Accessibility in Heuristic Evaluations

Blade Frisch

ABSTRACT

Heuristic evaluations can be used for rapid and early feedback on the usability of an interface design, but many of the current evaluation frameworks do not consider accessibility. This proposal presents a plan to develop a set of heuristic evaluations that target the accessibility of a user interface that is agnostic of implementation details. This plan involves evaluating designs at three levels of fidelity: production, high, and low, in order. A card sort is performed on these data to generate categories of issues, which will be turned into the final set of heuristics. These heuristics will then be validated by performing a heuristic evaluation on a design with known accessibility issues to determine if the heuristics are capable of identifying the issues.

BACKGROUND

Heuristic Evaluation

The Interaction Design Foundation defines a heuristic evaluation as a process where experts use a set of guiding principles and rules to evaluate the usability of an interface design [7]. This process typically involves gathering a small number of experts who will walk through the functionality of an interface design and compare that functionality to a set of heuristics that act as guidelines for what the desired functionality is. The design could be a low-fidelity prototype, high-fidelity prototype, or a final product. Jakob Nielsen explains that this process is used because no one individual could ever catch all the usability issues in a design and that this process can occur at different iterations of the design processes because the evaluator is not actually completing a task [3].

Existing Heuristic Frameworks

The key to performing a heuristic evaluation is the set of heuristics chosen for the task. These heuristics will guide the evaluation and can be tailored to target different goals. One popular set of heuristics is the Nielsen-Norman 10 usability heuristics for user interface design [4]. Originally developed by Jakob Nielsen in 1994, these heuristics are designed to evaluate the general usability of an interface. While certainly popular, there are other sets of heuristics available. Amélie Boucher created a set of heuristics that target simplicity and ease of use in an interface design [8], while Arhippainen created a set of heuristics that focus on the respecting the user's choices in how they interact with the interface, even if it means offering more options so the user can choose how they want to interact with the interface design [1]. Bastien and Scapin provide a set of detailed heuristics grounded in human factors and ergonomic psychology [9] while Shneiderman distills his experience down to eight heuristics he dubs "golden rules of interface design" [10].

ACCESSIBILITY

We can see there are a variety of heuristics available to choose from when validating a design, but there is one key area overlooked by many of these frameworks: accessibility. Boucher lists accessibility as one heuristic in the framework and Kaniasty's CARMEL guidelines [11] provide technical suggestions for accessibility, but Boucher does not provide insight into what accessibility means and Kaniasty's accessibility guidelines, while providing examples of what to look for in making a design accessible, are more targeted towards the technical and programmatic implementation of design.

In order to continue discussing accessibility, a definition must be provided. The US Government's Digital Accessibility website defines accessibility as "…usability for people who interact with products differently" [12]. Merriam-Webster's dictionary gives definitions including "capable of being used or seen", "capable of being understood or appreciated", and, most saliently, "easily used or accessed by people with disabilities OR adapted for use by people with disabilities" [6]. We will combine these definitions and narrow the scope to design, creating the working definition: a design that is capable of being used and understood by people with a range of abilities and disabilities and using a variety of tools to interface with the design. This definition allows us to evaluate how usable a design is as well as account for the different ways a disabled person could interface with the system, such as using a screen reader, eye tracking, or single-switch access.

It is critical that designs of software as well as other forms of digital media are accessible. In the US, it is required by Section 508 of the Rehabilitation Act to make digital media accessible [13:508]. As

such, the government provides guidance for making various forms of digital media (e.g., front end, user experience, content, etc.) accessible [14]. This guidance comes from the US Access Board, which also provides accessibility guidelines for more than software [15]. Implementation guidelines have been provided by Google and Apple for Android [16] and iOS [5] development to comply with these laws, and the W3C Consortium has created the Web Content Accessibility Guidelines [17] to make the web accessible. To follow these guidelines, the perspectives of disabled people must be centered in the design process [18].

