never used EasyAsk or Mailing, the percentage of prompted participants of EasyAsk is even lower than Baidu, which is relatively familiar for most participants. Participants reported that compared to the other two methods, using EasyAsk to find tutorials requires fewer steps, which is more friendly for older participants with cognitive and memory decline.

6.1.2 Efficiency Comparison. We measured the efficiency of users completing tasks with multiple metrics. First, we calculated the mean time of task completion. According to the steps of searching tutorial, we split the time into two parts: searching time (from opening the method to clicking on the "Search" button) and browsing time (checking all search results to determine if there is a tutorial that matches their needs).

We performed the Friedman's test on the searching time and browsing time of the three methods, and both metrics showed significant differences ($p < 0.001$; $p < 0.001$). Then we conducted post-hoc tests for paired comparisons. The result showed that participants spent significantly least time searching and browsing in the task completion with EasyAsk compared with Baidu and Mailing ($p < 0.001$; $p < 0.001$). However, there are no significant differences between Baidu and Mailing on these two metrics. The significantly shorter time required for searching tutorials using EasyAsk is mainly due to its user-friendly interface, allowing participants to ask questions directly, saving the cost of switching interfaces. For example, P19 said "*EasyAsk is easy to use, quite practical, and does not require switching interfaces.*"

When faced with numerous search results, participants need to find the correct one. The average order of correct results for Mailing is 3.77 (SD = 3.12, Max = 12), higher than Baidu ($Mean$ = 1.76, SD = 1.21, Max = 6). It can be observed that the browsing time for participants when using Mailing to complete tasks is also longer than with Baidu. However, EasyAsk provides only one result, so this figure was omitted. Even when the correct result is found, if it appears later in the search results, users demonstrate difficulty in identifying or locating this correct result. Only 73.53% of users can find the correct result in Mailing's search results, while for Baidu, this figure is 83.93%. As for EasyAsk, when the search result is correct, only one user (P25) considered it as incorrect (P25 asked "*How to shoot funny videos*", and EasyAsk's search result mentioned "filter", a term unfamiliar to the participant, leading to the misperception).

6.1.3 Effectiveness Comparison. To measure the effectiveness of the three methods in helping older adults access tutorials that can address their problems, we compared their accuracy in search tasks. As the definition of "accurate search" is mentioned in section 5.3.2, we define accuracy as **the number of accurate searches divided by the total number of searches**, and we calculated it as the overall accuracy in Table 6. For precise accuracy calculation, we define a single accurate search as one where there is at least one result meeting the participant's needs among the top-k (k = 1, 3, 5, 10) results. The maximum value of k is set to 10 because both Baidu and Mailing default to displaying only 10 results, and most participants (except for P20) only check whether there were tutorials that meet their requirements among the default top-10 search results. Additionally, since EasyAsk only returns one result, only the top-1 accuracy is calculated. Recorded by one of the researchers during the tutorial search tasks, we calculated the accuracy of the three methods (shown in Table 6).

We applied the chi-square test with Bonferroni correction for pairwise comparisons to determine if there were significant differences in the accuracy of searches across three methods. Results showed that EasyAsk had a significantly higher accuracy of 98.94% ($p < 0.001$) compared to Mailing and Baidu, with only one incorrect search result. In this incorrect case, participant P27 wanted to ask how to search for a product in Pinduoduo, and she

asked "*I want to buy bananas*", while EasyAsk returned the tutorial of buying some product in the shopping cart. The returned tutorial slightly deviated from the participant's needs. The participant is more focused on how to search, but her query was actually ambiguous. The accuracy is influenced by multiple factors: queries, search algorithms, and recommendation algorithms. Among these, queries play a vital role. The accuracy of Mailing is significantly lower than Baidu ($p < 0.001$), because participants were unfamiliar with this method, so they feel nervous each time using it, leading to a decrease in the quality of their queries. For Baidu, although many participants were familiar with it and had a general idea about the recommended results, it provided too many results, and the quality was relatively low. The most important reason why EasyAsk reached high accuracy is that EasyAsk automatically supplements and utilizes the contextual information and formats the query, as Baidu's accuracy indicates that the queries of participants still lack problem-related context even some participants have been familiar with Baidu.

Table 6. Accuracy of EasyAsk and Baseline Methods.

	Total Searches	Overall Accuracy	Top-k Accuracy			
			1	3	5	10
EasyAsk	94	98.94%	98.94%	-	-	-
Mailing	94	36.17%	7.45%	22.34%	25.53%	31.91%
Baidu	94	59.57%	36.17%	53.19%	56.38%	57.45%

To further validate the necessity and effectiveness of supplementing contextual information and query formatting in our method, we conducted an ablation study of EasyAsk. We set three additional models (E00, E01, and E10) based on whether contextual information is used and whether the query is formatted. We calculated the accuracy of these three models on the same dataset collected in the user experiment, and the results are shown in Table 7. We then conducted the chi-square test with Bonferroni correction for multiple comparisons to determine if there were significant differences between the models. For model E00, we used participants' raw voice input without any contextual information as queries to search for the tutorial. The accuracy of model E00, not surprisingly, turned out to be the lowest as 68.09%, which is significantly lower than E10 and E11 ($p < 0.05$; $p < 0.05$). However, it is still higher than Baidu and Mailing, due to the power of the large language model. For model E10, we used a new prompt to optimize participants' voice inputs only based on verbal questioning patterns, without contextual information. Then the optimized queries with the prompt are sent to the GPT-4 model to search for tutorials. Model E01 only improved the accuracy by one inquiry compared to E00. For model E10, the queries are unformatted, containing raw voice inputs and contextual information. The accuracy of E10 is 94.68%, significantly higher than E00 ($p < 0.05$) and E01 ($p < 0.05$), but still slightly lower compared to EasyAsk. The comparison of the E00 and E01 showed that without contextual information, query formatting was unable to significantly improve accuracy. Comparing the accuracy of E10 and E00, we concluded that contextual information played a crucial role in understanding users' intentions. While contextual information is important, query formatting is indispensable according to the accuracy gap between E10 and EasyAsk.

