Evaluation of An Augmented Reality Approach to Better Understanding Diverse End User website Usage Challenges

Minh Hieu Vu, Joshua (Shuki) Wyman and John Grundy[a]

Department of Software Systems and Cybersecurity, Monash University, Australia

vuminhhieu1105@gmail.com, jwym2@student.monash.edu, john.grundy@monash.edu

Keywords: Human computer interaction, Accessibility, Human Aspects, Disability Simulation, Software Engineering

Abstract: End users with disabilities face many limitations with online accessibility. Efforts to tackle the problems of testing for these problems have not yet been sufficient. In this paper we describe an evaluation of Funkify, an augmented reality tool to simulate diverse end users' experiences and challenges when interacting with software. The research evaluates the tool's effectiveness for software engineers use in emulating four diverse end user challenges using a heuristics-based evaluation and cognitive walk-throughs on three target websites. Our results show that the tool successfully simulates the target personas' challenges varying extent, with some noticeable limitations. We provide suggestions for extensions to Funkify as well as AR tools in general.

1 INTRODUCTION

An estimated 1 in 6 Australians lives with a disability [Green et al., 2021]. In an online world that demands ever greater online access to services since the onset of the COVID-19 pandemic, it's imperative that websites be conscious and accommodating of these diverse end users. Unfortunately, the problem is pervasive. Almost all of the top 1,000 free Android apps have been shown to have many severe accessibility issues [Alshayban et al., 2020]. The top 100 banking websites in the US have approximately 6 accessibility violations each on average [Wentz et al., 2019].

Research into accessibility has traditionally focused on the development of tools to automate issue identification and verify compliance with accessibility guidelines such as the Web Content Accessibility Guidelines (WCAG) [Abascal et al., 2019]. However, the issue continues to be severe. The US saw 3,500 website accessibility lawsuits filed in 2020, and the rate has risen 64% in the first half of 2021 [Alcántara, 2021, Lazar, 2019]. A possible underlying cause of this trend is a lack of understanding amongst software engineers for people different from themselves. Developers may lack the insight to fully grasp how different types of end users would be interacting with their system, making it very difficult to design and build suitably accessible products [Bi et al., 2022].

In this paper we evaluate the suitability of an Augmented Reality (AR) approach to providing developers with an experience of how diverse end users interact with their system. In contrast with the body of prior research, rather than highlight suggested changes or accessibility improvements, the aim is to elicit a lasting empathetic response from the developer and a deeper understanding of their users' needs. To do this, we used a browser plugin to mimic a variety of diverse end user challenges. Funkify is a publicly available extension for Google Chrome that offers a range of simulators for vision, motor, and cognition impairment as well as dyslexia [Funkify, 2021]. The extension breaks users down into personas, with each representing a class of challenged end users. Some available personas are shown in Figure 1.

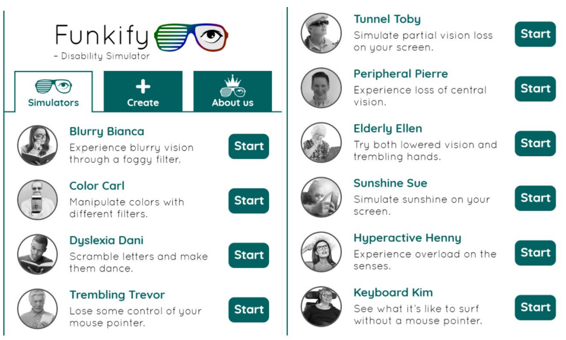

Figure 1: Examples of Funkify Personas

We evaluated Funkify Premium, which is available on a subscription basis and provides access to 4

[a] https://orcid.org/0000-0003-4928-7076

additional personas, the ability to manually adjust the sensitivity of each of the personas, and the ability to define new personas using a combination of any of the simulators. The free version of Funkify limits the colour-blindness simulator to a single type of colour blindness, whereas Premium allows the user to select between 7 different types of colour blindness to filter the screen for. Similar levels of customisation are available for each of the personas.

Section 2 outlines key related work to date in helping address accessibility challenges. Section 3 describes our research methodology, and Section 4 our study design details. Section 5 presents the details of our Funkify evaluation, and Section 6 discusses key findings and makes a number of recommendations for enhancement of the tool.

2 RELATED WORK

Humans are different and many have diverse challenges when making use of software solutions [Grundy, 2021]. Some of these are related to physical and mental challenges of the users and software not designed taking these into account will suffer accessibility problems [Rutter et al., 2007, Harper and Chen, 2012]. Currently software engineers lack tools and techniques for adequately modelling end users with diverse challenges that impact accessibility.

Alshayban et al. [Alshayban et al., 2020] developed an automated accessibility evaluation tool that they used to evaluate 1135 free Android applications and conducted a follow-up survey of Android developers. They found that accessibility issues are prevalent and pervasive throughout all categories of apps, and that these issues are not limited to affecting a single type of user disability. Although they surveyed a relatively small sample of 66 developers, it was found that the respondents lacked awareness of accessibility issues or training to address them. It's noteworthy that the minimum amount of development experience among respondents was 0.25 years, a very short amount of time that may raise questions about the ability to generalise this result. An interesting observation was that no existing accessibility analysis tools are prioritising issues by severity or importance, leading developers to either overlook critical flaws or be overwhelmed with the quantity and variety. They identified accessibility issues within half of the templates provided directly by Android Studio. Developers view these templates as a trustworthy foundation and build their apps around them, thereby propagating the issues. This raises the question of whether a similar phenomenon may be occurring in the website space, for example with front-end component libraries or popular website templates.

Bai et al. [Bai et al., 2019] investigated and compared six accessibility testing methods for software development teams. They evaluated these methods in terms of their usefulness, satisfaction, ease of use and ease of learning to determine the methods that would yield the best engagement for the participants, investigated how different software roles and development phases would affect the choice of appropriate testing methods. The sample size of participants was relatively small (53 participants) with 74% were male, and it would be interesting to see how this fact relates to the empathy of developers towards end-users. A notable observation was that there are no set rules to determine the types of issues a testing method best identifies, which made it difficult to choose a method and resulted in developers neglecting testing entirely. The de facto method for the participants was WCAG walk-through. However, the assessment of this method yielded the lowest result in terms of satisfaction and ease of use with a hostile attitude in many responses. This poses a question of whether enforcing use of a specific testing method is appropriate to engage the developers. The research highlights how different software roles prefer different methods.