However, a nondisabled person is not capable of creating heuristics to evaluate such aspects of a design. Disabled people must be included in the process of creating these heuristics in a central role, as only their lived experiences and expertise can adequately inform these heuristics. This aligns with the adage “nothing about us without us”, which is a rallying call from the disability rights movement to ensure that disabled people are meaningfully involved in any choices that are being made of their well-being [2]. Accessibility guidelines and heuristics must come from the community in order to serve the community.

As such, this proposal presents a methodology for developing a set of heuristics to evaluate the accessibility of a design. There is a need for accessibility to be present in a heuristic evaluation and these heuristics must come from the disabled community. However, the disabled community is far from a cohesive monolith, so no truly universal set of heuristics can be created, and the heuristics chosen for a specific evaluation should reflect the needs and wants of the target community or communities over attempting to make a universally accessible design.

METHODS

Recruitment

Because the disabled community is not a cohesive whole and each disability comes with its own set of unique characteristics, people with a wide range of disabilities will need to be recruited. While no set of heuristics will capture all possible accessibility issues nor even all the accessibility issues encountered by a single population, leveraging the experiences of a wider range of people will help make a more generalizable set of heuristics to identify issues that affect several populations. People with a physical disability, mobility disability, and/or cognitive disability should be recruited in this proposed methodology.

Cognitive Walkthroughs

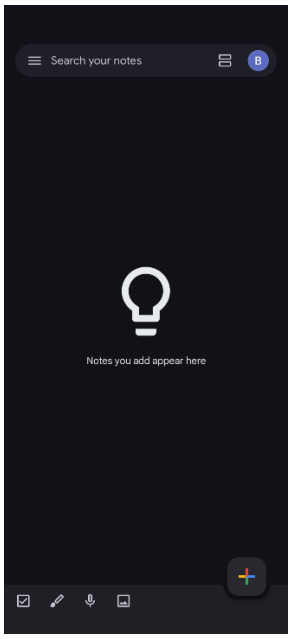
Cognitive walkthroughs will be used as the tool for identifying accessibility issues. The participants in the study will be given a design and a set of tasks to complete using the design. Using a cognitive walkthrough will allow designs to be evaluated without the need to create functional prototypes, which will also allow the focus to remain on the designs themselves rather than the technical implementation of the designs. A progressive approach will be used, starting with production-level designs and moving to low-fidelity prototypes. At each stage, the same cognitive walkthrough and tasks will be used.

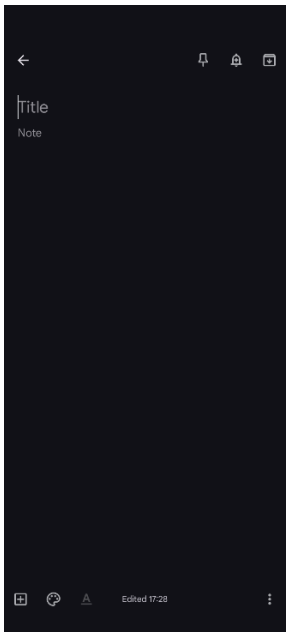
Production App

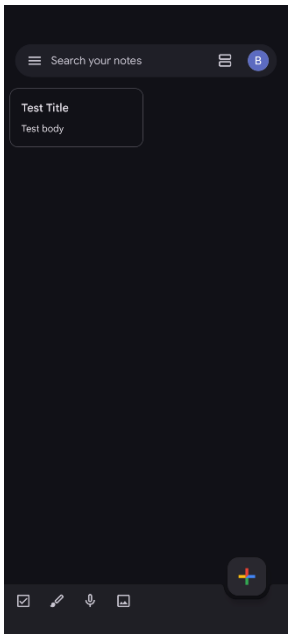
The study will begin with a design that is production ready. A note taking app will be used here as an example, such as the app shown in Figures 1 through 6. The participants will be asked to create a note, edit an existing note, and delete a note using this design. While this is a fairly trivial app, it is a useful example of a commonly used app. This methodology can, and should, be used with more designs to create the set of heuristics, which is discussed in the generalization section below. As the participants complete

the tasks, they will be asked to note any accessibility issues they uncover in the process, describing the issue and the details of the design that led to the issue separately. Each participant will provide notes for each of the three tasks, generating a dataset of accessibility issues in the design.