Overall, EasyAsk demonstrated the highest accuracy, and we validated the necessity of supplementing contextual information and query formatting, as two key schemes in our method.

6.2 Subjective Evaluation

In this section, to present the subjective evaluation of participants, we first illustrate the quantitative results from the 7-point Likert scale ratings and then present qualitative results recorded in the post-stimulus survey.

Table 7. Ablation Study of EasyAsk.

Model	Contextual Information	Query Formatting	Accuracy
E00			68.09%
E01		✓	69.15%
E10	✓		94.68%
EasyAsk	✓	✓	98.94%

6.2.1 Subjective Ratings. Figure 3 and 4 showed participants rating scores for EasyAsk, Mailing and Baidu in the post-stimulus survey. Overall, EasyAsk has achieved the best performance in five ratings (Figure 3). In comparison with Mailing and Baidu, participants rated EasyAsk with significantly less mental demand (*Mean* = 1.78, *SD* = 1.44 vs. *Mean* = 3.39, *SD* = 1.89 in Mailing vs. *Mean* = 3.22, *SD* = 1.90 in Baidu), less physical demand (*Mean* = 1.33, *SD* = 0.57 vs. *Mean* = 2.44, *SD* = 1.77 in Mailing vs. *Mean* = 2.17, *SD* = 2.25 in Baidu). Compared with Mailing, EasyAsk demonstrated significantly better performance (*Mean* = 1.61, *SD* = 1.06 vs. *Mean* = 2.67, *SD* = 1.20 in Mailing), significantly less effort (*Mean* = 1.67, *SD* = 1.41 vs. *Mean* = 2.94, *SD* = 1.72 in Mailing), as well as significantly lower frustration level (*Mean* = 1.24, *SD* = 0.55 vs. *Mean* = 2.22, *SD* = 1.72 in Mailing). We calculated the significance of the five ratings between EasyAsk and the baseline methods with the Wilcoxon Signed-Rank test.

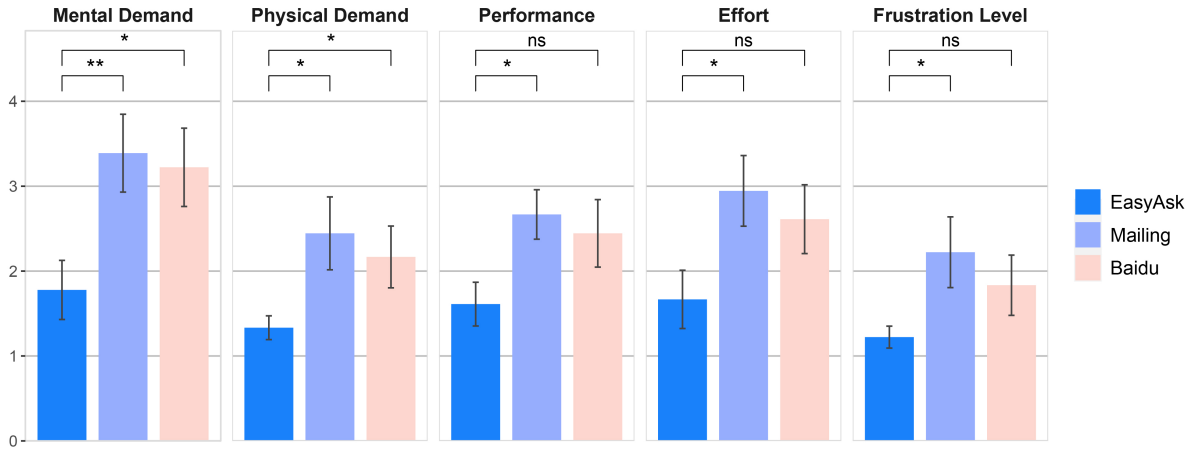

Fig. 3. The mean and standard deviation of rating scores for EasyAsk and baseline methods. $^{*}p < 0.05$, $^{**}p < 0.01$, ns = no significance.

6.2.2 Qualitative Feedback. Overall, all participants stated that they would like to use EasyAsk to search for tutorials when encountering smartphone problems in the future: "*I hope to be able to use this app in my daily life as it is very convenient*" (P18), "*I feel that EasyAsk is more practical, and compared to the other two, I prefer EasyAsk more. With this method, I won't have to trouble others when I have problems with my phone*" (P21), "*I appreciate this style of questioning; it doesn't require much effort mentally for me to describe the problem, which makes me feel quite at ease*" (P27).

Participants acknowledged the effectiveness of EasyAsk in searching for tutorials as 12 participants indicated the search results totally matched their intentions. P17 thought the search results of EasyAsk were "*more perfect*".

P20 said "*EasyAsk's results are more accurate. Although Baidu and Mailing returned more results, they are not precise.*" 10 participants reported the results were easy to understand. P18 noted that Baidu provided too many search results, which made it hard to understand each result, and he also reported the inaccurate results of Mailing and Baidu results made him feel frustrated.

Specifically, participants emphasized that our method alleviated their inquiry burden mentally and physically. 4 participants emphasized the questioning is straightforward and less frustrating (P27, P30, P21, P18). P27 said, "*I appreciate this style of questioning; it does not require much mental effort to describe the problem, which makes me feel quite at ease.*" P30 similarly mentioned, "*I do not need to say too many words, and I do not have to describe the question very clearly.*" P21 said that he had to mention the app using Mailing and consider how to organize his query using Baidu, and he had to say more stuff using these two methods than with EasyAsk, which increased his mental demand.