Schulz and Fuglerud [Schulz and Fuglerud, 2012] introduced methods to create personas more comprehensively. They presented the potential barriers, proposed multiple techniques, and discussed the application in their studies. A notable observation made was that if personas are simply recycled or imitated, then the empathy and engagement from developers is lost. Also, personas are not complete replacements of true end users, and misconceptions are possible when personas are developed without prior real interactions. Schulz and Fuglerud provided in-depth details to some commonly known methods and conducted research on real-world use cases. They made suggestions for persona creators to be mindful when using different Assistive Technology versions, and aim at creating personas with the same learning attitudes as focus groups.

Several design approaches have been developed specifically for visually impaired and blind users, encompassing a range of sight related accessibility challenges [Sierra and Togores, 2012, Jacko and Sears, 1998]. Dyslexia is another accessibility challenge faced by a range of users but with limited research on solutions to date [Wery and Diliberto, 2017]. A number of recent works have looked to to help software engineers take better account of diverse end users during development. This includes use of enhanced personas capturing user differences that allow software engineers to more easily examine accessibility considerations at all stages of development without requiring the overhead of finding many varied live users [Li et al., 2021]. It also includes work extending modelling languages documenting different end user accessibility-related and other human-centric related challenges [Jim et al., 2021]. Recent work has investigated how to support users with mild cognitive impairments [Jamieson et al., 2020].

Some work has investigated use of augmented reality-based interfaces to aid improved accessibility in software development. Biswas et al. [Biswas et al., 2012] discussed different user modelling approaches in regard to designing inclusive interfaces for elderly and disabled people. They proposed a simulator to address the limitations of existing modelling techniques in predicting likely interaction patterns and estimating the time needed to complete an action for users with disabilities. Stearns at al. [Stearns et al., 2018] design and evaluate an AR-based magnification aid for sight challenged end users to improve accessibility. Such approaches aim to fill the gaps in limited existing user modelling techniques by breaking interactions up into smaller components and combining different approaches in each component, while considering the needs of challenged people with visual or mobility impairments. In Biswas et al., roughly 30% of situations had more than 50% relative error. This raises the question of whether these case studies are sufficient to determine the validity of such simulators.

3 METHODOLOGY

In this work we wanted to answer the following research questions:

RQ1: Can a browser-based augmented reality simulation tool such as Funkify personas give software engineers a useful experience of diverse end user web accessibility challenges?

RQ2: How comparable are Funkify's simulations to documented experiences of web accessibility challenges?

RQ3: What extensions to tools such as Funkify would enhance their usability for more human-centric software engineering?

3.1 Our Approach

The research was divided into three major steps.

- Identification of a range of exemplar websites requiring diverse end user accessibility.
- Identification of a range of Funkify personas representing diverse end users with quite different accessibility challenges.
- Evaluation of Funkify Premium personas on several representative websites.
- Identification of needed improvements to the current state of Funkify.

The range and number of personas and websites evaluated were chosen with regards to a diverse, representative set of digital services needed by many in the community and relatively common end user accessibility challenges.

3.2 Funkify Persona Selection

Personas are fictional profiles of characters, created to represent different types of users with the aim to provide a perspective, or observations about different challenges or experiences diverse users face, ultimately building empathy from developers and designers towards their end users [Pruitt and Grudin, 2003]. In our research, these personas represent users with a disability or multiple disabilities that affects the users' experience while interacting with software, whether they be visual, motor and/or cognitive impairments.

Our set of target personas were selected on the basis of (i) covering a wide range of user's challenges, and (ii) relatively common disabilities that would have the most effect on users' experience and would be difficult to simulate effectively and thoroughly with an augmented reality tool such as Funkify. From the 10 personas that Funkify provides, 4 personas were selected to conduct our evaluation, consisting of:

Blurred vision: Blurred vision can have a negative effect on a persons' entire line of sight or partially affect one's vision. It could include peripheral vision, or how a person sees to the left or right of their field of vision, it can also be possible to experience blurred vision in one eye only. It is often caused by refractive errors (nearsightedness, farsightedness), abrasions to the cornea, age-related macular degeneration, migraine, trauma or injuries to the eye, infectious retinitis [Zhou et al., 2020].

A person with blurred vision can have their functional status and overall well-being severely impacted, according to P. Lee et al. They stated that blurred vision can have a substantially greater impact in role limitations due to physical health problems than that of hypertension or type II diabetes. It is also stated to be of significantly worse impact on energy than type I diabetes mellitus, on social function than indigestion, and on physical functioning than trouble urinating [Lee et al., 1997].

Dyslexia: Dyslexia is considered as a language-based learning difference that affects the organization in the brain which controls the ability to process the way language is heard, read, spelled, or spoken. Dyslexia can also have negative effects on a person's working memory, attention, and organization. Davis's research reported 37 common traits of a dyslexic that spans across vision, speech, hearing, writing, motor, math and time management, cognition, behaviour, personality... Davis stated that people with dyslexia exhibit about 10 of the common traits and behaviours and these characteristics vary inconsistently [Davis, 2004]. According to Davis, some notable traits and characteristics of people with dyslexia include:

- Reading or writing shows repetitions, additions, transpositions, omissions, substitutions, and reversals in letters, numbers and/or words.
- Complains of feeling or seeing non-existent movement while reading, writing, or copying.
- Extended hearing: hear things that are not said or apparent to others; easily distracted by sound.
- Difficulties with fine and/or gross motor skills and tasks.
- Poor memory for sequences, facts and information that has not been experienced.
- Extremely disorderly or compulsively orderly.
- Mistakes and symptoms increase dramatically with confusion, time pressure, emotional stress, or poor health.