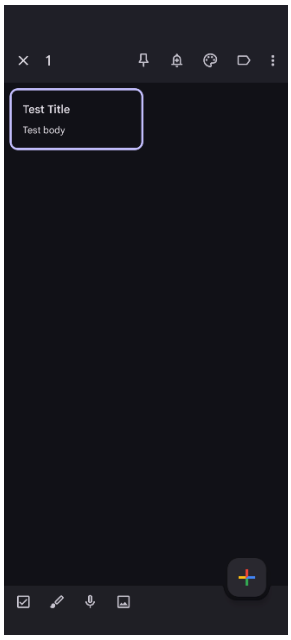
These data will then be analyzed using a card sort. The participants will be given a set of cards where each card will have the issue description and a second set with the design details sequentially. The participants will be asked to take the set of cards and organize them into groups, which they will then label and describe. This process will allow for categories of accessibility issues to emerge from the raw data. These categories will act as the first pass at a set of accessibility heuristics and be refined in the next step of the evaluation process.

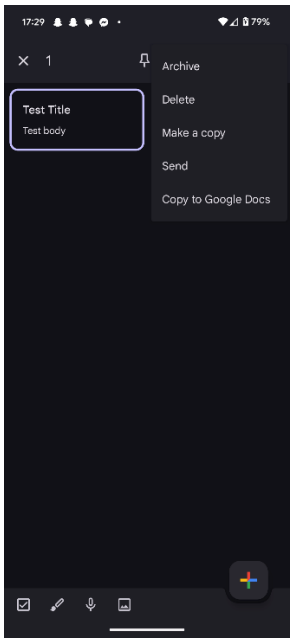
(1) 

(2) 

(3) 

(4) 

(5) 

(6) 

Figures 1 through 6, which show the designs of the note taking app Google Keep screens for creating, editing, and deleting a note.

High-Fidelity Prototype

Once the first pass of the heuristics has been generated, the evaluation process will be repeated with a high-fidelity prototype of a design, such as the one in Figures 7 through 11. Moving to a high-fidelity prototype from the production design will remove some of the finer implementation details and present things in a more abstract sense. This will also allow for the issues to be more removed from the technical implementation level and be more rooted in the design itself. The note taking app will be reused for this task as well. However, this could introduce a familiarity bias in the data. Since the participants have already worked with the production design, they will already have an idea of the accessibility issues in the

design. This is acceptable in this proposed methodology as it is focused on generating a set of set of implementation agnostic heuristics, and, as such, using the same app will allow of the heuristics to become more removed from the implementation as the designs also become more removed from the implementation.

During the evaluation, the participants will be asked to complete the same tasks: create a note, edit an existing note, delete an existing note. Like the first step, they will be asked to take note of any accessibility issues they encounter and provide a description of the issue and the design feature(s) that led to the issue separately. When all the data have been collected, the participants will then perform a card sort on these new issue descriptions and generate groups. These new groups will be compared to the groups previously generated once all evaluations have been completed.

