IRemember: Memorable CAPTCHA Method for Sighted and Visually Impaired Users

Mrim Alnfiai, Sahar Altalhi, Duaa Alawfi
Information Technology Department
College of Computer Science and Information Technology
Taif University, Taif, Saudi Arabia

***Abstract*—A CAPTCHA is used to automatically differentiate between human users and automated software to prevent bots from accessing unauthorized websites. Most proposed CAPTCHAs are not accessible to visually impaired users because of the memorability of the CAPTCHA's numerical digits. Recalling six random spoken digits is a difficult task for any human. Visually impaired users must typically play the audio several times to memorize the spoken digits in the correct order. The authors reviewed existing CAPTCHAs for visually impaired users and concluded that the high cognitive load is more susceptible to response errors due to extensive challenge digits intended for visual users. Thus, the authors proposed a novel method that improves current audio CAPTCHA by enhancing the display of the challenge and improving the memorability of its phraseology. The proposed CAPTCHA presents short common phrases, such as "piece of cake." After hearing or seeing the phrases, the users are required to type the first letter of each word from the presented phrases, such as POC for a piece of cake. The study results of 11 visually impaired users concluded that the memorability and success rate for the IRemember CAPTCHA was 82.72%, compared to the audio CAPTCHA at only 48.18%. It has also demonstrated higher memorability and less workload than the traditional audio method. This research indicates that using common knowledge and experience in the design process for a CAPTCHA method for these users can enhance performance and minimize workload and, hence, error rates.**

***Keywords*—CAPTCHA; blind users; visually impaired users; memorability; accessibility**

I. Introduction

Today, we can consider web services as the backbone of life, especially since the inception of the COVID-19 pandemic. Many services are provided through web-based applications. The Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) is used as an authentication protocol. It has become a necessary step for online services, especially ones designed for public use, free usage cost, or registration to the end user. If you use the internet regularly and deal with a web form, you will typically encounter a CAPTCHA test. Being able to utilize the services on these applications is a right for everyone worldwide. A study by the World Health Organization shows that about 217 million people have some kind of visual impairment, and almost 36 million are totally blind [1]. It is also determined that the rate of people with visual impairments will increase because of the population expansion. CAPTCHAs are annoying for humans to solve and can be difficult or impossible to figure out for those with disabilities. CAPTCHAs have to work for all people, disabled or not, because this tool prevents bots and viruses, but people must be able to use them with ease. Therefore, it is essential to develop new technologies that people with visual impairment can effectively use.

CAPTCHA, initially developed in 1997 by Alta Vista, is a test used to determine whether or not the user is human. It is a significant step when dealing with a web form to prevent automatic bots and harmful invasions. It does this by offering tests that humans can pass, but computer programs cannot. The term CAPTCHA was instituted in 2000 by Luis von Ahn, Nicholas J. Hopper and Manuel Blum of Carnegie Mellon University and John Langford of IBM [2]. Essentially, CAPTCHA blocks robot software from submitting fake or misused online requests. It is also used to protect the integrity of online polls by preventing hackers from sending in repeated false responses using robots.

Currently, CAPTCHA is categorized into five main categories: text-based, image-based, video-based, puzzle-based, and audio-based [20], the text-based type being the most widely deployed. Google, Yahoo, and Microsoft websites have deployed text-based CAPTCHAs for years. Many techniques are used for creating text-based CAPTCHAs, such as Gimpy, which provides an arbitrary number of words and displays them in a distorted manner. The EZ-Gimpy technique uses only one word. The Gimpy-r technique uses random distorted letters with a noise background. Simard's HIP CAPTCHA selects random letters and numbers and then uses arcs and colors to distort them [3]. The problem is that text-based CAPTCHAs are not suitable for individuals with visual impairments since they have a hard time remembering letters and typing them.

Image-based CAPTCHAs require users to select matching images or images that don't fit. A Braille display cannot present an image; therefore, a user who needs to use a Braille display cannot solve the challenge. So, if screen reader software could display the CAPTCHA to a blind user that would defeat the purpose of preventing automatic bots and harmful actions since bots would be able to solve the challenge as well. This is a severe limitation of image-based CAPTCHAs, which cannot be used wherever accessibility has to be guaranteed by governmental institutions.

The third category is video-based CAPTCHAs, which rely on typing information from the video that only humans can notice.

Puzzle-based CAPTCHAs have pictures divided into parts with different techniques and ask the user to find the missing part.

Audio-based CAPTCHA is the last category, and that presents an audio recording with a noisy background of a series of letters or numbers [14]. This CAPTCHA helps visually impaired users, but it is difficult to interpret the results. Many studies and schemes have been done in this area to develop it. One of these studies did try to help blind people by not making them type the word. This mechanism is called HearAct CAPTCHA, and it uses a tap or swipe to allow the user to determine if it is a specific letter in a sound-maker name. Users with comprehensive spelling and English vocabulary can solve this challenge. So, remembering how to spell the word is the gist of this type of CAPTCHA [1]. reCAPGen schemes depend on choosing audio clips from old radio programs, podcasts, and YouTube lectures and adding the amount of noise. Noising minimally affects the human ability to solve the generated audio CAPTCHAs [4].

The Last Two Words (LTW) scheme achieves a success rate of 78 percent with sighted users and 81 percent with visually impaired users with an average response time of 15 seconds [5]. These CAPTCHAs exploit the human effort to generate transcriptions for audio files with high accuracy [6]. This scheme relies on remembering the last two words and knowing how to pronounce them, which is difficult for non-native English speakers.

This paper introduces a new technical solution to address visually impaired individuals' challenges in answering CAPTCHA questions. This solution allows browsing and accessing internet content while maintaining website security. Using an approach to help users recall the response to the CAPTCHA question improves the accessibility. It makes solving the CAPTCHA challenge the idea behind audio, improves accessibility, and allows people with visual impairments to solve the CAPTCHA independently. This approach uses common memorable phrases to enable users to memorize the answer quickly without repeating it several times. In this technique, we will use these sentences in the audio CAPTCHA and ask the user to input the first letter of each word. For example, if we use the common phrase, "How are you?" they are required to input the first letter of each word, in this case, "hay." We will display common phrases to enhance the memorability of the CAPTCHA. The length of the phrases must not exceed four words. Developing this new form of CAPTCHA allows visually impaired users to memorize it.

The objectives of the proposed CAPTCHA method are:

1) To overcome the limitations associated with traditional audio CAPTCHA methods for visually impaired users.