Tremor: Tremor is characterised by shaking movements in a part of the body caused by involuntary muscle contractions. One of the most common neurological diseases, tremor can occur on its own or in conjunction with another neurological disease such as Parkinson's, MS, or stroke. For Parkinson's alone, 1 in every 100 Australians over the age of 60 lives with the disease [Dorsey et al., 2007]. Tremors are classified as either rest or action, with action further subdivided to give 7 types of tremors, each with their own typical frequency and amplitude. [Sirisena and Williams, 2009]. Designing apps to support people with tremor has been shown to be challenging [Zhong et al., 2015].

Tunnel vision: Tunnel vision is a colloquial name for peripheral vision loss, a narrowing of the field of view to the extent that the individual can only see directly ahead. [tun, 2019] It is most commonly seen in patients with retinitis pigmentosa or glaucoma, with glaucoma alone affecting over 300,000 Australians. Limited work to date has been done on design guidelines and support for users suffering from tunnel vision [Kamikubo et al., 2018].

3.3 Target websites

To perform the evaluation on the selected personas, a set of target websites were chosen. Our approach to choosing these websites were to cover a wide range of topics, layouts, content of the sites (content-heavy news sites, websites with a lot of interactions, pictures...) The websites were chosen to also differ in user needs when using them, spanning from recreational purposes to daily fundamental needs. In the end, a set of 3 such target websites were selected:

Commonwealth NetBank: Commonwealth Bank of Australia (CBA) is one of the most popular banks in the country and has an online system to support all the fundamental banking tasks like checking your balance, making a transaction, finding an ATM or branch, and so on. These are tasks that most people do almost every day, and can have direct detrimental effects on a person who is unable to do them effectively and accurately.

Reddit: Reddit is a social news platform with a large user community, covering a wide range of topics and subjects. Its website is very content-heavy consisting of mainly text and pictures with a lot of interactions between interactions. Reddit was selected to assess the experiences of diverse end users on a content-heavy site where most of the interactions and content are casual using a lot of colloquial language.

Amazon: Amazon is the world's largest e-commerce website with more than 1.5 million transactions every day. Amazon was selected to evaluate how one of the largest and most visited websites in the world design their platform in regard to accessibility for diverse end users.

4 STUDY DESIGN

4.1 Method

We conducted a heuristics evaluation of our selected personas using Funkify Premium applying a set of evaluation criteria, including:

- What range of diverse end user challenges does Funkify support? How do such challenges manifest in the browser?
- How well does the tool work with our selected websites when performing tasks?
- How is the AR environment produced by the tool for developers comparable to the documented experiences of end users with these challenges?
- Does the modification of website interaction appear to be based on actual evidence or literature?

- Do Funkify personas provide a software developer a good idea of how someone using the website with this challenge would find the experience? Can the software developer "empathise" with this target end user's accessibility-related challenges?
- Can users with multiple accessibility-related challenges be addressed, and how feasible and efficient does the tool address users with such multiple challenges?
- What new challenges (combination of challenges) are we able to add to the tool? What challenges are not possible to add?

To support the evaluation of this set of criteria, we conducted cognitive walk-throughs with each persona on all the target websites. From the results of the cognitive walk-throughs, we assessed how effective the persona's challenges manifest in the browser with Funkify, what the notable limitations of the tool for a persona/target website are, and how the personas challenges relate to real users' experience based on existing accessibility studies and literature on the disabilities. From these we identified opportunities for improvement in Funkify's augmented reality approach for better supporting software engineers in designing and building interfaces for diverse end users.

4.2 Cognitive walk-through

Cognitive walk-through is a method primarily used in usability evaluation to look for usability issues in interactive systems, with a focus on task completion for novice users [Blackmon et al., 2002]. Its emphasis is on studying how easy it is for new or infrequent users to learn a system. It was first used as a tool to evaluate systems such as ATMs or interactive exhibits in museums where users will generally have little to no prior experience or training. Due to its ease of use and feasibility, the method's usage has been extended to complex software systems including CAD and software development tools. [Rieman et al., 1995]

Cognitive walk-throughs were conducted, one for each selected persona and website combination. We chose a set of tasks for each website and conducted the walk-throughs. Our approach to selecting these tasks for each target website was to select those that were the most basic and relevant for all the websites (e.g., logging on, registering) and the most major tasks for each site (e.g., make a transaction). A series of 4-6 major tasks were chosen for each of the target websites. The next step was to define a goal or success criteria for each of the tasks to determine how effective can the task be done in the augmented environment created using Funkify for each of the personas. After finalising the tasks and their respective definition of success, we would then conduct the walk-throughs for each of the personas on all the target websites and tasks.

Each persona was be assessed as to how it affects the tasks' feasibility, whether the task could still be done with the Funkify filter active at different intensity settings or if it proved to be impossible to complete the task when Funkify is active, based on the defined success criteria. The results of each task and target website from the walk-throughs could then be generalised into the feasibility and effectiveness of the Funkify persona in terms of how well it achieves the simulation of the persona's challenges and what the major limitations of the tool are or in general the limitations of augmented reality tools in simulating diverse end users' experience.

The evaluators of the websites and Funkify personas were software engineering students with real-world software practitioner experience. Each has expertise in software engineering processes and tools, UX/UI design, web and app development, but no particular training in designing for accessibility.

4.3 websites

All of the target websites have similar tasks for logging on and registering a new user where the definition of success is straightforward and consistent among all websites. Key tasks which differ across the websites are as follows:

Commonwealth NetBank tasks:

Make a transaction: this is the most common and fundamental task to be completed on a banking website. The definition of success was to be able make a transaction for the correct amount of money from a specific account to the correct recipient. The user would need to navigate to the transaction page, choose the correct 'To' and 'From' accounts and input the correct amount to complete the task.

Find a branch or ATM: Commonwealth NetBank provides a feature to show a map of ATM or branches nearby or to selected set of filter options. How this task is considered as done can vary between use cases, so we chose to define success for this task to be able to navigate around the map effectively and be able to find a specific CBA branch on the map.