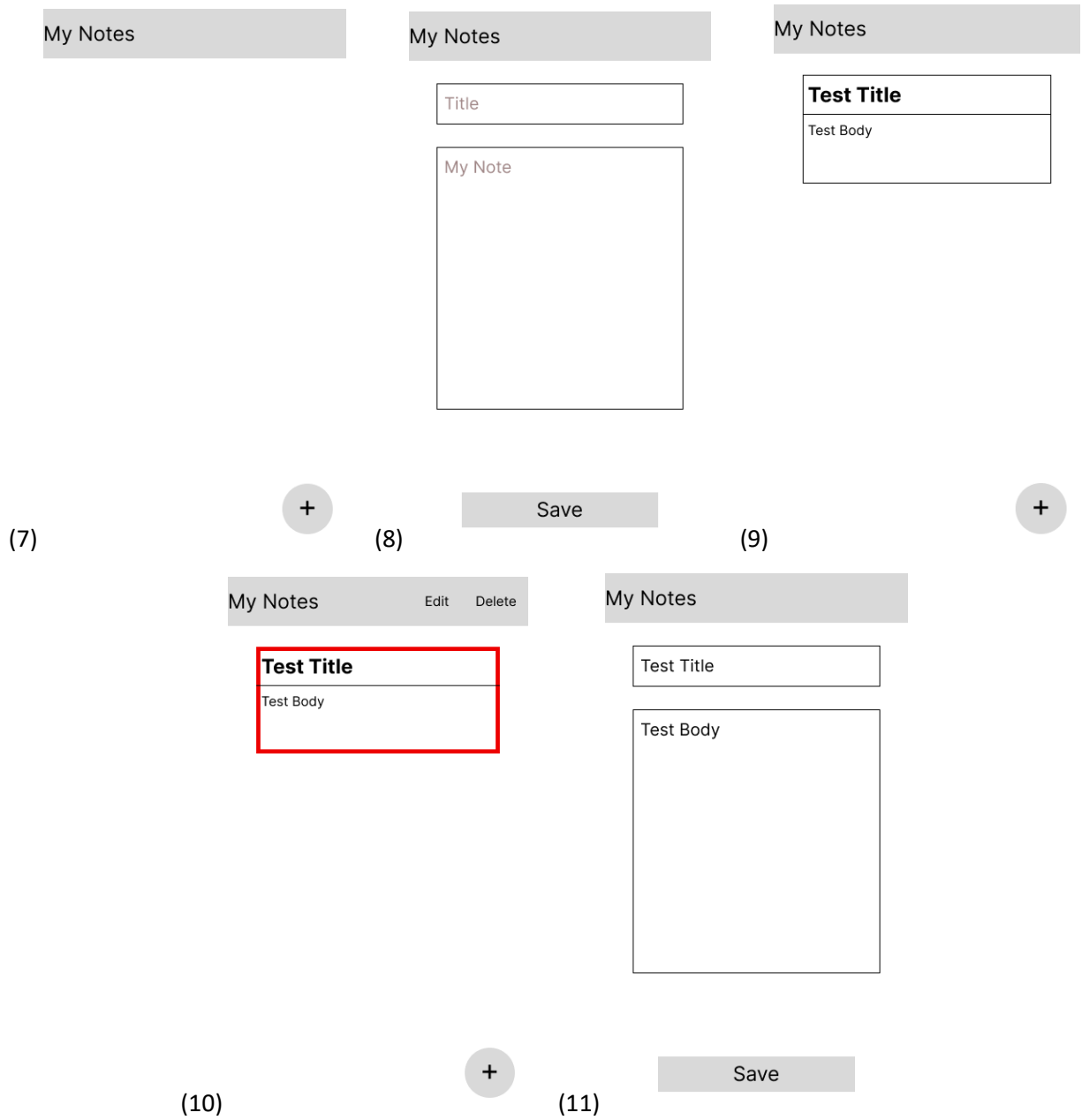
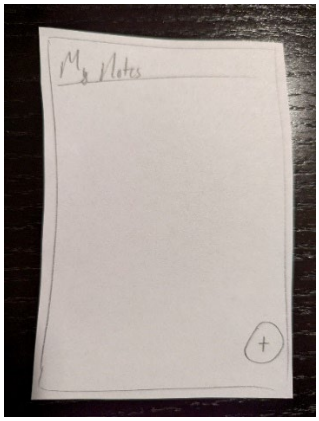
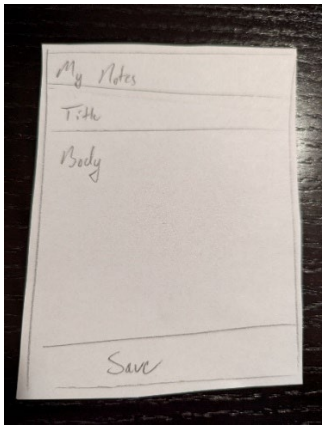


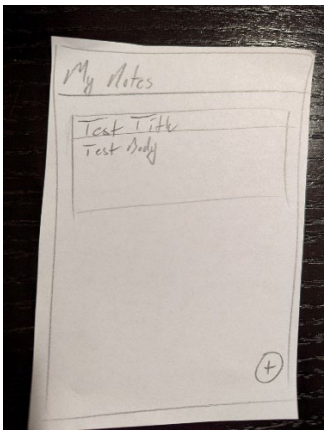
Figure 7 through 11, which show a high-fidelity prototype of a note taking app.