2) To build accessible and secure CAPTCHA challenges for both visually impaired and sighted users.

3) To improve memorability for visually impaired users by reducing the mental workload needed to solve the CAPTCHA and minimizing their cognitive load needed to recall the CAPTCHA digit.

4) To minimize error rate and completion time to solve the challenge.

The contribution of this research is as follows:

- Creating a novel and memorable CAPTCHA method that uses common phrases in creating CAPTCHA challenges, allowing users to respond with few characters to minimize completion time.
- Developing a CAPTCHA method that is memorable and easy to solve for visually impaired individuals.
- Presenting a user study to investigate the performance of the proposed CAPTCHA and measure the accuracy, usability, and cognitive load required to complete the task. Participants solved a set of CAPTCHA challenges during the experiment.

The paper is organized as follows: a review of the CAPTCHA systems previously designed for visually impaired people, an explanation of the IRemember CAPTCHA design, the methodological approach to testing the proposed CAPTCHA method, results, conclusions, and considerations for future research.

II. Related Work

With the spread of the Internet and its technologies, it has become necessary to propose and develop online verification methods, CAPTCHA, and facilitate it to all society members in proportion to their special needs, such as people with autism spectrum disorders [7], learning disabilities [8], and blinds or visual impairment [9]. For Internet users with a visual disability, our target in this research, the audio CAPTCHA overcomes limitations of other CAPTCHA types because it is based on what they hear, not what they see [1].

Human-Interaction Proof, Universally Usable (HIPUU) with both versions [10] and [11] is an audio CAPTCHA designed to overcome the traditional audio CAPTCHA limitations. The task is based on choosing the appropriate word to describe an image from a drop-down list. In addition, an alternative for the image is provided by playing an audio file corresponding to the same image. In this way, the user can choose the preferred media representing CAPTCHA. In version 1, the list contains 15 different choices and some false decoy answers, which makes it vulnerable to brute force attacks because of the small number of choices. However, HIPUU version 2 increases the security level by requiring solving multiple tasks, each with 35 choices and some false decoy answers. In contrast, version 3.0 requires users to type in the solution using a keyboard instead of choosing from the drop-down menu, which is a time-consuming task for visually impaired users. Similarly, SoundsRight [12] is another audio-based CAPTCHA that asks the user to press the space button each time the users hear a particular sound.

In [13], the CAPTCHA is made of a simple text-mathematical problem converted into speech/audio using text to speech (TTS) system. The user has to listen to the question and then answer by only typing the answer as a number. On

the one hand, this method relies on computer limitations to solve numerical spoken questions, but it is challenging to generate a large number of tests. Thus, it is not a practical solution [14]. Similarly, in [15], the authors propose four scenarios of audio CAPTCHA includes: ask users to calculate a running total, count the occurrence of a character in an alphanumeric series, transcribe the alphanumeric characters they heard, and last which is the categories prototype that asks users to count the number of sounds, in a series, that belonged to a specific category. Three out of four designs have increased vulnerability against random guessing attacks, which makes it limited to situations where high security is not critical. HuMan [16] asks the user a question from his/her favorite field and types the answer in a text box. It is easier than the text or image-based CAPTCHAs for blind users, but still has a high probability of spelling mistakes since it requires a full written answer.

In 2021, Mathai et al. improved the existing audio CAPTCHA as they stated that the traditional audio CAPTCHA is complicated to solve by visually impaired users. In their developed method, users hear a distinctive sound, and they are required to count how many times they hear the intended sound. The distinctive sound is a combination of sound, background noise, and music produced using Generative Adversarial Networks to make it difficult for bots to distinguish the answer. The main advantage of this method is allowing users to solve the challenges using numbers. However, it is too complicated and time-consuming to prepare the sound challenges.

Another distinct study where the CAPTCHA method is developed using an OTP-based QR code [17]. Users can decrypt the OTP using a unique key. It is a secure method, but it requires having another phone device to scan the QR code. Noorjahan designed a CAPTCHA method using fingerprints (2019). This method is accessible for visually impaired users, but it requires external hardware to complete the CAPTCHA solving task, which is a fingerprint scanner. Another drawback of this system is that it is applicable to smartphone devices.

HearAct [18] is an audio-based solution that does not require typing text but reaction. The user listens to a sound of what is called a "sound-maker" and answers a spoken question by tapping if the answer is true. Otherwise, swap left or right. The work achieved a success rate of about 82.05%, which is twice as much as the success rate of the traditional audio CAPTCHA and a faster completion time of 55.35 seconds than 65.64 seconds in the traditional audio CAPTCHA.

A recent and related work [19] proposes Machine Learning (ML) model to classify humans from bot behavior. To solve the CAPTCHA task, the user will be asked by audio to draw a simple shape. Then, the cursor movement will be captured and converted to an image and sent to the backend ML model to decide if it was a human or not.

The Generative Adversarial Network (GAN) algorithm is used in [20] to generate music with unique sound samples and a white noise layer. The CAPTCHA challenge asks the user to count the number of unique sounds made by percussion instruments such as drums. This solution simplifies the task for blind or impaired users by requesting to type the answer as a number. In addition, auto-generation of music and adding the noise layer make it more secure against speech-to-text bots and APIs.

The work [21] is a simple but efficient solution developed to allow visually impaired users to verify their humanity by walking at least five steps. If this step is done successfully, then the user will proceed to enter a username and password. Otherwise, the verification process will fail, and the user will be prevented from accessing the service that he/she wants.

By reviewing these studies and taking in mind the few numbers of research on audio CAPTCHAs and their accessibility and usability for people with vision impairments, the door is opened for more investigation and encouraged more studies to improve the accessibility, memorability, usability, and speed of solving CAPTCHA tasks for blind Internet users. This motivated us to propose a new solution based on commonly spoken phrases to increase memorability and requires entering only the first letter from each word to speed the task for blind or visual impairment users.

III. IREMEMBER CAPTCHA DESIGN

The proposed method is audio and text-based, where users can see or listen to a short common phrase and then identify the beginning letter of each word in the phrases displayed and spoken. The proposed method is built based on the theoretical concept that people can remember phrases better than random and varied letters [22]. The IRemember CAPTCHA features include:

1) Requiring users to listen to the provided phrase.

2) Recognizing the beginning letter of each word in the phrase.

3) Typing only the first letter of each word in the response field.

4) Ability for user to double tap when updating the CAPTCHA challenge.