Reddit tasks:

Join and navigate subreddits: Reddit is a social news platform where each subreddit represents a community that focuses on a specific category or topic. It was difficult to define a tangible success criterion for navigating through the subreddits, so we focused on joining a subreddit while assessing how the Funkify filters affect the experience of navigating through subreddits.

Read comments on posts: comments are a large proportion of all content and interactions on Reddit posts. Users need to successfully read comments, navigate through all comments and replies, and be able to follow and understand comment threads.

Post on subreddit: A Reddit post can be simple as a sentence or it can consist of pictures, links, various font options... We decided to assess the task's completion on the user's accessibility and effectiveness in using these additional options when posting.

Post comments: this task involves composing a comment relating to a post and is relatively simple, so a definition of success is not required.

Amazon tasks:

Find a product (example - an HDMI cable): Being an e-commerce site, Amazon's most fundamental task would be to find a product. The definition of success was to be able to search for a product (an HDMI cable was selected for this task) and effectively compare all the search results (in terms of descriptions, prices, quantities...).

Buy a product: conventionally buying a product on Amazon involves selecting the product, comparing and choosing the products' options if any, adding it to the user's cart, and checking out with the required user billing details. For this research, we define this task as completed when user has selected the right product, read its descriptions effectively, compared between product's options, and add it to their cart.

5 EVALUATION RESULTS

For each persona we describe how Funkify attempts to simulate the underlying accessibility issue, report key findings from the cognitive walk-throughs [RQ1], and discuss the extent to which the experience provided by Funkify mirrors the documented experiences of diverse end users with these challenges [RQ2]. Notably, Funkify only augments the browser window underneath the title bar, so any changes to the url or open tabs are not concealed. A limitation of the tool is that it only applies to the view within the page so all the other components in the browser UI are not affected such as tabs' names, browser menu, settings. This is of course not consistent with the real-world experience of a person with blurred vision.

5.1 Blurred Vision Augmentation

Blurred vision is the most common sight challenge, but varies in level. Funkify applies a blur filter to the entire window, with intensity being able to be varyied on a 0-10 scale. Figure 2 shows the persona and Figure 3 how this manifests for the CommBank website ATM locator page.

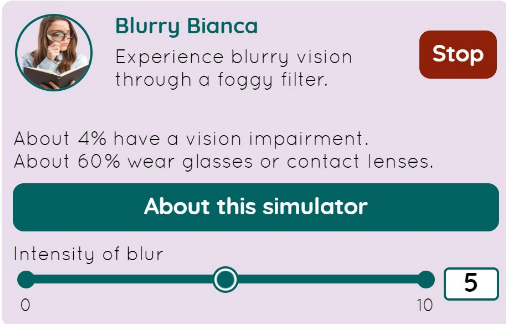

Figure 2: Blurry Bianca UI

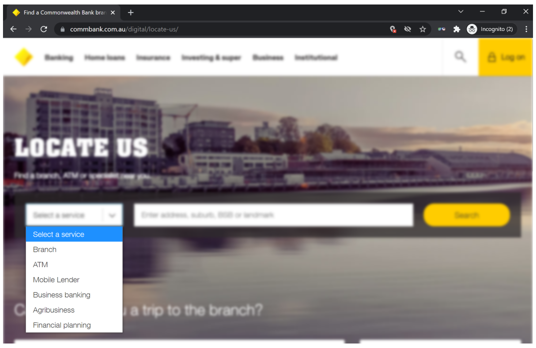

Figure 3: NetBank ATM search dropdown

While there were isolated instances like Figure 3 of components such as drop-down selections appearing as normal on top of the filter, overall the simulator performed consistently well. It immersed the user in the persona and allowed them to identify areas where the site design led to user frustration.

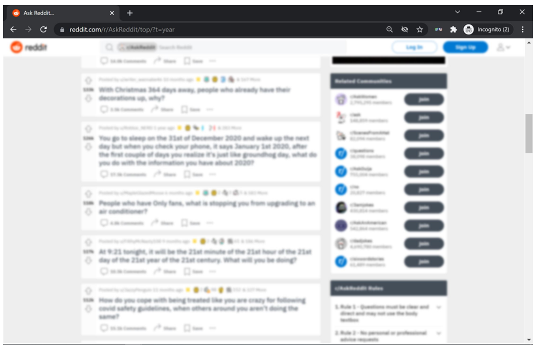

Figure 4: Reddit with Blurry Bianca filter, 2.5 intensity

For content-heavy websites like Amazon and Reddit (Figure 4), the visual strain made browsing particularly difficult and tiring. However, fatigue was also experienced when navigating NetBank, with lower information density. Amazon, with many busy pages, description blocks and copious amounts of small, clickable, text was very taxing to navigate. Product titles and costs, as the top elements of Amazon's typographic hierarchy, are the only readable parts of the page at 1.5 intensity. Even Amazon's captcha posed a difficulty. Although the registration process allows the user to select an audio challenge, the explanation of what is required is shown in unreadably small font.

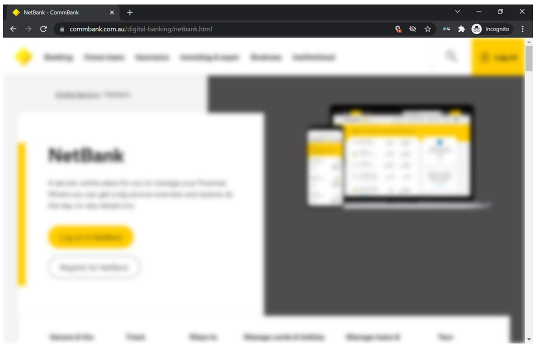

Figure 5: NetBank login colour highlights the most important interactions

As navigation was tiring, accessible and bold design choices stood out for their ease of use. The use of contrasting and vibrant colours drew the user's focus more easily, and was used to good effect by NetBank's yellow log in button as shown in Figure 5. It also highlighted the benefit of iconography and logos as opposed to reliance on text. Reddit's login modal has the Google and Apple logos alongside their login options. The text became unreadable at 1.8 intensity, whereas the logos remained recognisable until 3.0, reducing the cognitive strain.