Participants also affirmed that the interaction and interface design of EasyAsk made it easier for them to learn and use. As P19 mentioned, she could "*stay on the current interface to ask questions*", which was really helpful as EasyAsk saved the cost of frequently switching interfaces to make inquiry like baseline methods. Some participants also indicated that the feature of our method returning only one result reduced their fatigue and frustration when confirming tutorials. P18 reported that the results recommended by Baidu and Mailing were too many, causing fatigue in his eyes. P26 pointed out that "*EasyAsk is easy to use and presents few results, which is convenient*". However, 2 participants (P25, P31) emphasized familiarity with the methods was important and they were not certain about their feelings towards EasyAsk ("*Despite EasyAsk being straightforward, I am not familiar with it.*" – P25).

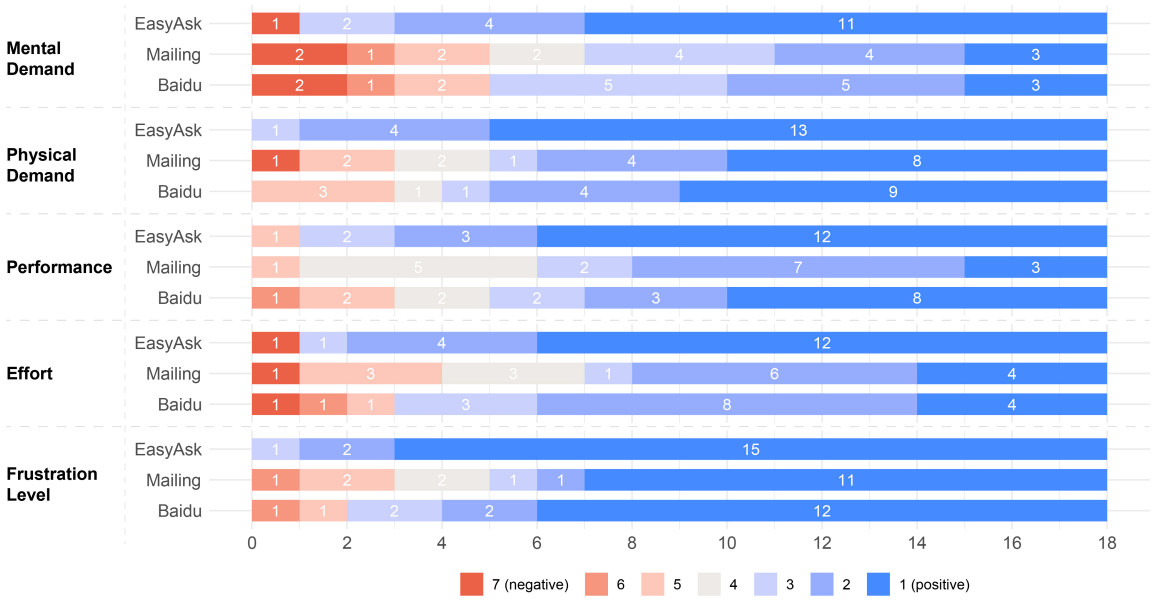

Fig. 4. Distribution of the rating scores for EasyAsk and the baseline methods. 1 = positive, 7 = negative.

7 DISCUSSION

In this section, we reflect on the formative study, development of EasyAsk and user experiment, and summarize our findings and implications. We also discuss the limitations of our work and future work for research aimed at helping older adults seek help for smartphone problems.

7.1 Older Adults' Behavioral and Verbal Questioning Patterns

Teaching older adults how to use mobile phones is valuable [15, 17, 48, 52]. However, previous research focused primarily on teaching methods and neglected the issue of older adults seeking help. In the formative study, we summarized older adults' current practice of seeking help, investigated their ideal way of seeking help, and found the behavioral and verbal patterns of older adults when seeking help for smartphone-related problems. We found that older adults turn to others for help (Table 2), which confirms the findings of previous studies [8, 51]. Moreover, it often leads to unsolved problems as their help-seeking behavior depends on others, consistent with the conclusion of a previous study [3]. Therefore, enabling older adults to quickly and accurately find the tutorials they need is highly beneficial for them to use mobile phones. Tools that can provide technology help may benefit from the ideal way of seeking help as it reveals older adults' needs. Most importantly, our findings about behavioral and verbal patterns of older adults when seeking help can serve as a reference for creating tools intended for older adults in the future, providing inspiration for both interaction and interface design.

7.2 Contextual Information Supplementation

To address errors related to patterns of older users' voice text, we automatically supplemented and utilized the in-app contextual information, which effectively help understand the needs of older users. The method we used for contextual data may inspire more research on integrating contextual information with users' queries. However, our contextual information is not comprehensive and is limited to the apps only. In the future, we hope to supplement more contextual information such as GPS and time data. Additionally, we aim to include non-contextual information, such as users' inherent preferences and demographic information [45], to enhance user experience and improve the accuracy of search results. Moreover, in the future, EasyAsk could integrate vision models to better understand of users' interface details and provide more contextual information [83].