For example, users will receive a question like, "Please type the first letter in each word in the following phrase." Then, users are required to listen to the given phrase, "How are you?" and to determine the first letter of each word within the phrase. Once users recognize the beginning letters of the phrase, they must type "hay." If users want to update the challenge, they can double tap anywhere on the screen to receive a new challenge (see Fig. 1).

The proposed CAPTCHA differs from existing methods in two ways: the identification stage and solving stage. In the identification stage, users listen to the phrase and are required to recognize the first letter of each word in the phrase. In the solving phase, users can type 3 to four letters that represent the first letter in each word. Doing so will minimize the time spent to solve the CAPTCHA as users have only to type three digits.

To improve the accessibility of CAPTCHA, two sounds are used. The first speaks the CAPTCHA question, "Type the first letter of each word in the phrase," and the second speaks the phrases. This feature can enhance the method's usability and simplify the differentiation between the question and challenge. To minimize cognitive load, the CAPTCHA

method has one basic CAPTCHA statement, which is "Type the first letter of each word in the CAPTCHA challenge."

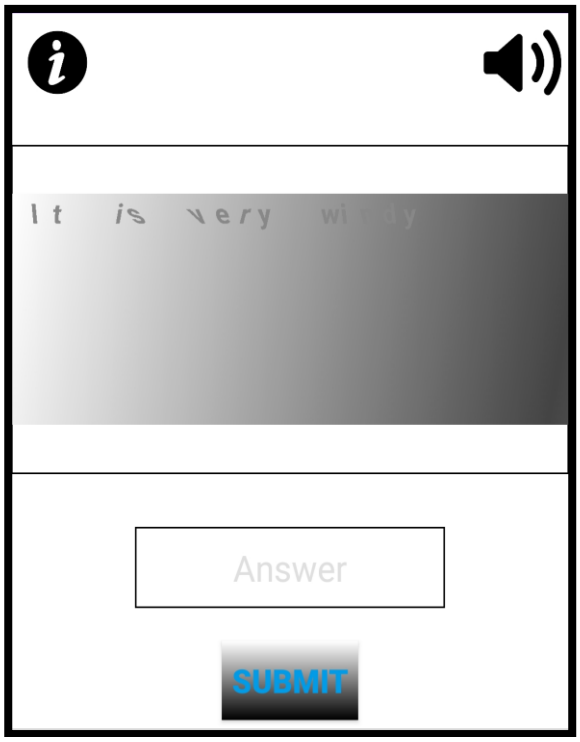

Fig. 1. IRemember CAPTCHA Method Interface.

The proposed approach is different from existing methods in both the identification and solving phases. In the initial phase, the challenges are represented as a phrase, which is easier to recall than presenting different letters. In the second phase, users are required only to type a few letters. To enhance security and reduce confusing recognition software, the CAPTCHA method presents the entire phrase that is distorted using text and audio. With this CAPTCHA style, it is easier for users to identify the first letter of each word of the phrase while simultaneously being difficult for the software to determine the letters.

A more secure and accessible CAPTCHA method is provided by developing the proposed CAPTCHA method. The new method will also benefit a wide range of users as it will simultaneously provide visual and audible challenges. Doing so makes CAPTCHA universal and meets usability and accessibility standards. Another valuable benefit of developing IRemember CAPTCHA is creating an accessible and memorable CAPTCHA method that will enable visually impaired individuals to solve CAPTCHA challenges by themselves without asking for assistance from others to pass this stage. By doing so, the visually impaired will be independent and able to access online materials without support. Another critical benefit of this CAPTCHA is minimizing the time spent to complete the challenge and reducing the cognitive load needed to recall the CAPTCHA digits. In addition, developing an easy-to-recall challenge will minimize the error rate and enhance users' satisfaction with the developed method.

The proposed method will be easier for sighted users to solve as they are only required to type the first letter of each word, which is much easier to recognize and type since its only six digits. From the website security perspective, the proposed method will be more secure as it will use different background noise to prevent software recognition and determine the challenge response. As a result, developing this method will provide websites with a more resistant and secure approach to various attacks. It is more confusing for software to recognize the content of a challenge, which is the entire phrase. Additionally, providing different methods for presenting CAPTCHA challenges will improve accessibility for humans as it considers people's disabilities. In general, building accessible CAPTCHA for visually impaired users and other individuals with disabilities prevents the technical barriers they have faced when browsing web pages and enables them to do their online tasks without complications.

A. Selection of Phrases

The authors gathered a set of common phrases or idioms that individuals typically use in their daily lives that are recognizable by general users without requiring too much concentration and mental effort to memorize the given digits. The used phrases were collected from the EnglishSpeak website [22]. In the implementation phase, the authors used 250 common English phrases to build the first version prototype's library.

B. IRemember CAPTCHA Task Complexity

The primary goal of the IRemember CAPTCHA is to create an accessible and memorable method for individuals with no or low vision. The method presents a phrase that contains at least three words to enhance the complexity of the audio CAPTCHA challenge and, thus, the security. In addition, background noise was added to the text and audio clip to make it more challenging for software to recognize the CAPTCHA content.

C. Implementation of IRemember CAPTCHA

The proposed CAPTCHA is implemented on the Android platform. The authors chose to implement the prototype on the Android platform because it is open source and allows designers to integrate available functions with future iterations. In addition, the proposed CAPTCHA can be implemented on other platforms.

IRemember has one interface that presents the challenges of using textual and audible approaches. At the bottom of the challenge, a large text box appears where users can type the requested response. As a result, it is easier for users to locate the answer box, and they can immediately type the response upon hearing the challenge. For example, the users hear the prompt, "Type the first letter of each word in the phrase" with a particular pitch of a sound. Then, the application will immediately present the phrase "Cup of Joe" on a different pitch. After users hear the phrase, they should recognize the first letter of each word (COJ) and type those letters into the response box. If the response is correct, users pass it onto the website's main page; if it is wrong, users receive a new CAPTCHA challenge.

Fig. 1 shows an example of an IRemember CAPTCHA as described above. Then, if users enter an incorrect answer, the CAPTCHA says, "Try again," and users receive a new challenge. Users follow the same steps to solve the additional challenge.

The key features of the audio CAPTCHA file used to implement IRemember CAPTCHA are:

- Common phrases.
- Phrases ranging between three to four words.
- Background noise.
- Challenge duration differs based on phrase length.
- Randomly presented phrase.