The experience produced by the tool was very comparable to that of a short-sighted person when not using prescription glasses. A noticeable takeaway is that for most of the evaluated websites and tasks, it is almost impossible to read any content or text on the screen when the intensity reaches 3.0, so a large portion of the intensity spectrum will yield the same results when using this augmentation. And the characteristic of the blur filter is the same for all the websites and all the intensity levels (a blur effect to the entire screen) which might not be the experience for all users with blurred vision (partial blurred vision, left or right blurred vision, short-sightedness compared to age-related causes...).

5.2 Tunnel Vision Augmentation

The Tunnel Toby Funkify persona, shown in figure 6, shrinks the visible area in the browser window to either a circle or rectangle centered at the current position of the user's mouse. The remainder of the window is covered by an opaque black filter. A sensitivity slider from 0-10 is provided to change the amount of vision loss desired. We evaluated the circular option intended to simulate peripheral vision loss.

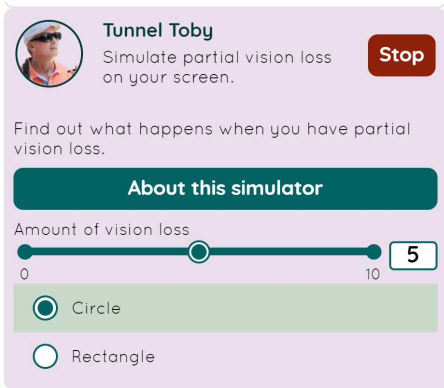

Figure 6: Tunnel Toby UI

Evaluating this simulator was found to be a very confronting and intensely uncomfortable experience. Although solely isolated to the browser window, and with the user able to stop at any time, prolonged exposure to this persona repeatedly led the evaluators to spikes of anxiety and claustrophobia. On unfamiliar sites navigation times were dramatically slowed due to the need to systematically scan each page, building a mental picture of the structure and ensuring key details were not overlooked.

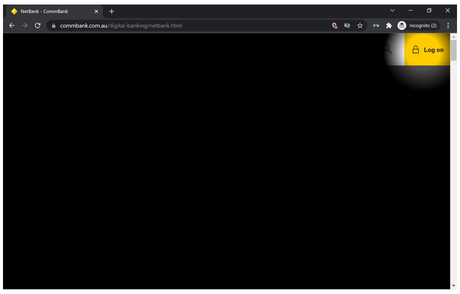

Figure 7: NetBank login button placement

Figure 8: Reddit hierarchy lines

To compensate, there was a noted reliance on established UI norms. For example, assuming a button to login would be in the top right corner of the window (shown in Figure 7). Similarly, visual identifiers of page structure helped with maintaining an understanding of position. As shown in Figure 8, Reddit displays its nested comment hierarchy with parallel vertical lines signifying the level of indentation, where clicking on a line navigates the user to the parent comment of that level. Without such a visual aid it would be extremely difficult to follow conversations.

This confusion manifested when evaluating Amazon's product search. Products were arranged inconsistently, alternating between a single product per row and three separate products side-by-side. Advertisements and sponsored products were interlaced at unpredictable intervals. Product images were larger than the visible circle, and product names were long and technical in nature. All of these factors combined to make comparisons between products difficult.

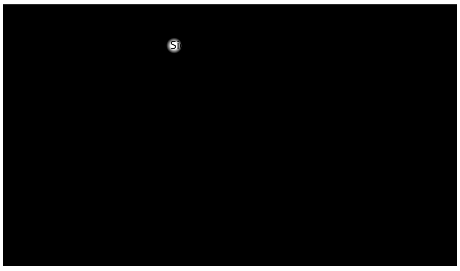

Figure 9: Tunnel Toby sensitivity 10

A number of bugs were identified with Funkify's tunnel vision augmentation. When using the persona, the mouse could not enter the Reddit login modal and lagged significantly when Reddit posts were opened. Scrolling on a page moved the visible area in tandem, as opposed to it sticking to the mouse position. This resulted in most scrolling occurring with an entirely black screen, only refreshing on mouse move.

Funkify provides a 0-10 slider for amount of vision loss, however this scale is quite a narrow range. 0 is no vision loss at all, but 0.1 is already significant impairment. 10, demonstrated in Figure 9 is almost complete vision loss. To better reflect reality this range should be much wider, with lower numbers just showing some darkening around the edges of the screen and gradually decreasing the vision.

5.3 Dyslexia Augmentation

Funkify attempts a simulation of dyslexia by scrambling text on the web page. Scrambling is localised to a word with the exact characters swapping at random. Figure 10 illustrates this persona.

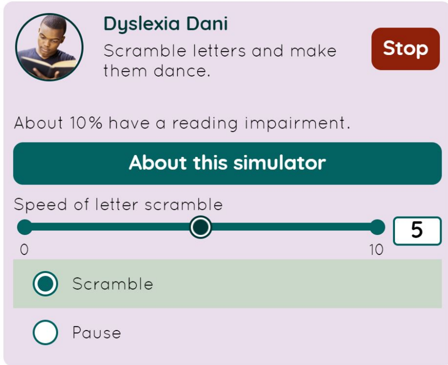

Figure 10: Dyslexia Dani UI

It was found that content-heavy sites such as Amazon are difficult to read, and small site elements escalate the challenge. It became difficult to ascertain and remember product names when trying to find a product that had been seen previously. Evaluating Reddit (Figure 11) exposed a set of challenges around comprehending slang and internet colloquialisms, which proved to be barriers to entering the conversation.

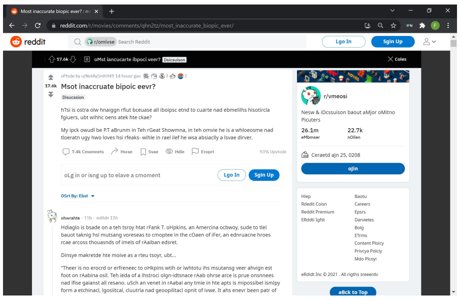

Figure 11: Browsing Reddit with Dyslexia Dani

The simulator only changes page text, and therefore it does not scramble any text in images, branding, tooltips, and certain buttons. It fails to work on Reddit's login modal and seems to not scramble two-digit numbers, even when they appear as text. Additionally text typed into websites remains unaltered. Figure 13 shows the ATM search stage of a NetBank evaluation, where neither the input text or dropdown list are scrambled. These limitations cumulatively limit the immersive experience of the simulator, as in Figure 12 where the focal point of the screen is text embedded in an image.

