Examining Visual Semantic Understanding in Blind and Low-Vision Technology Users

Venkatesh Potluri, Tadashi E Grindeland, Jon E. Froehlich, Jennifer Mankoff
<vpotluri,tadg99,jonf,jmankoff>@cs.uw.edu
Paul G. Allen School of Computer Science and Engineering, University of Washington
Seattle, WA, USA

ABSTRACT

Visual semantics provide spatial information like size, shape, and position, which are necessary to understand and efficiently use interfaces and documents. Yet little is known about whether blind and low-vision (BLV) technology users want to interact with visual affordances, and, if so, for which task scenarios. In this work, through semi-structured and task-based interviews, we explore preferences, interest levels, and use of visual semantics among BLV technology users across two device platforms (smartphones and laptops), and information seeking and interactions common in apps web browsing. Findings show that participants could benefit from access to visual semantics for collaboration, navigation, and design. To learn this information, our participants used trial and error, sighted assistance, and features in existing screen reading technology like touch exploration. Finally, we found that missing information and inconsistent screen reader representations of user interfaces hinder learning. We discuss potential applications and future work to equip BLV users with necessary information to engage with visual semantics.

CCS CONCEPTS

• **Human-centered computing → Accessibility theory, concepts and paradigms**; **Empirical studies in interaction design**.

KEYWORDS

Accessibility, visual design, blind and low-vision creators

ACM Reference Format:
Venkatesh Potluri, Tadashi E Grindeland, Jon E. Froehlich, Jennifer Mankoff. 2021. Examining Visual Semantic Understanding in Blind and Low-Vision Technology Users. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3411764.3445040

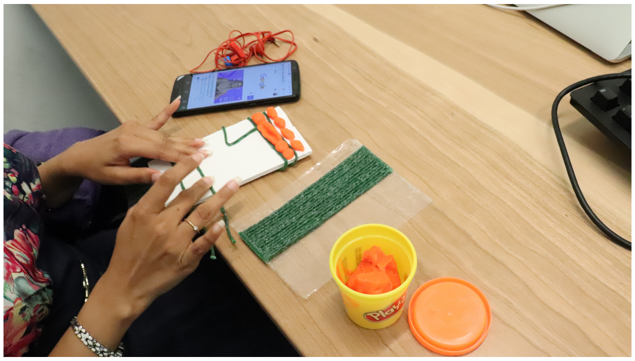


Figure 1: The image above shows a participant using Wikki Stix (green) and Play-Doh (orange) to reconstruct the smartphone web Google homepage.

1 INTRODUCTION

User interfaces (UIs) intermix interactive widgets, text, and blank space, all carefully designed and positioned to convey meaning and afford use. For screen reader users, the size, shape, spatial layout, and perceptual affordance of a UI fundamentally differs from sighted users. Though powerful, screen readers linearize access to interactive widgets and content, making it difficult for blind and low vision (BLV) users to interact with the visual semantics of UIs.

In this paper, we explore whether access to these visual semantics is important to the BLV community and, if so, why. Though a rich literature examines how to enhance BLV access to and understanding of spatial information, it focuses on specialized interfaces or tasks, such as maps and navigation [2], access to diagrams and graphs [19, 32, 42] and touchscreens [36, 38]. Our focus is complementary: investigating how BLV users identify and interact with spatial characteristics of modern UIs, such as the size, shape, and relative and absolute position of widgets and content. Because screen readers often function differently across device platforms–from conventional computers like desktops and laptops (linear access) to touchscreens (direct + linear access) and native applications *vs.* websites–we examine how these differences change BLV users' perception and understanding of UIs.

Our research builds on recent work by Li *et al.* [44], who conducted a formative study of seven BLV users to explore how they engage in layout design (*e.g.*, to make slides in PowerPoint and create websites with WordPress). They incorporated their findings into an accessible layout design tool that uses auto-generated tactile

---

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445040

sheets overlaid on a tablet to enable the editing of HTML templates. Our focus is broader: we are interested not only in how BLV users engage in creative design tasks that require spatial understanding, but also in how they perceive and interact with spatial information in modern interfaces in their everyday use. We address two primary research questions:

- When, how, and why do BLV technology users want to interact with the visual semantics of UIs? How do they interpret visual semantics of UIs? Relatedly, do these users even *want to* learn about these semantics?
- What does prototyping an interface reveal about their understanding of the visual semantics of UIs? How do lo-fi prototypes created by BLV users and the thought process of creating these prototypes, *e.g.*, think-aloud comments, inform about their understanding of visual semantics?