D. IRemember CAPTCHA Security

The IRemember CAPTCHA method provides another level of layer after the authentication process for users. It has been implemented to build an accessible approach for people with no vision to enable them to solve the given challenge in a short time with less effort.

In the implementation process of the proposed CAPTCHA, we considered the essential security requirements that make the CAPTCHA method more secure and eligible [23, 24]. The IRemember CAPTCHA is implemented based on these requirements, including distortion levels, challenge type, randomness, time constraints, and size. Most blind users find it extremely difficult to solve a CAPTCHA following a loud background noise or distortion. Thus, according to the HearAct, CAPTCHA challenge type and size are the predominant design factors to consider and have been used to make the IRemember solution secure and accessible. Users receive the whole phrase and need to recognize the first letter of each word in the phrase; doing this makes it very difficult and costly for an automated bot to attack. The addition of background noise makes it complicated and requires time from software bots to provide the answer. Randomness is another IRemember method presented randomly to increase the hardness of distracting the correct response. In addition, time is another approach to restrict the software bots from recognizing the answer. After a short time, the method will generate a new challenge. Adding a specified time frame for answering the CAPTCHA challenge prevents the bots from having time to identify the right answer. Another factor that enhances the security level of a CAPTCHA method is the size and length of the CAPTCHA challenge. The IRemember CAPTCHA method provides phrases that have at least three or four words.

The steps that are required to break or automatically solve the IRemember CAPTCHA method would include the following procedures:

1) Understand the main question (e.g., "Type the first letter of each word in the phrase?"

2) Separate or remove the random noise from the audio content.

3) Convert the spoken phrase into textual format.

4) Extract concepts from the transcribed text and recognize the first letter of each word from the given phrase.

5) Type the answer by analyzing the challenge and specify the response.

Theoretically, the software bots need to transcribe the audio into a textual format in the traditional audio CAPTCHA. However, the software bots need more than three steps to break the IRemember CAPTCHA: removing the background noise, understanding the main question, and determining the first letter of each word in the phrase. These steps make it more challenging to break the CAPTCHA method and require more time.

E. Advantages of IRemember

The proposed CAPTCHA system is easy to solve and reduces the time users will spend typing by only asking for the first letter of each word. This CAPTCHA is unlike the traditional audio CAPTCHA, where users need to type at least six digits for each challenge [25]. Additionally, this CAPTCHA approach is efficient because once users hear the audio challenge and identify the CAPTCHA, they can immediately solve the challenge without any need to hear the audio clip in its entirety. When users identify the phrase, they can type the response and pass the challenge seamlessly.

The most important advantage of the proposed CAPTCHA is that it is simple to recall the response phrases because they are commonly used in everyday life. So, memorizing a phrase is far easier for users than a string of random digits.

IV. Methodology

This section aims to investigate the usability and memorability of the IRemember CAPTCHA method for visually impaired users. Our methodology focuses on mixed methods to perform measurements. The quantitative method was used to measure solving time and success rate and to clearly determine user experience through a workload (NASA TLX) Subjective Questionnaire. On the other hand, the qualitative method will be used to collect strengths, limitations, and suggestions through interview questions. The performance measures were used to analyze the efficiency and effectiveness of the examined CAPTCHAs. This section explains the user study procedures and the obtained results in detail.

A. Study Goals and Hypotheses

The primary goal of the proposed IRemember CAPTCHA method is to improve memorability as well as maximize the success rate. The main goal of the user study is to analyze the usability and memorability of the proposed method. In particular, we compared our proposed method with the traditional audio CAPTCHA method taken from [25]. The audio CAPTCHA method is an alternative to text-based CAPTCHA, where it speaks out the given digits, and the users are required to memorize it and type it using the keyboard on the answering box (see Fig. 2).

Fig. 2. Common Audio CAPTCHA Method.

The user study hypotheses are:

1) Completing time: The IRemember CAPTCHA can be solved faster than the traditional CAPTCHA due to the easiness of remembering the given challenges and the few numbers of digits that need to bed entered. The IRemember CAPTCHA removes the need to repeat the CAPTCHA challenges several times to memorize the digits limitations of the touchscreen entry method, including locating the location of keys on a touchscreen keyboard.

2) Success rate: The IRemember CAPTCHA will be solved correctly more than the traditional audio CAPTCHA due to the memorability of the given challenges.

3) Workload: The IRemember CAPTCHA will require less workload to complete solving the challenge than the traditional audio CAPTCHA.

B. Participants

To evaluate the proposed CAPTCHA and audio CAPTCHA, eleven participants (six female and five male) were recruited from the Taibh University Accessibility Center and from the Center for Disabled People at Taibh. Participants' ages ranged from 18 to 33 years old and participated in the study voluntarily. All participants were completely blind, have experience using touch screen devices with the support of a screen reader, and were proficient in the English language.

C. Procedure

The study aimed to examine the usability and memorability of two CAPTCHA methods, including IRemember CAPTCHA and the traditional audio CAPTCHA. At the beginning of the experiment, researchers explained the study goals and tasks and described the functionality of both examined CAPTCHAs. They also introduced an overview of the study requirements and study steps and answered participants' queries and provided a clear overview of the study procedures and requirements. After answering participants' questions, they signed a consent form that disclosed that their interaction with the smartphone device while solving CAPTCHA challenges is recorded and their private information is protected. The study contains four main sessions, which are a demographic questionnaire, a training session, a test session, and a Workload Questionnaire.

1) A Demographic questionnaire: To have a clear overview of user experience and performance when using the proposed CAPTCHA method, we collected participants' demographic data, including age, gender, vision level, and touchscreen experience, and English language proficiency.

2) Training session: After collecting participants' demographic data, researchers trained participants on how to solve challenges in both CAPTCHA methods. They were also encouraged to explore both methods and use them for five minutes to ensure they were familiar with the CAPTCHA methods' interfaces. The data from the training session were not included in the study results.

3) Test session: Participants were required to answer 10 CAPTCHA challenges without any assistance or without using other tools to help them record the response to a CAPTCHA question. To avoid order effect, CAPTCHA methods were randomly presented with a 2*2 Latin Square. Challenges for each CAPTCHA method were presented randomly. After listening to a CAPTCHA challenge, participants were requested to speak out the CAPTCHA answer before typing it in order to determine the causes of error when solving a CAPTCHA challenge, whether due to the input method or the difficulty of recalling the given digits. Then, they can type their response. At the end of the test session, participants were required to answer the questionnaires orally.