7.3 Implications for Interaction Design for Older Adults

Based on a profound understanding of older adults' behavioral and verbal patterns in the formative study, we carried out the interaction design of EasyAsk. We found that the majority of the participants prefer immediate inquiries, asking questions as they encounter questions. Therefore, EasyAsk exists in the form of a fixed FAB (Floating Action Button) button [4], providing just-in-time service. As older users' inquiries often involve natural language and gestures, our design facilitates the users in searching tutorials with voice and touch. Due to a decline in cognitive abilities, decision-making, and choice-making skills in older adults, EasyAsk presents only one result to reduce users' cognitive load. In the user experiment, we observed a gap between "subjective" and "objective" accuracy, originating from the cognitive abilities of the elderly. When the correct result is found, if it appears later in the search results, users have difficulty in identifying or locating this correct result, leading to a lower "subjective" accuracy. This finding may provide implications for future tools when presenting information to older users. Besides, as older users tend to prefer step-by-step tutorials [30, 82, 89], the tutorials we provide are exactly catering to this learning preference. Results in subjective evaluation between EasyAsk and the baseline methods validated the usability of EasyAsk, proving the effectiveness of our interaction design.

7.4 Limitations and Future Work

Although we succeeded in developing a method to make tutorial search accessible for older adults, our study has some limitations.

Given the expanding functions of apps and diverse user demands, the current 578 tutorials from 14 applications supported by EasyAsk are certainly insufficient. Considering that tutorials can be recorded with smartphones, future efforts will explore diverse approaches like crowdsourcing to create and collect tutorials, enabling more users to contribute as help-givers. Inspired by PromptRPA [29], we consider generating tutorials automatically only based on inquiries and in-app information with RPA tools. We will also improve the quality of tutorials, supporting various tutorial formats such as Trial-and-Error mode [30].

To assist less tech-savvy older adults with lower cognitive ability to search and confirm tutorials, our method only returns a single result. Therefore, when the tutorial provided does not solve their problem, they have to initiate a new inquiry. Although the current method has achieved a high level of accuracy, and user experiments have confirmed the rationality and effectiveness of this design choice, we will further improve the interaction process in the future to better handle cases where searches fail. For example, multi-turn inquiries [59] or providing related tutorials could be helpful when search results are incorrect.

Regarding privacy and security, researchers found that older adults lack a nuanced understanding of mobile security and privacy [18, 81]. We acknowledge that our method has prioritized effectiveness and usability over adequately addressing the privacy and security concerns of older adults. First, customized training and educational efforts are necessary to address privacy and security concerns and misconceptions among older adults [81]. Second, to protect the privacy and ensure the safety of older users when using EasyAsk, we consider improving transparency and control [18] by standardizing and being transparent regarding the types, amount, and granularity of information collected and shared. The interface of EasyAsk should be improved to explicitly state when data is transmitted over the cloud as opposed to being processed locally [18]. Further, to alleviate concerns about uploading personal information to the cloud, we also plan to employ on-device LLM such as Gemini Nano [74] and Phi-3-mini [47] that can be executed locally on mobile devices, allowing performing tutorial search tasks offline.

8 CONCLUSION

In this paper, we presented EasyAsk, a method that incorporates in-app contextual information to help older adults search for interactive tutorials through voice and touch. Guided by a formative study with 16 older participants, we implemented EasyAsk as an Android app, enabling users to ask questions naturally and directly, and obtain task-specific step-by-step interactive instructions to complete tasks. To empower EasyAsk, we used large language models to understand the voice text and contextual information provided by older users, and search for tutorials to help users solve their problems. Our user experiment, involving 18 older participants, demonstrated the usability and accuracy of EasyAsk, providing superior support compared to baseline methods, with an accuracy rate of 98.94%. We hope that our work will catalyze future work in providing technology assistance to older adults, transcending the limitations of smartphones.

ACKNOWLEDGMENTS

This work is supported by the Natural Science Foundation of China under Grant No. 62132010, Beijing Key Lab of Networked Multimedia, Institute for Artificial Intelligence, Tsinghua University (THUAI), and Beijing National Research Center for Information Science and Technology (BNRist). The authors gratefully acknowledge insightful feedback on draft versions from Yuntao Wang.

REFERENCES

[1] Ali Abdolrahmani, Ravi Kuber, and Stacy M Branham. 2018. " Siri Talks at You" An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In *Proceedings of the 20th international ACM SIGACCESS conference on computers and accessibility*. 249–258.

[2] Maria Bakardjieva. 2005. Internet society: The Internet in everyday life. *Internet Society* (2005), 1–232.

[3] Statistics Bureau. 2010. Ministry of internal affairs and communications, Japan. *Annual Report on Current Population Estimates* (2010).

[4] Xiaoou Chen, Fei Wang, Zhenwei You, Xiaochun Wang, Chunjing Tao, Jian Liu, et al. 2017. Design of interactive tutorials on mobile applications for chinese middle-aged and older adults. *Art and Design Review* 5, 03 (2017), 162.

[5] Pei-Yu Chi, Sally Ahn, Amanda Ren, Mira Dontcheva, Wilmot Li, and Björn Hartmann. 2012. MixT: Automatic Generation of Step-by-Step Mixed Media Tutorials. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/2380116.2380130

[6] Wonchan Choi, Julia Carranza, and Melissa Fox. 2013. Guidelines for older-adult-friendly online tutorial for Facebook: Content, design, and training principles. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–4.

[7] Leigh Clark, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garaialde, Justin Edwards, Brendan Spillane, Emer Gilmartin, Christine Murad, Cosmin Munteanu, et al. 2019. What makes a good conversation? Challenges in designing truly conversational agents. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[8] Teresa Correa, Joseph D Straubhaar, Wenhong Chen, and Jeremiah Spence. 2015. Brokering new technologies: The role of children in their parents' usage of the internet. *New Media & Society* 17, 4 (2015), 483–500.

[9] Cédric Courtois and Pieter Verdegem. 2016. With a little help from my friends: An analysis of the role of social support in digital inequalities. *New media & society* 18, 8 (2016), 1508–1527.

[10] Sara J Czaja, Walter R Boot, Neil Charness, and Wendy A Rogers. 2019. *Designing for older adults: Principles and creative human factors approaches*. CRC press.