To address these questions, we conducted a three-part qualitative study with ten BLV participants. First, we asked participants about their knowledge and use of visual semantics in UIs. We then probed participants' understanding when using smartphone apps with screen readers. Finally, we explored participants comprehension and perception of visual semantics in desktop and mobile web interfaces through interview questions and a "think aloud" lo-fi prototyping exercise. Here, participants reconstructed a UI of their choice using lo-fi craft materials and described the underlying visual semantics (Figure 1).

We collected and analyzed questionnaire data, observational notes, participant-generated artifacts, and study session audio and video recordings. Our study is the first to capture nuanced qualitative information of how BLV users understand visual semantics across device platforms and common usage scenarios and probes BLV users to express their understanding of UIs through prototyping.

Our research findings suggest that most BLV users *want* to understand the visual semantics of interfaces in mobile app contexts. Participants, while generally disinterested in the visual semantics of websites, still emphasized the importance of such semantics. Informed by these findings and the literature, we enumerate design recommendations for technology designers that would enable BLV technology users to better understand and engage more effectively with the visual semantics of web layouts. For example, we discuss integration of visual descriptions of interfaces into existing screen readers through a verbosity mode similar to settings currently available for punctuation. In summary, we make the following research contributions:

1. We demonstrate that BLV technology users already understand concepts of spatiality and visual semantics, especially on phones rather than desktops and use this understanding to navigate and collaborate.
2. We find that BLV users' perceptions of visual semantics (*e.g.* size, shape, location) are strongest for smartphone apps. These findings may be a result of device affordances and screen reader features but design implications extend to both platforms.
3. We find that BLV users currently rely on trial and error to discover information about visual semantics, and that mobile phone screen readers are particularly effective at supporting this exploration.
4. We make recommendations for technology designers to support BLV access to visual semantic information.

Our findings will help screen reader developers improve capabilities of both desktops and laptops, and touchscreen screen readers. To support the more efficient learning of visual information, we must address gaps in current tools (such as missing information or screen reader deficiencies) and pay attention to design elements (such as reference points), that helped our BLV participants create lo-fi prototypes. Our findings can inform the way that screen readers describe interfaces and the interaction options that they support, as well as how we support tasks that require spatial information.

2 BACKGROUND AND RELATED WORK

To better contextualize visual semantics and non-visual computer access, we discuss the interaction paradigms that are standard in desktop and smartphone (touchscreen) interaction and then enumerate prior work on screen reader interactions, access to visual information, and the creation of accessible visual layouts. As the interaction paradigms supported by screen readers vary significantly with platform and application, we encourage unfamiliar readers to try them. Screen readers are freely available for Windows[49], OSX [5], and both Android [29] and iPhone [4] smartphones.

2.1 Standard Screen Reading Access to Desktops

Screen readers convert visual semantics and content into a linear, ephemeral stream of output (typically, audio or Braille). They can also convey hierarchical information (such as the interactor hierarchy of an interface), tabular information (such as the title and contents of a table) and semantics (such as non-speech audio representations of content type, *e.g.*, beeps and tones, or spoken identification of a heading level). VoiceOver [7] and NVDA [53] convey interactor hierarchy by presenting the UI as a list of interactable elements as well as through a separate review cursor that uses the underlying accessibility tree exposed by Windows APIs, respectively. Likewise, several approaches have been explored to convey notions like tables. Screen readers allow non-linear access to tables by providing navigation capabilities horizontally (along the row) or vertically (along the column). Furthermore, they provide announcements that help users identify reference points in a table, *e.g.*, row and column headers if they are semantically marked up [34, 35].