4) Questionnaires: To measure workload perception [26] for CAPTCHA methods, participants completed the workload questionnaire (NASA TLX) at the end of the study (see Table II). Participants were also asked to identify the strengths and limitations of CAPTCHA methods and recommend any improvements for the memorability aspect. This session took five to ten minutes.

5) Apparatus: We developed a prototype of IRememeber CAPTCHA method on an Android device called Galaxy Nexus phone that has a 4.65-inch-long screen. In the experiment, we asked participants to use the prototype on a portrait mode when solving challenges. In addition, participants were screen recorded when solving CAPTCHA challenges to track the method performance and participants' behavior when interacting with the device. At the end of the experiment, participants were audio recorded to report their answers to the questionnaires.

D. Usability Measurements

Success rate, completing time and the subjective NASA TLX were measured to evaluate the usability and memorability of both CAPTCHA methods.

1) Success rate: The effectiveness (success rate) of the IReamamber CAPTCHA and audio CAPTCHA were evaluated by calculating the challenge completion rate and the number of errors. The total number of challenges completed successfully were divided by the total number of challenges to calculate the CAPTCHA success rate (see Eq. 1).

$$Success\ Rate = \frac{Challenges\ completion\ Successfully}{Total\ number\ of\ Challenges} \quad (1)$$

2) Completing time: In order to determine the average solving time taken to solve CAPTCHA challenge for each participant, each CAPTCHA challenge was tracked from the start time (the time when Run button was pressed) and and the

completing time (the time when Submit button was pressed). The difference between the completing time and start time determine the solving time, which involves the time spent listening to the challenge and the time spent typing a response.

3) Subjective measure: The Workload NASA TLX questionnaire was used at the end of the study, it contains six factors which are mental demand, physical demand, temporal demand, performance, effort, and frustration (see Table II).

4) Memorability measure: To determine whether participants remember the CAPTCHA content, we asked them to speak out the answer before typing it. In the IRemember CAPTCHA, users cannot repeat listening to the challenge, but in the audio CAPTCHA method, users can repeat several times and then they can speak out their responses.

V. Result

The aim of the conducted user study is to evaluate the usability and memorability of the proposed CAPTCHA IRemember. The success rate, speed, user satisfaction, and memorability were computed and discussed in detail in this section.

A. Success Rate

Fig. 3 shows the average success rate of each participant for each CAPTCHA method. The average success rate of the IRemember CAPTCHA method is 81%. Whereas the average success rate for audio CAPTCHA is 48% (see Table I). The results indicate that the success rate for IRemember CAPTCHA is higher and more efficient than the traditional audio CAPTCHA. The results also disclosed that users tend to make more typing errors when using the traditional audio CAPTCHA. Fig. 3 also shows that all participants perform better when solving CAPTCHA challenges using the IRemember CAPTCHA method, which may contribute to the easier recalling of a common phrase than random digits.

One way ANOVA value shows a significant difference between IRemember and the traditional audio CAPTCHA ($F(1, 21) = 78.47$; $p < 0.000476$). The t-test value indicates a significant difference between the success rate for IRemember and the traditional audio CAPTCHA ($p= 0.000476$).

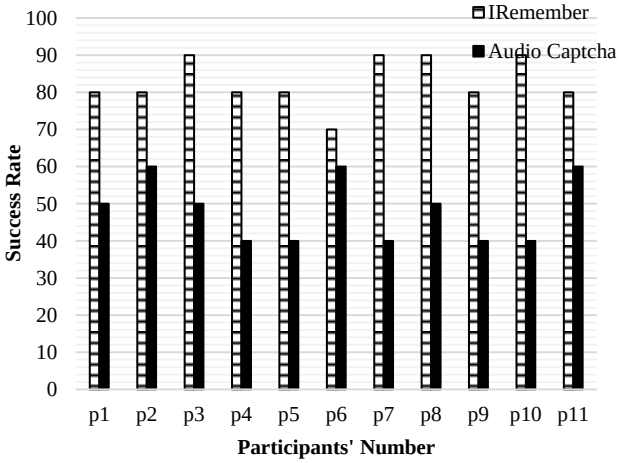

Fig. 3. Average Success Rates for IRemember and Audio CAPTCHA.

TABLE I. Summary of the Results of Both CAPTCHA Methods for Visually Impaired Users

Methods	Average success rate (standard)	Average solving time (standard)
Audio CAPTCHA (6 digits)	48.18% (8.73)	138.75 s
IRemember CAPTCHA	82.72% (19.20)	54.96 s

B. Completing Time

Fig. 4 shows the average completing time each participant spent on solving IRemember and Audio CAPTCHAs. 54.36 seconds is the average time participants spent Completing IRemember CAPTCHA, and 137.5 seconds for the traditional audio CAPTCHA. The results reveal that all participants spent more time completing the traditional audio CAPTCHA method than IRemember method. That may be due to the need to listen to the CAPTCHA challenge several times to memorize it, unlike the IRemember CAPTCHA, which can recognize the phrase immediately and do not need to listen to it again.

One way ANOVA value shows a significant difference between IRemember and the traditional audio CAPTCHA ($F(1, 21) = 384.13$; $p < 0.000261$). The t-test value indicates a significant difference between the success rate for IRemember and the traditional audio CAPTCHA ($p= 0.000261$).

Solving the traditional audio CAPTCHA method requires more time to solve than the IRemember CAPTCHA. This may be due to playing the audio challenges several times to memorize the random digits in the correct order. Unlike IRemember CAPTCHA, the participant recognizes the answer once they hear the common phrase, which eliminates the repetition needs.

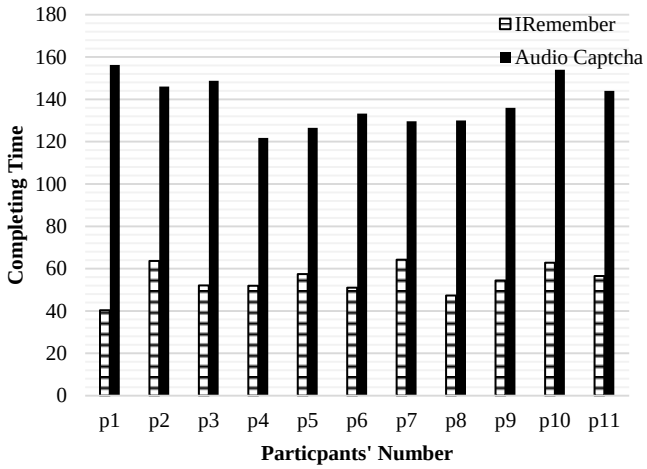

Fig. 4. The Average Time each Participant Spent on Solving IRemember and Audio CAPTCHAs.