[11] Auriol Degbelo and Sulaxan Somaskantharajan. 2020. Speech-based interaction for map editing on mobile devices: a scenario-based study. In *Proceedings of Mensch und Computer 2020*. 343–347.

[12] Katharina V Echt, Roger W Morrell, and Denise C Park. 1998. Effects of age and training formats on basic computer skill acquisition in older adults. *Educational Gerontology: An International Quarterly* 24, 1 (1998), 3–25.

[13] Mira El Kamali, Leonardo Angelini, Denis Lalanne, Omar Abou Khaled, and Elena Mugellini. 2020. Multimodal conversational agent for older adults' behavioral change. In *Companion Publication of the 2020 International Conference on Multimodal Interaction*. 270–274.

[14] Rebecca Eynon and Ellen Helsper. 2015. Family dynamics and Internet use in Britain: What role do children play in adults' engagement with the Internet? *Information, Communication & Society* 18, 2 (2015), 156–171.

[15] Mei Lan Fang, Sarah L Canham, Lupin Battersby, Judith Sixsmith, Mineko Wada, and Andrew Sixsmith. 2019. Exploring privilege in the digital divide: implications for theory, policy, and practice. *The Gerontologist* 59, 1 (2019), e1–e15.

[16] Sónia Ferreira and Ana Isabel Veloso. 2019. Older adults in ICT contexts: recommendations for developing tutorials. In *Human Aspects of IT for the Aged Population. Design for the Elderly and Technology Acceptance: 5th International Conference, ITAP 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part I 21*. Springer, 376–387.

[17] Thomas N Friemel. 2016. The digital divide has grown old: Determinants of a digital divide among seniors. *New media & society* 18, 2 (2016), 313–331.

[18] Alisa Frik, Leysan Nurgalieva, Julia Bernd, Joyce Lee, Florian Schaub, and Serge Egelman. 2019. Privacy and security threat models and mitigation strategies of older adults. In *Fifteenth symposium on usable privacy and security (SOUPS 2019)*. 21–40.

[19] Syed Ghayas, Suziah Sulairnan, Jafreezal Jaafar, and Saipunidzam Mahammad. 2014. Motivational scaffolding: Tackling the challenges of using mobile applications. In *2014 3rd International Conference on User Science and Engineering (i-USEr)*. IEEE, 48–51.

[20] David Gollasch and Gerhard Weber. 2021. Age-related differences in preferences for using voice assistants. In *Proceedings of Mensch und Computer 2021*. 156–167.

[21] Marjolein den Haan, Rens Brankaert, Gail Kenning, and Yuan Lu. 2021. Creating a social learning environment for and by older adults in the use and adoption of smartphone technology to age in place. *Frontiers in Public Health* 9 (2021), 568822.

[22] Toshiyuki Hagiya, Keiichiro Hoashi, and Tatsuya Kawahara. 2018. Voice Input Tutoring System for Older Adults using Input Stumble Detection. In *23rd International Conference on Intelligent User Interfaces*. 415–419.

[23] Toshiyuki Hagiya, Toshiharu Horiuchi, and Tomonori Yazaki. 2016. Typing tutor: Individualized tutoring in text entry for older adults based on input stumble detection. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 733–744.

[24] Eszter Hargittai, Anne Marie Piper, and Meredith Ringel Morris. 2019. From internet access to internet skills: digital inequality among older adults. *Universal Access in the Information Society* 18 (2019), 881–890.