Though these approaches enable access, table navigation remains problematic for screen reader users [33, 69]. Complementary to this work [33, 69], Williams *et al.* [68] investigate advantages of representing web interfaces as tables and find that participants rated their table navigation experience more positively in terms of effort, memorization, ease of navigation, understanding of page information, and confidence in submitted answers. To improve information access, Khurana *et al.* [39] explore non-linear interaction techniques on the keyboard to enable BLV users to navigate tables and webpages spatially. They show improved task completion times in information-seeking tasks by BLV users in a lab study. User expertise also impacts which features are used during non-visual access. For example, studies of beginning users highlight challenges with understanding visual context, such as when the cursor has switched to a new application [8].

2.2 Touchscreen and Smartphone Access

Unlike traditional computer interfaces, touchscreens are inherently more spatial because they are designed for *direct* rather than *indirect* interaction (*i.e.,* touch *vs.* mouse). On touchscreens, visual semantics, like the size, shape and spatial arrangement of interactors are fundamental to their use. Touchscreen-based screen readers also reflect this shift. Both Apple's VoiceOver [7] and Google's TalkBack [30] enable screen reader users to directly access elements spatially under their touch and to also interact indirectly using gestures.

In seminal early work, Kane *et al.* [36] presented design recommendations and gestures for non-visual touch interactions, many of which can be seen in mainstream screen readers today. In follow-up work, Kane and colleagues performed a gesture elicitation study with both sighted and blind participants to co-design usable gesture sets [38]. Their findings categorize touchscreen interaction techniques into menu browsing, discrete gestures, and fixed regions. They further inform quantifiable measures of gestures, differentiating between sighted and BLV individuals. Their work [37] expands this knowledge to larger touchscreens and demonstrates three access overlays: edge projection, neighborhood browsing, and touch-and-speak. Despite these advances, screen reading remains predominantly linear and ephemeral.

While prior work on touchscreen accessibility investigated how BLV users interact with spatial layouts, these explorations have been limited to accomplishing interactions and information seeking. Our work explores visual semantics beyond this specific scope.

2.3 Tasks That Require Spatial Information

While traditional GUIs require generalizable access to a wide range of interfaces in webpages and applications, some inherently spatial tasks, such as image and graph exploration, cannot be supported by screen readers. Many of these tasks have been addressed using specific, non-generalizable techniques. However, we can still learn much about how to best support visual semantics by studying what has been done in other domains. Work in the context of BLV access to visual information includes picture books [65], picture and scene descriptions through apps like *SeeingAI* [24] and *Aira* [22], graphs and graphics [19, 32, 42], cross-word puzzles [60] and work in the domain of wayfinding (*e.g.*, digital interfaces [2, 9, 67], tactile maps [26, 63], and environment exploration [15, 28, 50]).

Of these, wayfinding has received the most attention, perhaps because it is so central to independence. Wayfinding is also highly relevant to our work since considerable attention has been paid to studying the impact of maps and other navigation tools on spatial understanding (*e.g.*, [27]). For example, one survey found that haptic imagery (*i.e.*, mental representations generated on the basis of previous haptic experience) can be almost as accurate as one based on visual imagery in many different cognitive tasks [21]. Relatedly, onset of blindness does not impact spatial memory ability [46], and spatial memory can be acquired independent of visual perceptual abilities. In a study of sequential representations of environments based on step-by-step actions and points of interest, BLV users were able to build mental representations of new places [27]. This work shows that BLV users can create mental models of spatially organized information through sequential representation.

BLV exploration of visual semantics of UIs differs from exploration of maps since the alternative representations of user interfaces between BLV and sighted users differ significantly, depending on semantics, and yet BLV users can fully utilize accessibly designed interfaces. Unlike touchscreen screen reader interactions that support dynamic exploration of interfaces, tactile modalities to explore maps do not permit exploration of dynamic content.

2.4 Accessing and Understanding Visual Information

Apart from maps and graphics, most prior work in accessible visual information for BLV individuals focuses on images [13, 47, 52, 57]. While this work explores access and rich representations of visual information, the BLV individual is limited to the role of a consumer of visual information provided by sighted producers or describers. Furthermore, this information seldom addresses the aesthetic value of producing and consuming these artifacts. Likewise, deployed human visual interpreter services like Aira [22] and automated ones like SeeingAI [24] are limited by policy and technology to describing the aesthetics of the visual environment around us. For example, Aira agents cannot offer opinions (personal communication) and SeeingAI is focused on functional information. Thus, visual semantics and UI appeal have remained inaccessible to BLV technology users.