C. Memorability Response

The difficulty of remembering the CAPTCHA digits of the traditional audio CAPTCHA was the main cause of error and confusion. Solving CAPTCHA requires memorization of the challenged content. Thus, the primary goal of designing the

IRemember CAPTCHA is to enhance remembering the CAPTCHA responses. So, we asked participants to listen to the CAPTCHA challenges and speak out the response before typing it to figure out the cause of the error, whether it is related to recalling the CAPTCHA response or it is related to the use of the input method.

The result shows that most participants were able to remember the response for the IRememeber CAPTCHA correctly (95%). The reason that caused an error at the stage of speaking the answer out was ignoring the preposition word that is in the provided phrase, for example, the phrase "I waited for a while." Most participants ignored the word "for" and only type "Iwaw". However, the main cause of error for the proposed method was the use of the input method that might lead them to double tap a letter which caused entering the letter twice. As a result, the success rate drops to 84% because the input method is not accessible and error-prone [17].

Regarding the traditional audio CAPTCHA, most participants could not remember the given random digits in the correct order, and sometimes they forgot a digit. Additionally, the other causes of error are that when they type their answer, they also make typing errors using the keyboard as they get confused when they listen to the screen reader and recalling what they should type, they got interfered and confused unlike on the proposed method they can just recall the phrase and determine the letter they should type.

D. Workload (NASA TLX) Questionnaire

The workload questionnaire contains six indicators that are applied to measure mental demand, physical demand, temporal demand, performance, effort, and frustration (see Table II). There is a significant difference based on all six aspects for both CAPTCHA methods. The average score of the mental demand indicator is 33.57 for IRemember CAPTCHA and 87.43 for the audio CAPTCHA method, and it shows the largest significant difference between the IRemember CAPTCHA and the audio CAPTCHA ($F(1, 21) = 262.877$, $p < 0.000159$). Because participants need to memorize the random six digits in the correct order, they must listen to the screen reader when it reads the letter under the user's fingers. The Temporal demand is the second highest significant difference between the IRemember CAPTCHA and the audio CAPTCHA ($F(1, 21) = 302.46$, $p < 0.000708$). The average score of the temporal demand indicator is 33.57 for IRemember CAPTCHA and 84.28 for the audio CAPTCHA method. It means there is high time pressure when solving the audio CAPTCHA. The third highest significant difference between the tested CAPTCHA is the frustration indicator; ANOVA result shows that ($F(1, 21) = 84.309$, $p < 0.000893$). Participants reported a frustration rate of 33.57 for IRemember CAPTCHA and 80.71 for audio CAPTCHA. It means participants feel frustrated when solving the traditional audio CAPTCHA. The fourth highest significant difference between tested CAPTCHA is in the effort indicator ($F(1, 21) = 136.7088$, $p < 0.000645$), where the average score of the effort indicator is 42.85 for IRemember CAPTCHA and 85.71 for the audio CAPTCHA method. It means that participants need to accomplish all tasks at one time, such as listening, memorizing the giving challenge, locating the answer box, locating letters on the keyboard, and locating submitting button. There was also a statistically significant difference between IRemember CAPTCHA and audio CAPTCHA in terms of performance indicators ($F(1, 21) = 132.923$, $p < 0.000754$). The performance score (80) shows that participants were not satisfied with the performance for audio CAPTCHA. Physical demand was also an essential indicator of workload, and the test showed a significant difference between the Physical demand of IRemember CAPTCHA and audio CAPTCHA ($F(1, 21) = 45.375$, $p < 0.000208$). In general, the t.test values reveal a statistically significant difference between IRemember CAPTCHA and audio CAPTCHA in the six indicators of NASA TLX workload ($p=0.000760$).

Overall, the interpretation of all six indicators is very high for the traditional audio CAPTCHA, where each indicator receives a very high score, which means users are not satisfied with the audio CAPTCHA. On the other hand, the IRemember CAPTCHA receive a score between 30 and 50 which is interpreted as somewhat high, and users might face some usability issue as well when solving the CAPTCHA.

TABLE II. AVERAGE NASA TLX SCORES FOR IREMEMBER CAPTCHA AND AUDIO CAPTCHA

TLX aspect	IRemember CAPTCHA	Audio CAPTCHA	T-test
Mental demand	33.5714286	87.4285714	0.00029
Physical demand	52.8571429	84.2857143	0.00170
Temporal demand	33.5714286	84.2857143	0.000377
Performance	45.7142857	80	0.000877
Effort	42.8571429	85.7142857	0.000211
Frustration	33.5714286	80.7142857	0.000596
Average of workload	40.3571429	83.7380952	
Interpretation of score	Somewhat high	Very high	

E. Strengths

The fundamental strength of IRemember CAPTCHA is that it is easy to remember and recall the CAPTCHA challenge response as it provides common phrases that users can easily recall what they just heard, whereas the traditional audio CAPTCHA requires users to remember the six digits as well as the sequence of the digits to pass the CAPTCHA challenge. Another strength of the proposed CAPTCHA is that it reduces the need to locate the repeat button to repeat listening to the CAPTCHA several times. It eliminates the need to repeat the CAPTCHA challenges several times to memorize the digits in the correct order. Users hear the CAPTCHA challenge once and begin typing the response based on memorization. It also removes the need for extra movements like locating the playing button or the response box; users need to interact with the screen to respond to the given challenge. In the proposed CAPTCHA, the cursor is directed at the response box to allow users to type their answers straightforwardly without a requirement to locate the response box.

It also provides a universal design as its interface offers both audio and visual interfaces for sighted people with no

vision. The most important point that makes IRemember CAPTCHA accessible for people with no or low vision is that it overcomes the limitation of memorizing the CAPTCHA content and eliminates the need to locate the position of the play button or the update button. These points improve the efficiency of the proposed CAPTCHA where it enhances the success rate because users can easily remember the answer to CAPTCHA questions in the correct order, as well as reducing the time spent responding to a challenge and minimizing the needed efforts to type the response where users can type three or four digits that are extracted from a given phrase. It also minimizes accidentally pressing an unwanted button, which leads to an error and much extra effort. The IRemember CAPTCHA is also easy to learn, and all participants were able to learn how to accomplish its task from the first trial.