With each letter generally not moving far relative to their correct position; words with 2 or less characters stay the same during the simulation. The tool's capability is only limited to the visual aspects of dyslexia, so it is not comprehensible enough to generalize the whole set of challenges a dyslexic person experiences. The tool is not able to cover other aspects related to the persona including hearing, writing, mo-

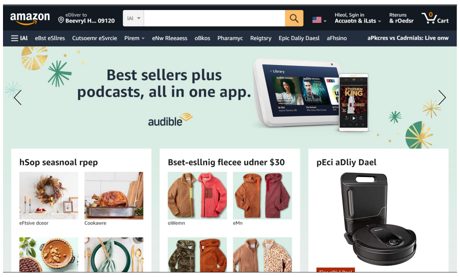

Figure 12: Amazon home page - banner unaffected by dyslexia filter

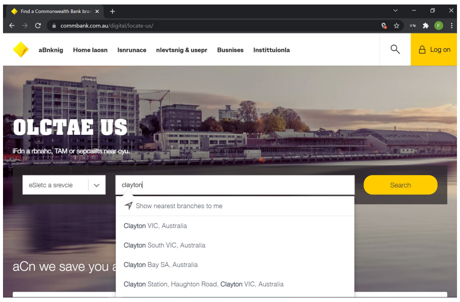

Figure 13: Dropdown item and form inputs do not scramble

tor, especially behavioral traits. It is not clear that all people with dyslexia experience reading in this manner, and some have simulated dyslexia by removing lines from normal lettering [Bai et al., 2019].

5.4 Tremor Augmentation

The Trembling Trevor Funkify persona simulates tremor by moving the mouse involuntarily and unpredictably in all directions. As the mouse continues to move without user input it would be classified as a resting tremor, as is commonly seen in Parkinson's disease. The amount of tremor is adjustable on a 0-10 scale. Figure 14 shows the persona from Funkify.

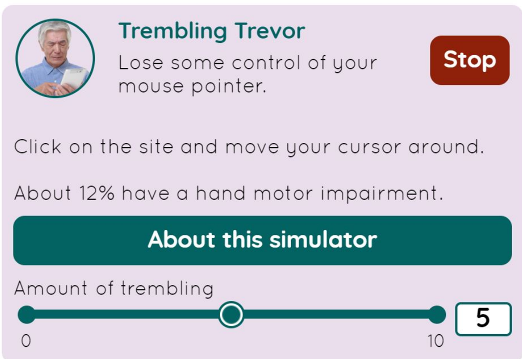

Figure 14: Trembling Trevor UI

While using the persona, accurately selecting small targets posed a significant challenge. This was most prevalent in the Amazon walk-through, due to the abundance of clickable text. When searching for products the filter options are inaccessibly small, as is the breadcrumb navigation in the user profile section.

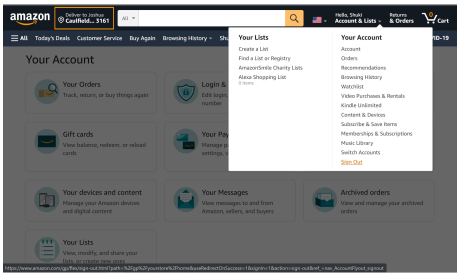

Figure 15: Amazon sign out button

Critically, signing out of an account necessitates clicking on a small piece of text at the bottom of a dropdown menu, as shown in Fig. 15. It requires fine motor control to accurately click the correct button, with the additional frustration that unintended movements that cause the cursor to move outside of the box results in the dropdown menu collapsing.

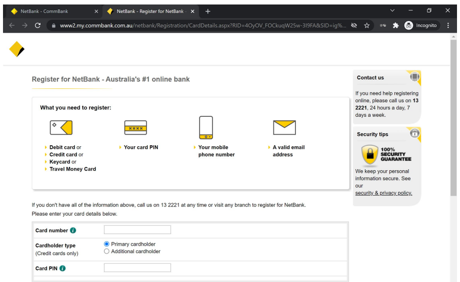

Figure 16: NetBank registration form radio buttons

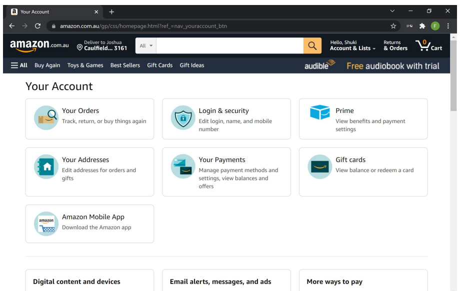

Figure 17: Amazon profile page

Similarly, when applying to register for NetBank (Figure 16), the user is prompted to select between radio button options, which were basically unclickable with this Funkify persona enabled. The most accessible pages had large, separated buttons, like on Amazon's user profile page shown in Figure 17.

The sensitivity adjustment is a 10-point sliding scale of an ill-defined "amount of trembling". Although probably sufficient to emulate the experience of navigating with a tremor, this could be broken down into frequency and amplitude to better align with the medical literature. An advanced feature set could also break it down by tremor type.

6 DISCUSSION

Cognitive walk-throughs with Funkify provided both valuable insights into accessibility issues on the evaluated sites and also a lasting impression on the evaluators of the constraints experienced by end-users who live with the examined disabilities. The lessons learned were applicable beyond the local scope of one website, and the tool elicited the desired empathetic response above and beyond accessibility issue identification. We recommend its use in further research that can explore this relationship further and to identify whether it can lead to improved long-term improved accessibility outcomes.

Funkify as it currently exists is a straightforward and easy to use extension that offers flexibility and customisation across a wide range of disabilities. Personas are toggled with a single button click and any combination of existing simulators can be combined into a new persona. Having all of the functionality neatly packaged in a single plugin allows Funkify to add value if included in future projects that use personas to model diverse end users. However, a number of potential limitations were found, and a number of extensions to Funkify identified that would enhance its suitability as a basis for future research [RQ3].