Many efforts related to visual information focus on consumption and do not discuss creation without help by sighted users to fill in accessibility gaps (*e.g.* [40, 55, 66]). For example, Kuber *et al.* [40] found that BLV and sighted users working together could form a mental model of UIs. Some of this may be due to the tools available: Li *et al.* found that BLV users chiefly rely on sighted assistance to create visually appealing layouts [44]. Other work has explored BLV users as experts, who provide instructions to sighted users. To facilitate collaborations between BLV and sighted users, Bigham *et al.* [14] present a tool that helps create annotated instructions for task completion on webpages. Similarly, [61] surfaces the communication difficulty between BLV and sighted users in virtual interfaces. They focus particularly on deictic directions used in commands, which proved problematic to BLV users.

Related to visual layout creation, emerging work explores BLV engagement in design through workshops exploring ideation [11], toolkits to facilitate BLV participation in design [10], and re-imagining co-design with BLV people [18]. Recent work also explores how to enable BLV individuals to independently create visual layouts and tactile artifacts like 3D models and graphics. *ShapeCad* [64] uses a 2.5D display to enable BLV users to complement the programming of visual layouts. Other efforts have focused on enabling editing of visual layouts with non-tactile interfaces [16, 59]. Lastly, Potluri *et al.*present a vision for semi-automatic visual design where AI and BLV users work together to create user interfaces [58]. Though this body of work addresses the important issue of engaging BLV users in design, it does not specifically discuss visual semantic understanding, and the role it plays in design.

Building on prior work, we are left with several open questions: How and why do BLV users use visual semantics in everyday interactive computing tasks? What is their current visual semantic understanding? How can better visual semantics descriptions help BLV users use computers, engage in design work, and more easily collaborate with sighted users? We begin to address each of these in our study.

3 STUDY

To investigate how BLV users understand, learn, and potentially use visual semantics in interfaces across devices and contexts, we performed a three-part interview study with 10 BLV participants: in Part 1, we examine perceptions, usage, and the importance of visual semantics via formative questions; in Part 2, we observe how BLV participants perform everyday tasks with smartphone apps, smartphone webpages, and desktop webpages and the role of visual semantics therein; finally, in Part 3, we assess how BLV users interpret and interact with visual semantics via lo-fi UI reconstructions.

3.1 Method

3.1.1 Participants. We recruited ten BLV participants (four women; six men) through email, social media, and snowball sampling. As summarized in Table 1, participants were on average 35 years old (*SD*=11.18, range 24-58) and compensated $15 per hour, for a total of $22 for our 90-minute study.

3.1.2 Procedure. The study took 90 minutes to complete and included three parts: (1) a semi-structured formative interview, (2) smartphone and desktop-based screen reader tasks, and (3) UI reconstructions using lo-fi prototyping. Study sessions were conducted by two researchers (one facilitator, one note taker) and video recorded.

Part 1: Formative interview. We began with a semi-structured interview to examine how BLV users think about, make sense of, and use visual semantics in everyday interactions on smartphones and desktops (*i.e.* with traditional GUIs). We also asked about the perceived importance and motivations for learning such semantics.

Part 2: Screen reader tasks. To investigate how BLV users rely on visual semantics in their UI interactions and to understand usage across contexts, we asked participants to perform common tasks using *smartphone apps*, *smartphone webpages*, and *desktop webpages*. For each device platform and usage scenario, we requested that participants choose a task that they perform every day using a screen reader, such as sending a text message, requesting a ride share, or planning a route. We also asked participants to perform two prescribed tasks: *adding a contact* (smartphone app only) and *searching for a video in YouTube* (all three contexts). During the tasks, we asked them to 'think aloud,' and we directly observed their interactions (*e.g.,* use of gestures and spatial patterns on the smartphone). For these interactions, participants could use screen readers or voice assistants (*e.g.,* Siri). Importantly, these tasks were intended as a probe to evoke reactions with regards to visual semantics in everyday computer and smartphone use and not meant to derive specific findings related to task completion–something that has been explored in prior work on web browsing [12, 66], programming [3], and digital visual layout creation [44].