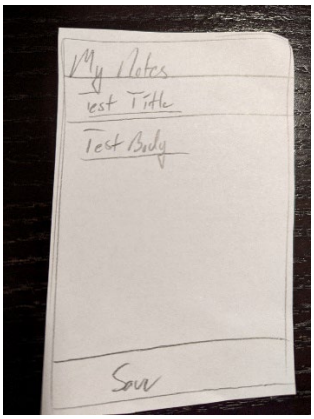
Low-Fidelity Prototype

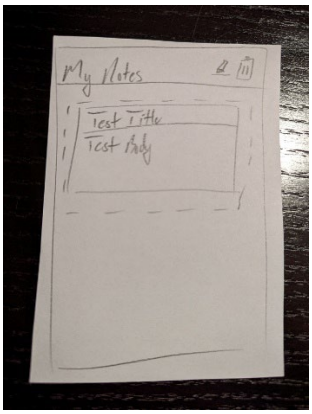
In this final step, a low-fidelity prototype of the note taking app will be evaluated, such as the one shown in Figured 12-through 16. This will be in the form of a paper prototype. The same process will be repeated, where the participants will be asked to make a note, edit and existing note, and delete a note. The accessibility issues will be described, and a card sort will be performed. In this step of the process, the design is far removed from implementation and therefore the generated issues will be in the most general and abstract form. The data collected from the previous reviews can provide the context for understanding these general forms, while the general forms will act as the basis for the final set of heuristics.

(12) 

(13) 

(14) 

(15) 

(16) 

Figures 11 through 15, which show a low-fidelity prototype of a note taking app.

Generalization

The proposed methodology above uses a simple note taking app as the example app to illustrate the process. This methodology should be repeated on multiple designs in order to create a generalized set of heuristics rather than a set of heuristics tailored to a note taking app. The designs evaluated should also cover a variety of app categories, such as note taking, form submission/data collection, browsing, and e-commerce to provide some examples. Once multiple designs across multiple categories have been evaluated and sets of issues have been identified, these data can be used to make the final set of heuristics.

Once this process has been completed for all the desired designs, there will be three sets of data for each design: categories for the production prototype, the high-fidelity prototype, and the low-fidelity prototype. One final card sort will be performed on these three sets of categories. Since the users performed the same set of tasks across each design, there will most likely be underlying themes across the three datasets, allowing for a final set of categories to emerge from these data. These final categories will be the final set of heuristics. These final heuristics will be discussed with the participants to ensure that they are reflective of their experiences.

VALIDATION

Now that a set of heuristics has been generated, they will need to be validated. To validate these heuristics, a new design will be presented to the participants. This design will be a production-ready design

with known accessibility issues placed into the design. These issues will be based on the set of heuristics and should be intentionally placed into the design without being exceedingly obvious. A domain from the exploration steps (e.g., note taking, data collection, e-commerce) or a novel domain could be used for the validation. The participants will be given the design and the final set of heuristics and be asked to perform a heuristic evaluation on the design.

As an example, a navigation app design could be given. The participants will then be asked to search for a location (e.g., a specific restaurant), get directions from their desired starting location, and add a stop to the route. These are common tasks for a navigation app and, as such, should be evaluated for accessibility. The participants will use the set of heuristics to determine the accessibility of the designs, noting the places where accessibility fails, and which heuristic(s) identify the issue. Once all the participants have completed their evaluation, the number of correct identifications, missed identifications, and false identifications can be tallied. Any false identifications should be evaluated and discussed with the participants after the fact to determine if they were truly false identifications or if the heuristics have a gap in the issues they can identify.