F. Limitations

The main limitation of the proposed CAPTCHA and audio CAPTCHA was that the CAPTCHA challenges and questions were provided in English. Non-native speakers might require more time to memorize the given phrases. Thus, they will spend more time responding to the given challenge. To overcome this limitation, the challenges should be introduced based on the user's native language, preferences, and culture to enhance the user experience. Another drawback associated with the traditional audio CAPTCHA is requiring users to memorize the given challenges in order, which is a difficult task for users with no vision as they cannot see the challenge while typing; they need to depend completely on their memory.

G. Suggestions

At the end of the study, participants suggested multiple recommendations to improve the memorability and usability of IRemember CAPTCHA. They recommended allowing users to customize CAPTCHA challenges based on users' suggestions, personalities, culture, language, and region. Doing so, it can make the CAPTCHA challenge easier to remember. Another critical suggestion was to eliminate the need to type the first character from a preposition word and it is short, and it is the main cause of error in the IRemember CAPTCHA. Finally, they suggested integrating the proposed CAPTCHA into different platforms.

VI. Discussion

The IRemember CAPTCHA method outperformed the traditional audio CAPTCHA method. The success rate for the IRemember CAPTCHA method was higher than the traditional audio CAPTCHA. 82.72% was the success rate for the IRemember CAPTCHA, and it was 48.18% for the traditional method. The solving time was lower for the IRemember CAPTCHA, as well as it required less workload when solving its challenges. 55 seconds users spent to solve the IRemember challenge, and they spent 138.75 seconds solving the traditional audio CAPTCHA. According to prior research, the success rate of the traditional audio CAPTCHA is 46% and the completing time to solve the challenge is 65.64s [27, 28]. These findings show that the IRemember CAPTCHA achieved high performance in terms of success rate and solving time. Overall, the IRemember CAPTCHA achieved high performance in terms of success rate and solving time. Overall, the IRemember CAPTCHA approach is useable and accessible for visually impaired people, making it suitable for implementation as a real-world CAPTCHA in an application.

According to the NASA TLX scale, the IRemember CAPTCHA method requires fewer mental and physical temporal demands than the traditional audio CAPTCHA. The IRemember CAPTCHA requires users to recall the phrase and type the first letter of each word in the phrase. This task is simple and does not cause cognitive overload and memorization efforts. Thus, all users were able to recall the answer once they heard the given phrase. Unlike traditional audio, CAPTCHA requires users to remember six random digits in the correct order; these steps increase cognitive load as well as physical, temporal, and mental demand. Usually, visually impaired users tend to repeat listing to the given challenge trying to memorize the random digits. Thus, it causes a high level of effort, is demanding and is time-consuming [29, 30]. Therefore, the traditional audio CAPTCHA leads to high frustration once the users are required to repeat until memorize the digits as well as locate the repeat button and then the answer box, which needs more effort and mental demand. Whereas the IRemember CAPTCHA is less frustrating and requires less physical and mental effort than the traditional audio CAPTCHA. It eliminates the causes of confusion and the need to memorize the random CAPTCHA digits by allowing users to hear the common phrase and then recognize the intended character.

Findings show that most existing CAPTCHA requires intensive mental and recalling efforts to remember all digits. Unlike the IRemember CAPTCHA method requires less workload efforts [4-16].

Regarding the security of the proposed method, as the study result shows, IRemember CAPTCHA has the same level of security as the traditional audio CAPTCHA. In addition, IRemember CAPTCHA questions are easy to recall by humans and take only a few seconds to listen to the provided phrase and begin typing the answer [31, 32]. However, software bots find it complicated and time-consuming to solve these questions as the software must be able to determine the first character of each word in the phrase and extract the background noise from the CAPTCHA challenge [33].

VII. Conclusion

This paper discusses the implementation and evaluation of a novel audio CAPTCHA method called IRemember CAPTCHA. The IRemember CAPTCHA uses various approaches to build an accessible CAPTCHA for an extensive range of users. The proposed method applies a recognition-based CAPTCHA where the user is asked to listen carefully to the given phrase and determine the first letter of each word from the spoken phrase. Then, users are required to type or record the determined first letters. In addition, users can use swiping gestures to repeat the challenge or submit the response. The result of the evaluation study shows that visually impaired users find the IRemember CAPTCHA easy to learn and solve. However, the most important outcome was that the CAPTCHA was easy to remember for users.

In the experiments, the designed CAPTCHA has demonstrated a higher memorability and recall over the traditional audio method. Through experimental analysis, recording the response for the CAPTCHA method is considered a more effective technique than other input mechanisms like typing or gestures in CAPTCHA design. Moreover, the success rate was high, and the error rate was low. The IRemember CAPTCHA challenges take a different amount of time-based on the length of the phrase and the user's ability to type the response and repeat the challenge. The researchers indicated that the IRemember CAPTCHA is a valuable alternative method as it is a memorable and usable mechanism for people with limited abilities.

In the future, there is a need to improve the CAPTCHA usability by allowing personalization and customization of the presented challenges that are most common and known to the end users and is provided in their own language. Further improvements would make CAPTCHA more accessible by taking into account the impact of culture and language to simplify the way of responding to the given challenges. We also need to integrate the proposed CAPTCHA into a website to ensure its effectiveness and efficiency with the existence of different tasks around the CAPTCHA mechanism.

The study's main limitation is that the sample number is few due to the difficulty of finding a large number of users from this kind of sample demographic of the participants used for the study. We recruited only blind users who have good English language. Thus, there is a need to verify the study results by conducting a longitudinal study with more individuals with no or low vision and sighted users as well to see the effectiveness of the proposed method for all users. Finally, conducting a similar study with users with other types of impairments like mental or mobility impairment.

ACKNOWLEDGMENT

This research work was supported by Abu Dhabi National Oil Company (ADNOC), Emirates NBD and Sharjah Electricity Water & Gas Authority (SEWA), Dubai Electricity and Water Authority R& D Center as the sponsors of the 3rd Forum for Women in Research (QUWA): Women Empowerment for Global Impact at University of Sharjah.

REFERENCES

[1] M. Alnfiai. "Evaluating the accessibility and usability of a universal CAPTCHA based on gestures for smartphones," Universal Access in the Information Society, vol. 20, pp. 817-831, 2020.

[2] S. Roy, S. Nag, I. K. Maitra, and S. K. Bandyopadhyay. International journal of advanced research in computer science and software engineering. International Journal, 3(6), 2013.