Part 3: UI reconstructions. Finally, to better understand how BLV users interpret and build mental models of visual semantics, we asked participants to reconstruct a UI of their choice for each context using lo-fi prototyping materials. Specifically, participants built UI reconstructions using pre-prepared poster board templates (cut in the size of a phone and laptop computer) along with Play-Doh, Wikki Stix, and Braille labels. We also asked them to reconstruct their own websites (for those that had them, *N*=4).

3.1.3 Data and Analysis. We collected interview session audio-video recordings, researcher notes, and participant-created prototypes. Video recordings were manually transcribed by the research team and analyzed using an iterative thematic coding approach with a mixture of inductive and deductive codes [17]. The unit of analysis was a segment of video containing either a participant comment or an observational note made during initial transcription.

To begin, we created an initial codebook derived from our study protocol. For each unit, we recorded the context (*e.g.*, the question that was asked and the location in the interview), timestamp, an initial code, and a participant identifier. To refine the codebook, six participants were randomly selected and re-coded by two researchers (three participants each, no overlap). The codebook was shared and mutually updated continuously. Using the updated codebook, 10% of the data was randomly selected and coded independently by the two researchers. To calculate inter-rater reliability, we used Cohen's Kappa, which resulted in κ=0.7. The researchers then met, resolved disagreements to consensus, and updated the codebook (where necessary). Finally, one researcher recoded all of the data using the updated codebook.

4 FINDINGS

We describe findings from our three-part qualitative study, including the perceived importance, understanding, and use of visual semantics across smartphone and desktop contexts. As a qualitative study, we are interested in capturing nuanced views and perspectives; however, we report numbers to indicate participant preferences and trends.

4.1 Importance of and Access to Visual Semantics

Towards addressing our initial research question about *when*, *how*, and *why* do BLV users interact with visual semantics, we report on perceived importance, touch *vs.* desktop interface use, and strategies and challenges to learning.

4.1.1 Importance of Visual Semantics. Despite not having full access to a UI's visual semantics, all participants felt that they were critically important, particularly to: (1) improving interaction with and navigation of UIs (*N*=6), (2) enabling collaboration with sighted users (*N*=6), especially to provide and receive instruction, and (3) building their own visually oriented artifacts (*N*=2), such as websites and blogs.

For UI navigation, BLV users felt that visual semantics improved interaction efficiency, especially on smartphones. For example, when discussing their smartphone, P8 said, *"For one thing, I know where some of the icons on the screen are. So, I can just touch the area where I believe it is."* (P8). Similarly, P6 noted, *"Spatially, if you know something is at the top vs. at the bottom, it helps you get to it."* (P6). These spatial connections are enabled by the touchscreen screen

Table 1: Demographic and technology use details of the participants. Note that P7-LV and P10-LV are both low vision which is indicated in their participant ID since it may be relevant to interpreting quotes from these participants. *Light* refers to light perception; *Prog* refers to programming experience (this data is missing for P1).

ID	Age	Gen	Level of Vision	Light	Frequently Used Access Tech	Occasionally Used Access Tech	Prog?	Web Presence
P1	30	F	Totally blind	No	JAWS, VoiceOver on OS X, VoiceOver on iOS	NVDA		None
P2	27	M	Totally blind	No	NVDA, Braille Display, VoiceOver on iOS	JAWS, Windows Narrator	Yes	Personal website
P3	28	F	Totally blind	No	JAWS, VoiceOver on iOS	NVDA, Braille Display	No	Personal website
P4	35	M	Visual acuity not measurable	Light	JAWS, VoiceOver on OS X, Braille Display, VoiceOver on iOS	NVDA	No	None
P5	24	M	Totally blind	No	JAWS, VoiceOver on iOS	NVDA	Yes	Personal website
P6	58	M	Extremely low	Light	JAWS, Braille Display, VoiceOver on iOS	Windows Narrator	No	None
P7-LV	49	F	Double and blurry vision	Light	JAWS, Screen Magnification, VoiceOver on iOS, ZoomText	NVDA	No	Personal blog
P8	27	M	Totally blind	Light	VoiceOver on OS X, Braille Display, VoiceOver on iOS	JAWS, Windows Narrator, NVDA	No	None
P9	26	F	Light perception	Light	JAWS, NVDA, TalkBack on Android	None	Yes	Personal blog
P10-LV	46	M	Peripheral; no center vision	Yes	VoiceOver on OS X, Screen Magnification, VoiceOver on iOS, CCTV	None	No	None