CONCLUSION

Heuristic evaluations are a method for rapidly evaluating the usability of a design and can be performed at various stages of the design process. There are several existing sets of heuristics, but few address accessibility issues and none focus on evaluating accessibility. Accessibility is something that all designs should be concerned about, and it can even be mandated by law. There are technical guidelines for implementing accessible designs but there is a need for implementation agnostic heuristics to evaluate designs themselves. A methodology is proposed where a cognitive walkthrough is performed at various stages of the design process and disabled participants identify and describe the accessibility concerns in each design. A card sort is performed on these issues to generate categories of accessibility issues that will be synthesized into a set of heuristics. These heuristics will then be validated by performing a heuristic analysis on a design with known accessibility issues to evaluate how well the heuristics identify the known accessibility issues.

REFERENCES

[1] admin. 2015. Ten User Experience Heuristics | Arhippainen.fi. Retrieved February 21, 2024 from http://arhippainen.fi/user-experience-heuristics/ten-user-experience-heuristics/

[2] James I. Charlton. 1998. *Nothing About Us Without Us: Disability Oppression and Empowerment*. University of California Press, Berkely, California.

[3] World Leaders in Research-Based User Experience. The Theory Behind Heuristic Evaluations, by Jakob Nielsen. *Nielsen Norman Group*. Retrieved February 20, 2024 from https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/theory-heuristic-evaluations/

[4] World Leaders in Research-Based User Experience. 10 Usability Heuristics for User Interface Design. *Nielsen Norman Group*. Retrieved February 21, 2024 from https://www.nngroup.com/articles/ten-usability-heuristics/

[5] Apple Inc. Accessibility. *Apple Developer*. Retrieved February 20, 2024 from https://developer.apple.com/accessibility/

[6] 2024. Definition of ACCESSIBLE. Retrieved February 21, 2024 from https://www.merriam-webster.com/dictionary/accessible

[7] What is Heuristic Evaluation? *The Interaction Design Foundation*. Retrieved February 20, 2024 from https://www.interaction-design.org/literature/topics/heuristic-evaluation

[8] Amélie Boucher's Ergonomic Criteria – Capian. Retrieved February 21, 2024 from https://capian.co/amelie-boucher-ergonomic-criteria

[9] Bastien & Scapin Ergonomic Criteria for the Evaluation of Human-Computer Interfaces – Capian. Retrieved February 21, 2024 from https://capian.co/ergonomic-criteria-bastien-scapin

[10] Ben Shneiderman. Retrieved February 21, 2024 from http://www.cs.umd.edu/~ben/goldenrules.html

[11] Kaniasty's CARMEL Guidelines - Heurio. Retrieved February 21, 2024 from https://www.heurio.co/kaniastys-carmel-guidelines

[12] Getting started | UX design | Accessibility for Teams. Retrieved February 21, 2024 from https://accessibility.digital.gov/ux/getting-started/

[13] Section508.gov. Retrieved February 21, 2024 from https://www.section508.gov/manage/laws-and-policies/

[14] Accessibility for Teams. Retrieved February 22, 2024 from https://accessibility.digital.gov/

[15] U.S. Access Board - U.S. Access Board - Home. Retrieved February 22, 2024 from https://www.access-board.gov/

[16] Make apps more accessible | App quality | Android Developers. Retrieved February 20, 2024 from https://developer.android.com/guide/topics/ui/accessibility/apps

[17] Web Content Accessibility Guidelines (WCAG) 2.1. Retrieved February 20, 2024 from https://www.w3.org/TR/WCAG21/#:~:text=Web%20Content%20Accessibility%20Guidelines%20(WCAG)%202.1%20defines%20how%20to%20make,%2C%20learning%2C%20and%20neurological%20disabilities.

[18] Humanize accessibility | UX design | Accessibility for Teams. Retrieved February 21, 2024 from https://accessibility.digital.gov/ux/humanize-accessibility/