[25] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[26] Matthew B Hoy. 2018. Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Medical reference services quarterly* 37, 1 (2018), 81–88.
[27] Yaxin Hu, Yuxiao Qu, Adam Maus, and Bilge Mutlu. 2022. Polite or direct? Conversation design of a smart display for older adults based on politeness theory. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
[28] Man Huang, Derek Hansen, and Bo Xie. 2012. Older adults' online health information seeking behavior. In *Proceedings of the 2012 iConference*. 338–345.
[29] Tian Huang, Chun Yu, Weinan Shi, Zijian Peng, David Yang, Weiqi Sun, and Yuanchun Shi. 2024. PromptRPA: Generating Robotic Process Automation on Smartphones from Textual Prompts. arXiv:2404.02475 [cs.HC]
[30] Xiaofu Jin, Xiaozhu Hu, Xiaoying Wei, and Mingming Fan. 2022. Synapse: Interactive Guidance by Demonstration with Trial-and-Error Support for Older Adults to Use Smartphone Apps. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–24.
[31] Abraham Sahilemichael Kebede, Lise-Lotte Ozolins, Hanna Holst, and Kathleen Galvin. 2022. Digital Engagement of Older Adults: Scoping Review. *Journal of Medical Internet Research* 24, 12 (2022), e40192.
[32] Caitlin Kelleher and Randy Pausch. 2005. Stencils-based tutorials: design and evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 541–550.
[33] Sunyoung Kim et al. 2021. Exploring how older adults use a smart speaker–based voice assistant in their first interactions: qualitative study. *JMIR mHealth and uHealth* 9, 1 (2021), e20427.
[34] Masatomo Kobayashi, Atsushi Hiyama, Takahiro Miura, Chieko Asakawa, Michitaka Hirose, and Tohru Ifukube. 2011. Elderly user evaluation of mobile touchscreen interactions. In *Human-Computer Interaction–INTERACT 2011: 13th IFIP TC 13 International Conference, Lisbon, Portugal, September 5-9, 2011, Proceedings, Part I 13*. Springer, 83–99.
[35] Jarosław Kowalski, Anna Jaskulska, Kinga Skorupska, Katarzyna Abramczuk, Cezary Biele, Wiesław Kopeć, and Krzysztof Marasek. 2019. Older adults and voice interaction: A pilot study with google home. In *Extended Abstracts of the 2019 CHI Conference on human factors in computing systems*. 1–6.
[36] Sanna Kuoppamäki, Riitta Hänninen, and Sakari Taipale. 2022. Enhancing older adults' digital inclusion through social support: A qualitative interview study. In *Vulnerable People and Digital Inclusion: Theoretical and Applied Perspectives*. Springer, 211–230.
[37] Sanna Kuoppamäki, Razan Jaberibraheem, Mikaela Hellstrand, and Donald McMillan. 2023. Designing Multi-Modal Conversational Agents for the Kitchen with Older Adults: A Participatory Design Study. *International Journal of Social Robotics* 15, 9 (2023), 1507–1523.
[38] Sanna Kuoppamäki, Sylvaine Tuncer, Sara Eriksson, and Donald McMillan. 2021. Designing Kitchen Technologies for Ageing in Place: A Video Study of Older Adults' Cooking at Home. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 69 (jun 2021), 19 pages. https://doi.org/10.1145/3463516
[39] András Láng, Béla Birkás, András N Zsidó, Dóra Ipolyi, and Norbert Meskó. 2021. It takes two to tango: development, validation, and personality correlates of the acceptance of sugar relationships in older men and women scale (ASR-OMWS). *Frontiers in Psychology* 12 (2021), 592138.
[40] Jaedong Lee, Changhyeon Lee, and Gerard Jounghyun Kim. 2017. Vouch: multimodal touch-and-voice input for smart watches under difficult operating conditions. *Journal on Multimodal User Interfaces* 11 (2017), 289–299.
[41] Jaewook Lee, Sebastian S Rodriguez, Raahul Natarrajan, Jacqueline Chen, Harsh Deep, and Alex Kirlik. 2021. What's This? A Voice and Touch Multimodal Approach for Ambiguity Resolution in Voice Assistants. In *Proceedings of the 2021 International Conference on Multimodal Interaction*. 512–520.
[42] Rock Leung, Charlotte Tang, Shathel Haddad, Joanna Mcgrenere, Peter Graf, and Vilia Ingriany. 2012. How older adults learn to use mobile devices: Survey and field investigations. *ACM Transactions on Accessible Computing (TACCESS)* 4, 3 (2012), 1–33.
[43] Rock Leung, Charlotte Tang, Shathel Haddad, Joanna Mcgrenere, Peter Graf, and Vilia Ingriany. 2012. How Older Adults Learn to Use Mobile Devices: Survey and Field Investigations. *ACM Trans. Access. Comput.* 4, 3, Article 11 (12 2012), 33 pages. https://doi.org/10.1145/2399193.2399195
[44] Rock Anthony Leung. 2011. *Improving the learnability of mobile devices for older adults*. Ph. D. Dissertation. University of British Columbia.
[45] Xiaojian Liu, Yi Zhu, and Xindong Wu. 2023. Joint user profiling with hierarchical attention networks. *Frontiers of Computer Science* 17, 3 (2023), 173608.
[46] Mailing. 2024 (accessed Apr 3, 2024). Mailing. https://mailing.coldlake1.com.
[47] Microsoft. 2024 (accessed Apr 28, 2024). Phi-3. https://azure.microsoft.com/en-us/blog/introducing-phi-3-redefining-whats-possible-with-slms.
[48] Hazwani Mohd Mohadisdudis and Nazlena Mohamad Ali. 2014. A study of smartphone usage and barriers among the elderly. In *2014 3rd international conference on user science and engineering (i-USEr)*. IEEE, 109–114.
[49] Linda Mortensen, Antje S Meyer, and Glyn W Humphreys. 2006. Age-related effects on speech production: A review. *Language and Cognitive Processes* 21, 1-3 (2006), 238–290.
[50] United Nations. 2017. Living arrangements of older persons: A report on an expanded international dataset.