reader, which offers direct access to widgets and information via touch. As an extreme case, P9 used only direct touch interactions on her phone: *"I only tap; I don't use gestures."* (P9).

When collaborating with sighted individuals, BLV individuals often find themselves receiving and providing instructions about visual artifacts, *e.g.*, how to find a button or feature in an application. This information exchange typically requires both sighted and BLV individuals to share an understanding of the overall UI as well as the current element in focus. P9 described the dual challenge of giving and receiving instructions with sighted users:

> *"It's two things. One, when I am trying to convey something to someone [sighted], and a lot of times the buttons that we hear are only labels, and it's not actually written [as a] home button [or] maybe a home icon. [. . .] And the second reason why I'd love to know how it's laid out is when other people are giving me directions. The other day I was using Microsoft Teams, and I wanted to figure out how to share my screen for a presentation. All my teammates said, 'Oh, it's like the red-colored button on the top right corner.' For me, there is no top-right corner."* (P9)

The visual modalities used by sighted users and the semantic-based understanding of layouts of BLV users lead to inconsistencies that make receiving and providing instruction challenging. These challenges cascade into layout creation, as explained in Section 4.3.2.

In summary, participants find utility in accessing visual semantic information to navigate, use interfaces, and to collaborate with sighted users. Despite the utility, they do not find value in visual semantics at all times, *e.g.*, when navigating webpages.

4.1.2 Touch vs. Desktop Interface Use. Despite differences in interaction modalities between smartphone and desktop screen readers, visual semantic access was desired across contexts, *e.g.*, participants learned about visual semantics in both smartphones and desktops although to a varying degree. As described in Related Work, smartphone screen readers support richer touch interaction compared to their desktop counterparts. Consequently, smartphones enable direct manipulation of user interface elements, thereby facilitating access to the spatial arrangement of the UI. These differences proved consequential in our data.

Due both to their inherently spatial nature and the direct manipulation aspect of smartphone screen readers, we observed that participants learned about spatiality and visual semantics predominantly on smartphones. However, learning took place on desktops as well, though was valued less. In contrast, desktop screen readers linearize UIs, which can obfuscate visual semantics. Indeed, while four participants wanted better access to visual semantics on desktop UIs, five could identify no benefit to having them. As P3 said, *"[When] I am using a keyboard, I use the find feature quite frequently, and so it doesn't really matter if I know the spatial layout."* (P3). This shows that P3 preferred seeking information without using visual semantics. In contrast, P4 describes the value of visual semantics for navigation when asked if he tried to understand the layout of the website when using a desktop: *"Yes. For the same reason. For navigation purposes. I may not use it in the same manner as when I am touching the screen, but it is important to know where the locations are if I need to get to it. Even navigation is quite different from touching the screen using the navigation keys on the keyboard vs. touching the screen."* (P4). P4 thus describes how he prefers being aware of visual semantics, though they deviate from traditional keyboard-based navigation.

Lack of access to visual semantics may be an important reason why participants feel unmotivated to learn and understand spatial layout information. When asked about the last time they benefited from knowing the spatial layout information of a desktop interface, P1 said, *"I don't know much about spatial layout information. I don't have access about how to know spatial layout information."* (P1).