6.1 Expanded personas

Prior research on the use of personas to understand end users includes more biographical and emotional context than Funkify currently allows. Other work has proposed personas that are tailored to the needs of the development team and express the background, goals, and frustrations of the modelled end user to specific problem domains [Li et al., 2021]. We propose that another button be created underneath "About this simulator" for each Funkify persona, where the user will be able to read and edit more of the human aspects of the persona. This ability to enrich Funkify personas by defining more contextual data would improve the quality and utility of the personas for testing purposes, and simultaneously provide a mechanism by which to instantly simulate the perspective of the previously text-based persona.

6.2 Sharing capability

As a Chrome extension, any custom personas or changes to sensitivity settings are localised to a single account. To facilitate repeatable testing and consistency amongst team members, support is needed for sharing personas with other Funkify users. This could be combined with extending persona human aspects to provide richer, more contextual personas.

6.3 Predefined intensities

At present, Funkify Premium always allows the user to adjust the simulator intensity. However as we move to richer personas we may, for consistency, want to specify the exact desired values and disable the ability of users to deviate from it. Therefore we suggest that the screen to make or edit a new persona should contain a field for whether the intensity values are locked. It will always be possible to return to the edit screen to unlock them, but this feature would prevent accidental miscalibration. If this implementation is not possible, we would suggest a button in the persona window that would return all values to a preset default.

6.4 Broader range of intensities

A finding of our cognitive walk-throughs was that the range of intensity options provided for each persona was too narrow, often resulting in the user solely utilising the lower end of the spectrum. For example the minimum amount of vision loss (0.1) for the Tunnel Toby persona shown in Figure 9 is already quite advanced peripheral vision loss whereas the circle size in upper half of the range (> 5) is too small to be useful. The sensitivity range should be re-calibrated to allow the minimums to represent a far more gentle form of the disability.

6.5 Combining personas

Funkify provides a limited ability to define new personas and some target end users have multiple challenges that would be good to be able simulate concurrently. As Funkify uses different approaches to its augmented reality-based simulations, this may be difficult if not impossible for some multiple challenge personas. This may require a more sophisticated augmented reality-based approach.

6.6 New simulators

Finally, we would add new simulators to Funkify targeting aspects of dyscalculia, hearing impairment, and developmental disabilities including autism. While these conditions may not be able to be replicated precisely, we would implement targeted parts of their documented end user experience from other studies and medical literature. As above, we would aim to leverage Funkify's ability to combine simulators in order to scaffold a more holistic persona from constituent symptom personas.

6.7 Limitations and Future Work

One major limitation in this research was limiting our evaluation to Funkify. This opens further potential research paths and questions as to whether all of the personas provided by Funkify yield the same evaluation results as other comparable tools, or if there are better performing tools compared to Funkify to simulate diverse end users' experience for some challenges. Although our selection approach was to cover the widest range of personas and websites possible, there are of course many other widely used websites and tasks that could be evaluated with Funkify's augmentation approaches. The small number of people doing the cognitive walk-through-based evaluations is also a limitation. This can result in the lack of integrity between each cognitive walk-throughs as members would have had some prior experience with the target websites and tasks. With the evaluation only done internally by research team members with experience in software engineering, UI/UX and web development, this may not reflect actual experience of target end users. However a major aim of our work was to see how software engineers perceived the augmented browser-based interfaces, and whether they could help identify key usability challenges for such target end users. We also wanted to see if software engineers could gain a more empathetic understanding of these target end users by using Funkify's persona and augmented browser-based interface approach.

Another limitation in our study was the lack of real end user involvement. This can be overcome in the future by conducting user studies with actual visual-impaired users. It would also help to overcome the limitation of possible bias in current participants since they were from a software engineering background. A comparison of task completion between real versus simulated user groups may reveal further interesting insights. However, we did compare Funkify's augmented interfaces to the impacts of different user challenges reported in prior studies and medical literature. Further studies into evaluating a wide range of users at different levels/manifestations of the disabilities are also necessary to assess the impact on their web usage and study different UI solutions required to cater to different range of users with the same disabilities (e.g., different levels of tunnel vision, different manifestations of dyslexia).

The effect on software developers and designers after using tools such as Funkify on their design process is an important area to study. Some studies have shown having able-bodied people experience disability simulators can reinforce negative stereotypes and attitudes to these challenged people. [Nario-Redmond et al., 2017]. It would be interesting to see how developers assess their existing UI designs using the Funkify personas, and how this might influence their designs to better accommodate more diverse challenged end users and empathise better with them after experience with these tools. A notable perspective raised by Huynh et al. was the use of personas of developers with their own challenges [Huynh et al., 2021], and how they would diagnose and fix UI problems differently to other developers. Future research could focus on using these different software engineer personas to evaluate Funkify.

7 SUMMARY

We conducted a heuristics-based evaluation of Funkify, an augmented reality-based simulator, to assertain its suitability as a tool to simulate diverse challenged end users' web experience. Four user personas (Dyslexia, Blurred Vision, Tremor, and Tunnel Vision) were selected to carry cognitive walkthroughs on three target websites (Amazon, Reddit, Commonwealth NetBank). We found the evaluated personas provided an overall good representation of users' challenges. There are notable limitations relating to specific components of the web page (modals, drop-downs...) where the tool might fail to apply the persona filter. Future work can investigate the other personas Funkify provides as well as other similar tools on a wider range of target websites. It would be good to compare the assessment of real users with target challenges to those of developers using Funkify simulations to see if the same problematic accessibility issues are detected. We also want to evaluate the simulation environment created by these tools as to their effect on eliciting developers' greater empathy towards diverse end users.

ANOWLEDGEMENTS

Grundy is supported by ARC Laureate Fellowship FL190100035.

REFERENCES

(2019). Tunnel vision: What peripheral vision loss feels like – https://www.webmd.com/eye-health/common-causes-peripheral-vision-loss.