[51] Barbara Barbosa Neves, Fausto Amaro, and Jaime RS Fonseca. 2013. Coming of (old) age in the digital age: ICT usage and non-usage among older adults. *Sociological research online* 18, 2 (2013), 22–35.
[52] Alan F Newell. 2011. *Design and the digital divide: insights from 40 years in computer support for older and disabled people.* Morgan & Claypool Publishers.
[53] Sharon Oviatt. 1999. Mutual disambiguation of recognition errors in a multimodel architecture. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 576–583.
[54] Sharon Oviatt. 2003. User-centered modeling and evaluation of multimodal interfaces. *Proc. IEEE* 91, 9 (2003), 1457–1468.
[55] Jari Pirhonen, Luciana Lolich, Katariina Tuominen, Outi Jolanki, and Virpi Timonen. 2020. "These devices have not been made for older people's needs"–Older adults' perceptions of digital technologies in Finland and Ireland. *Technology in Society* 62 (2020), 101287.
[56] Suporn Pongnumkul, Mira Dontcheva, Wilmot Li, Jue Wang, Lubomir Bourdev, Shai Avidan, and Michael F. Cohen. 2011. Pause-and-Play: Automatically Linking Screencast Video Tutorials with Applications. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 135–144. https://doi.org/10.1145/2047196.2047213
[57] Alisha Pradhan, Amanda Lazar, and Leah Findlater. 2020. Use of intelligent voice assistants by older adults with low technology use. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 4 (2020), 1–27.
[58] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. " Accessibility Came by Accident" Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on human factors in computing systems*. 1–13.
[59] Chen Qu, Liu Yang, W. Bruce Croft, Johanne R. Trippas, Yongfeng Zhang, and Minghui Qiu. 2018. Analyzing and Characterizing User Intent in Information-seeking Conversations. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (Ann Arbor, MI, USA) *(SIGIR '18)*. Association for Computing Machinery, New York, NY, USA, 989–992. https://doi.org/10.1145/3209978.3210124
[60] Anabel Quan-Haase, Carly Williams, Maria Kicevski, Isioma Elueze, and Barry Wellman. 2018. Dividing the grey divide: Deconstructing myths about older adults' online activities, skills, and attitudes. *American Behavioral Scientist* 62, 9 (2018), 1207–1228.
[61] Mary Elizabeth Raven and Alicia Flanders. 1996. Using contextual inquiry to learn about your audiences. *ACM SIGDOC Asterisk Journal of Computer Documentation* 20, 1 (1996), 1–13.
[62] Jorge Ribeiro and Ana Correia de Barros. 2014. Efficiency of a video and a tutorial in teaching older adults to interact with smartphones. In *Universal Access in Human-Computer Interaction. Aging and Assistive Environments: 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part III 8*. Springer, 34–45.
[63] Wendy A Rogers, Elizabeth F Cabrera, Neff Walker, D Kristen Gilbert, and Arthur D Fisk. 1996. A survey of automatic teller machine usage across the adult life span. *Human factors* 38, 1 (1996), 156–166.
[64] Daisuke Sato, Masatomo Kobayashi, Hironobu Takagi, Chieko Asakawa, and Jiro Tanaka. 2011. How voice augmentation supports elderly web users. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. 155–162.
[65] Sergio Sayago, Barbara Barbosa Neves, and Benjamin R Cowan. 2019. Voice assistants and older people: some open issues. In *Proceedings of the 1st International Conference on Conversational User Interfaces*. 1–3.
[66] Gianluca Schiavo, Ornella Mich, Michela Ferron, and Nadia Mana. 2017. Mobile multimodal interaction for older and younger users: exploring differences and similarities. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. 407–414.
[67] Gianluca Schiavo, Ornella Mich, Michela Ferron, and Nadia Mana. 2022. Trade-offs in the design of multimodal interaction for older adults. *Behaviour & Information Technology* 41, 5 (2022), 1035–1051.
[68] Stephan Schlögl, Markus Garschall, and Manfred Tscheligi. 2014. Designing natural language user interfaces with elderly users. In *Workshop on Designing Speech and Language Interactions*.
[69] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I Hong. [n. d.]. "Hey Alexa, what's up?": studies of in-home conversational agent usage,". In *Proceedings of the DIS*.
[70] Shradha Shalini, Trevor Levins, Erin L Robinson, Kari Lane, Geunhye Park, and Marjorie Skubic. 2019. Development and comparison of customized voice-assistant systems for independent living older adults. In *Human Aspects of IT for the Aged Population. Social Media, Games and Assistive Environments: 5th International Conference, ITAP 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part II 21*. Springer, 464–479.
[71] Khe Chai Sim. 2010. Haptic voice recognition: Augmenting speech modality with touch events for efficient speech recognition. In *2010 IEEE spoken language technology workshop*. IEEE, 73–78.
[72] AARON SMITH. [n. d.]. Mobile Access 2010. https://www.pewresearch.org/internet/2010/07/07/mobile-access-2010/.
[73] Hironobu Takagi, Masaki Ohno, Masatomo Kobayashi, and Takeo Nakada. 2018. Evaluating speech-based question–answer interactions for elder-care services. *IBM Journal of Research and Development* 62, 1 (2018), 6–1.
[74] Gemini Team. 2023. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805* (2023). https://arxiv.org/abs/2312.11805
[75] Carla Tubin, João Pedro M Rodriguez, and Ana Carolina B De Marchi. 2021. Voice and touch interaction: A user experience comparison of elderly people in smartphones. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*. 1–7.

[76] M Turk. 2013. Multimodal interaction: A review. Pattern Recognition Lett. (2013).
[77] Markku Turunen, Aleksi Kallinen, Iván Sánchez, Jukka Riekki, Juho Hella, Thomas Olsson, Aleksi Melto, Juha-Pekka Rajaniemi, Jaakko Hakulinen, Erno Mäkinen, et al. 2009. Multimodal interaction with speech and physical touch interface in a media center application. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*. 19–26.
[78] Eleftheria Vaportzis, Maria Giatsi Clausen, and Alan J Gow. 2017. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. *Frontiers in psychology* 8 (2017), 294305.
[79] Ravichander Vipperla, Steve Renals, and Joe Frankel. 2008. Longitudinal study of ASR performance on ageing voices. (2008).
[80] Christopher B. Mayhorn W. Morrell, Denise C. Park. 2000. Effects of age and instructions on teaching older adults to use Eldercomm, an electronic bulletin board system. *Educational Gerontology* 26, 3 (2000), 221–235.
[81] Zhiyuan Wan, Lingfeng Bao, Debin Gao, Eran Toch, Xin Xia, Tamir Mendel, and David Lo. 2019. Appmod: Helping older adults manage mobile security with online social help. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 4 (2019), 1–22.
[82] Cheng-Yao Wang, Wei-Chen Chu, Hou-Ren Chen, Chun-Yen Hsu, and Mike Y. Chen. 2014. EverTutor: Automatically Creating Interactive Guided Tutorials on Smartphones by User Demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 4027–4036. https://doi.org/10.1145/2556288.2557407
[83] Jiancan Wu, Xiangnan He, Xiang Wang, Qifan Wang, Weijian Chen, Jianxun Lian, and Xing Xie. 2022. Graph convolution machine for context-aware recommender system. *Frontiers of Computer Science* 16, 6 (2022), 166614.
[84] Bo Xie, Tom Yeh, Greg Walsh, Ivan Watkins, and Man Huang. 2012. Co-Designing an e-Health Tutorial for Older Adults. In *Proceedings of the 2012 IConference* (Toronto, Ontario, Canada) *(iConference '12)*. Association for Computing Machinery, New York, NY, USA, 240–247. https://doi.org/10.1145/2132176.2132207
[85] Qian Yu, Tonya Nguyen, Soravis Prakkamakul, and Niloufar Salehi. 2019. " I almost fell in love with a machine" speaking with computers affects self-disclosure. In *Extended Abstracts of the 2019 CHI conference on human factors in computing systems*. 1–6.
[86] Panayiotis Zaphiris, Sri Kurniawan, and Mariya Ghiawadwala. 2007. A systematic approach to the development of research-based web design guidelines for older people. *Universal Access in the information Society* 6 (2007), 59–75.
[87] Maozheng Zhao, Henry Huang, Zhi Li, Rui Liu, Wenzhe Cui, Kajal Toshniwal, Ananya Goel, Andrew Wang, Xia Zhao, Sina Rashidian, et al. 2022. Eyesaycorrect: Eye gaze and voice based hands-free text correction for mobile devices. In *27th International Conference on Intelligent User Interfaces*. 470–482.
[88] Mingyuan Zhong, Gang Li, Peggy Chi, and Yang Li. 2021. HelpViz: Automatic Generation of Contextual Visual Mobile Tutorials from Text-Based Instructions. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 1144–1153.
[89] Ziyao Zhou, Jia Zhou, and Fengli Liu. 2022. Fewer steps the better? Instructing older adults' learning and searching in smartphone apps. *International Journal of Human–Computer Interaction* 38, 9 (2022), 789–800.
[90] Randall Ziman and Greg Walsh. 2018. Factors affecting seniors' perceptions of voice-enabled user interfaces. In *Extended abstracts of the 2018 CHI conference on human factors in computing systems*. 1–6.