[3] A. A. Su¨zen. "Uni-captcha: A novel robust and dynamic user-non-interaction captcha model based on hybrid bilstm+ softmax," Journal of Information Security and Applications, vol. 63, pp. 103-136, 2021.

[4] G. Schryen, G. Wagner, and A. Schlegel. "Development of two novel face- recognition captchas: a security and usability study," Computers & Security, vol. 60, pp. 95–116, 2016.

[5] G. Mori and J. Malik. "Recognizing objects in adversarial clutter: Breaking a visual captcha," In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 1-1, 2003.

[6] M. Jain, R. Tripathi, I. Bhansali, and P. Kumar. "Automatic generation and evaluation of usable and secure audio recaptcha," In The 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 355– 366, 2019.

[7] A. Dattolo and F. L. Luccio. "Accessible and usablewebsites and mobile applications for people with autism spectrum disorders: a comparative study," EAI Endorsed Trans. Ambient Syst., vol. 4, pp. 1-11, 2017.

[8] R. Gafni and I. Nagar. "Captcha: Impact on user experience of users with learning disabilities," Interdisciplinary Journal of e-Skills and Life-long Learning, vol. 12, pp. 207–223, 2016.

[9] A. Jarry, C. Chapdelaine, S. Kurniawan, and W. Wittich. "Blind adults' perspectives on technical problems and solutions when using technology," Journal of blindness innovation and research, vol. 7, pp. 1-10, 2017.

[10] J. Holman, J. Lazar, J. H. Feng, and J. D'Arcy. "Developing usable captchas for blind users," In Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, pp. 245–246, 2007.

[11] G. Sauer, J. Lazar, H. Hochheiser, and J. Feng. "Towards a universally usable human interaction proof: evaluation of task completion strategies," ACM Transactions on Accessible Computing (TACCESS), vol. 4, pp. 1–32, 2010.

[12] J. Lazar, J. Feng, T. Brooks, G. Melamed, B. Wentz, J. Holman, A. Olalere, and N. Ekedebe. "The soundsright captcha: an improved approach to audio human interaction proofs for blind users," In Proceedings of the SIGCHI conference on human factors in computing systems, vol. A247, pp. 2267–2276, 2012.

[13] M. Shirali-Shahreza and S. Shirali-Shahreza. "Captcha for blind people. In 2007 IEEE International Symposium on Signal Processing and Information Technology," vol. A247, pp. 995–998, 2007.

[14] S. Shirali-Shahreza, H. Abolhassani, H. Sameti, and M. Hassan. "Spoken captcha: A captcha system for blind users," In 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, vol. 1, pp. 221–224, 2009.

[15] V. Fanelle, S. Karimi, A. Shah, B. Subramanian, and S. Das. "Blind and human: Exploring more usable audio CAPTCHA designs," In Sixteenth Symposium on Usable Privacy and Security, pp. 111– 125, 2020.

[16] K. Kuppusamy and G. Aghila. "Human: an accessible, polymorphic and personalized captcha interface with preemption feature tailored for persons with visual impairments," Universal Access in the Information Society, vol. 17, pp. 841–864, 2018.

[17] Yadav, J., Aggarwal, Y., & Goel, V."QR Captcha," International Journal of Engineering Research & Technology. Pp. 398-401, 2020.

[18] M. Alnfiai. "A novel design of audio CAPTCHA for visually impaired users," International Journal of Communication Networks and Information Security, vol. 2, pp. 168–179, 2020.

[19] S. Baheti, H. Shah, R. Sherathia, and P. Waghamode. "CAPTCHA for visually impaired people: A review" vol. 8, pp. 1-6, 2021.

[20] A. Mathai, A. Nirmal, P. Chaudhari, V. Deshmukh, S. Dhamdhere, and P. Joglekar. "Audio captcha for visually impaired," In 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), pp. 1-5, 2021.

[21] S. Kulkarni and H. Fadewar. "Pedometric captcha for mobile internet users," In 2017 2nd IEEE International Conference on Recent Trends in Electron- ics, Information & Communication Technology (RTEICT), pp. 600–604, 2017.

[22] EnglishSpeak. [Online]. "1000 Most Common English Phrases," Retrieved from https://www.englishspeak.com/en/english-phrases, 2022.

[23] Schryen, G., Wagner, G., Schlegel, A. "Development of two novel face-recognition captchas: a security and usability study," Computer Security. vol. 60, pp. 95–116, 2016.

[24] Meriem Guerar, Luca Verderame, Mauro Migliardi, Francesco Palmieri, and Alessio Merlo. "Gotta CAPTCHA 'Em All: A Survey of 20 Years of the Human-or-computer Dilemma," ACM Computer Survey. vol. 54, pp. 1-33, 2021.

[25] A. Mathai, A. Nirmal, P. Chaudhari, V. Deshmukh, S. Dhamdhere and P. Joglekar, "Audio CAPTCHA for Visually Impaired," 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), vol. A247, pp.1-5, 2021.

[26] S. G. Hart and L. E. Staveland. "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," Advances in psychology, vol. 52, no.5, pp. 139 -183, 1988.

[27] Y. Jyoti, A. Yash and G. Vaishali, "QR Captcha," International Journal of Engineering Research and Technology (IJERT), vol. 09, no. 12. 2020.

[28] S. Graig, L. Jonathan, H. Harry, F. Jinjuan, "Towards a universally usable human interaction proof: Evaluation of task completion strategies," ACM Transactions on Accessible Computing, vol. 32 no. 15, pp.1-15, 2010.

[29] S. G. Hart and L. E. Staveland. "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," Advances in psychology, vol. 52, no.5, pp. 139 -183, 1988.

[30] J. Lazar, J. Feng and H. Hochheiser, "Research Methods in Human Computer Interaction," John Wiley and Sons, Chichester, UK, 2013.

[31] G. Haichang, W. Wei, Q. Jiao, X. Wang, X. Liu et al. "The robustness of hollow CAPTCHAs'," In Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security, Association for Computing Machinery, New York, NY, USA, pp. 1075-1085, 2013.

[32] Alnfiai, M., Alassery, F. TapCAPTCHA: non-visual CAPTCHA on touchscreens for visually impaired people. J Multimodal User Interfaces 16, 385–398 (2022). https://doi.org/10.1007/s12193-022-00394-2.

[33] G. Sauer, J. Lazar, H. Hochheiser and J. Feng, "Towards A Universally Usable Human Interaction Proof: Evaluation of Task Completion Strategies," ACM Transaction Access Computing, vol. 2, no. 4, pp.1-30, 2010.