Abascal, J., Arrue, M., and Valencia, X. (2019). Tools for web accessibility evaluation. In *Web Accessibility*, pages 479–503. Springer.

Alcántara, A.-M. (2021). Lawsuits over digital accessibility for people with disabilities are rising.

Alshayban, A., Ahmed, I., and Malek, S. (2020). Accessibility issues in android apps: state of affairs, sentiments, and ways forward. In *ICSE*. IEEE.

Bai, A., Stray, V., and Mork, H. (2019). What methods software teams prefer when testing web accessibility. *Advances in Human-Computer Interaction*, 2019.

Bi, T., Xia, X., Lo, D., Grundy, J. C., Zimmermann, T., and Ford, D. (2022). Accessibility in software practice: A practitioner's perspective. *to appear in ACM Transactions on Software Engineering and Methodology*.

Biswas, P., Robinson, P., and Langdon, P. (2012). Designing inclusive interfaces through user modeling and simulation. *Int. J. of Human-Computer Interaction*, 28(1):1–33.

Blackmon, M. H., Polson, P. G., Kitajima, M., and Lewis, C. (2002). Cognitive walkthrough for the web. In *CHI*, pages 463–470.

Davis, R. (2004). The gift of dyslexia. *Education Horizons*, 8(3):12–13.

Dorsey, E. a., Constantinescu, R., Thompson, J., Biglan, K., Holloway, R., Kieburtz, K., et al. (2007). Projected number of people with parkinson disease in the most populous nations, 2005 through 2030. *Neurology*, 68(5):384–386.

Funkify (2021). Funkify – a disability simulator for the web - https://www.funkify.org/.

Green, C., Dickinson, H., Carey, G., and Joyce, A. (2021). Barriers to policy action on social determinants of health for people with disability in australia. *Disability & Society*, pages 1–25.

Grundy, J. C. (2021). Impact of end user human aspects on software engineering. In *ENASE*, pages 9–20.

Harper, S. and Chen, A. Q. (2012). Web accessibility guidelines. *World Wide Web*, 15(1):61–88.

Huynh, K., Benarivo, J., Da Xuan, C., Sharma, G. G., Kang, J., Madugalla, A., and Grundy, J. (2021). Improving human-centric software defect evaluation, reporting, and fixing. In *COMPSAC*, pages 408–417. IEEE.

Jacko, J. A. and Sears, A. (1998). Designing interfaces for an overlooked user group: Considering the visual profiles of partially sighted users. In *ACM Conf. on Assistive technologies*, pages 75–77.

Jamieson, M., Cullen, B., Lennon, M., Brewster, S., and Evans, J. (2020). Designing appltree: usable scheduling software for people with cognitive impairments. *Disability and Rehabilitation: Assistive Technology*.

Jim, A. Y., Shim, H., Wang, J., Wijaya, L. R., Xu, R., Khalajzadeh, H., Grundy, J., and Kanij, T. (2021). Improving the modelling of human-centric aspects of software systems: A case study of modelling end user age in wirefame designs. In *ENASE*, pages 68–79.

Kamikubo, R., Higuchi, K., Yonetani, R., Koike, H., and Sato, Y. (2018). Exploring the role of tunnel vision simulation in the design cycle of accessible interfaces. In *15th Int. Web for All Conf.*, pages 1–10.

Lazar, J. (2019). The potential role of us consumer protection laws in improving digital accessibility for people with disabilities. *U. Pa. JL & Soc. Change*, 22:185.

Lee, P. P., Spritzer, K., and Hays, R. D. (1997). The impact of blurred vision on functioning and well-being. *Ophthalmology*, 104(3):390–396.

Li, C., Yu, Y., Leckning, J., Xing, W., Fong, C. L., Grundy, J., Karolita, D., McIntosh, J., and Obie, H. O. (2021). A human-centric approach to building a smarter and better parking application. In *COMPSAC*, pages 514–519. IEEE.

Nario-Redmond, M. R., Gospodinov, D., and Cobb, A. (2017). Crip for a day: The unintended negative consequences of disability simulations. *Rehabilitation psychology*, 62(3):324.

Pruitt, J. and Grudin, J. (2003). Personas: practice and theory. In *Designing for User Experiences*, pages 1–15.

Rieman, J., Franzke, M., and Redmiles, D. (1995). Usability evaluation with the cognitive walkthrough. In *Human factors in computing systems*, pages 387–388.

Rutter, R., Lauke, P. H., Waddell, C., Thatcher, J., Henry, S. L., Lawson, B., Kirkpatrick, A., Heilmann, C., Burks, M. R., Regan, B., et al. (2007). *Web accessibility: Web standards and regulatory compliance*. Apress.

Schulz, T. and Fuglerud, K. S. (2012). Creating personas with disabilities. In *Int. Conf. on Computers for Handicapped Persons*, pages 145–152. Springer.

Sierra, J. S. and Togores, J. (2012). Designing mobile apps for visually impaired and blind users. In *5th Int. Conf. on Advances in Computer-human Interactions*, pages 47–52. Citeseer.

Sirisena, D. and Williams, D. R. (2009). My hands shake: Classification and treatment of tremor. *Australian family physician*, 38(9):678–683.

Stearns, L., Findlater, L., and Froehlich, J. E. (2018). Design of an augmented reality magnification aid for low vision users. In *SIGACCESS*, pages 28–39.

Wentz, B., Pham, D., Feaser, E., Smith, D., Smith, J., and Wilson, A. (2019). Documenting the accessibility of 100 us bank and finance websites. *Universal Access in the Information Society*, 18(4):871–880.

Wery, J. J. and Diliberto, J. A. (2017). The effect of a specialized dyslexia font, opendyslexic, on reading rate and accuracy. *Annals of dyslexia*, 67(2):114–127.

Zhong, Y., Weber, A., Burkhardt, C., Weaver, P., and Bigham, J. P. (2015). Enhancing android accessibility for users with hand tremor by reducing fine pointing and steady tapping. In *12th Int. Web for All Conf.*

Zhou, S., Carroll, E., Nicholson, S., and Vize, C. J. (2020). Blurred vision. *BMJ*, 368.