A PROMPT EXAMPLES

*Your task is to modify the query A spoken by the older user when using the phone, {**Context Information Description**}.*

Firstly, before making any modifications, check if query A contains the keywords listed in the following [Keyword List]. If it does, ensure to retain these keywords in the modified result. The noteworthy keywords are listed in the [Keyword List] below, separated by commas.

Next, while keeping the keywords from query A intact, modify the statement to be more concise, clear, and understandable. Replace overly specific details mentioned in A with generic references, such as specific locations, names, or product names. This will help readers understand which phone function the older user wants to use.

*Keyword list = [{**list of keywords**}]*

Fig. 5. The prompt used for query formatting.

Your task is to receive a query from elderly users while using their mobile phones, and based on understanding the context, predict the most likely function the user may intend to use or inquire about, outputting the function with the highest probability.

The query may includes the user's voice text and contextual information such as the app opened, the page the user is on, and the page components (such as a button or product) clicked by the user when asking the question. Therefore, combine the user's contextual information to understand the user's query in the question.

The context contains the names of different app functions and their explanations, separated by colons. The function name is before the colon, and the explanation of that function is after the colon.

Provide the function with the highest probability in the Answer format. The result should only contain the function name and the app where the function is located, separated by a semicolon.

Please note: Only answer based on the content of the query and context, and do not provide unrelated responses.

*Query: {**query**}*

*Context: {**context**}*

Answer: function name; app

Fig. 6. The prompt used to search for tutorials.

B VERBAL QUESTIONING PATTERNS

Table 8. The patterns of verbal questions spoken by older participants when seeking help for smartphone problems. *Problem* is the smartphone issue that the participant wants to solve; Red text in *Query* is the participant's original question, and the black text following it provides additional description.

Pattern	Detail	Problem	Query	Explanation
Incomplete information	Missing location information	Making a payment at a certain store	"I want to pay" (Missing payment location, some stores only support specific payment methods.)	Older adults have **weaker memory** and **get easily anxious** when encountering problems, which often results in forgetting to provide contextual information.
	Missing interface information	Deleting a product from the cart in an app	"I want to delete it" (Missing interface where the product is located.)	
	Missing app	Changing user profile picture in an app	"I want to change my profile picture" (Missing information about which app's user profile picture needs to be modified.)	
	Missing UI components	Collecting virtual coins in an app	"I want to collect this" (Missing target object, older participants use content-pointing actions to target instead of verbal description.)	Older adults have the **lower cognitive ability** and may not understand how visual patterns on the interface convey information.
Redundant information	Descriptions of why things happen	Taking a video and sharing it with friends	"I'm in a park and the view is very beautiful and I want my friends to see this too and I want to take videos and send them to do how to do it"	Older adults are **overly concerned** that others may not understand, so they tend to provide more information about the reasons behind what happened.
	Repetition of Interface Information	Copy a piece of text	"I want to copy 'next Friday to go hiking together and meet at the gate at 1 pm'" (Reading out all the text to copy directly)	Older adults have the **lower cognitive ability** and may not understand how components on the interface convey information.
Ambiguous descriptions	Description of visual information through describing visual style and intended operations	Feeling confused about an icon and wanting to remove it	"Sometimes there is a kind of icon on the phone screen, I don't know how to use it, I don't understand what you mean, I want to remove it, how to operate it"	Compared to younger individuals, older adults are typically **less tech-savvy** and may not know the proprietary names of visual components and patterns on the interface.
Overly specific vocabulary	Use specific vocabulary that is not outlined as specialized vocabulary	Searching for a singing blogger	"I want to listen to Wang Erni's songs" (Using the specific blogger name)	Older adults are not familiar with search techniques and have relatively **low search literacy**. They prefer to use specific nouns to express their intentions, instead of generalizing specific nouns into professional terms.
Wrong description of vocabulary	Misunderstanding of vocabulary	Downloading videos	"I want to collect videos" (Some older adults think that downloading and collecting are the same.)	Older adults have the **lower cognitive ability** and are not familiar with the functions of the mobile phone.