4.1.3 Strategies to Learn About Visual Semantics.

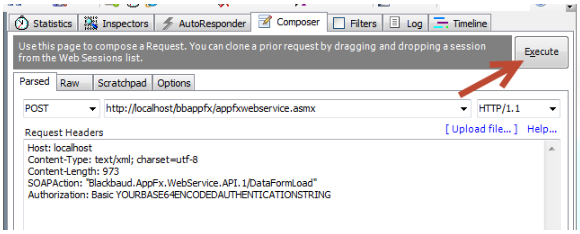
Despite there being no standard method or access technology to teach visual semantics to BLV users, participants reported using several learning strategies, including trial and error, sighted assistance, training seminars, and features in existing screen reading technology, such as exploring by touch or listening to a hierarchical representation of a webpage. On smartphones, participants primarily depended on trial and error, using existing screen reading features such as touch exploration (N=5). On desktops, sighted assistance (N=3) and trial and error (N=3) were mentioned.

This learning about visual semantics on desktops is interesting given that desktop screen readers do not surface spatial arrangement of interfaces or provide visual semantic information. When describing trial-and-error strategies on the smartphone, participants talked about the ability to explore by touch as well as by leveraging common UI design paradigms (*e.g.* tab bars at the bottom of iOS apps). As P3 said, *"When I download a new app, I touch parts of the screen where I expect certain things to be... I will tap on the bottom and see if there are tabs there... Once I start to understand the order of things, maybe I will touch and discover like 'oh that's near the top of the screen.'"* (P3). Participants clearly conveyed the value of these mental models. Interestingly, these function well even though participants were not overly accurate when asked to construct these models, as described in Section 4.2.

Trial and error on the desktop was also possible. For example, P5 describes how ... *"Screen readers themselves have given me [a] qualitative sense of what the layout looks like ... You would also know that this was a heading above that. This was a heading below that."* (P5). Here, P5 interprets the screen reader element order to imply something about a UI's visual semantics (*e.g.*, the relative position of elements).

While participants learned spatial layouts on smartphones using trial and error and sighted assistance, they used only sighted assistance to learn desktop layouts. This contrast in learning strategies is well emphasized in P9's words: *"With touchscreen interfaces, you can actually get a sense of how it is laid out... if you touch, I know what is on the top-right corner, what is on the top-left corner..."* (P9). Instead, participants mentioned using sighted assistance to learn spatial layouts on desktops. For example, P9 said: *"On Outlook [51], I used to have trouble uploading files to email once upon a time, so I had to take sighted help, that's how I know."* (P9). We observed that BLV users gained knowledge of visual semantics of specific interfaces from sighted friends and colleagues when trying to navigate inaccessible layouts. Similarly, P5 said:

> *"I was working with Fiddler [43] today. What I was trying to do was edit the Fiddler script, and it turns out that the way it's laid out is the button, and the button is followed by the text box to the right. It is the other way round... This did happen before that I didn't know the button was there, and I almost thought my focus was not going, and I asked my sighted friend, 'Hey, is there an accessibility bug here?'... You will almost think for a second... You will almost be sure that the screen reader is not reading that text box. So I had her look at it, and I was like, 'what's going on?', and she was like, 'that's how the interface is laid out.'"* (P5)



Use this page to compose a Request. You can clone a prior request by dragging and dropping a session from the Web Sessions list.

http://localhost/bbappfx/appfxwebservice.asmx

```
Host: localhost
Content-Type: text/xml; charset=utf-8
Content-Length: 973
SOAPAction: "Blackbaud.AppFx.WebService.API.1/DataFormLoad"
Authorization: Basic YOURBASE64ENCODEDAUTHENTICATIONSTRING
```

Figure 2: Screenshot of Fiddler, the web debugging software referred to by P5. The arrow shows the execute button for the script text box coming before the text box itself, which made it difficult for P5 to navigate the interface.

We see here that the BLV participant had a theory about the spatial layout of the interface that they revised when something did not work as expected, but they had to ask for sighted confirmation.

Other participants also mentioned receiving sighted help. For example, P1 describes learning Word, *"probably when I was attending some webinars about Word. That webinar was explaining how Word looks like on the screen."* (P1). Similarly, P6 said, *"In training, they would tell me this is here and that is there, but to get comfortable and efficient with it, it was just trial and error. And repetition."* (P6). Interestingly, this quote illustrates that even when sighted assistance was provided, trial and error was still important.

Finally, we discuss other non-dominant strategies participants used to learn visual semantics. P7-LV mentioned learning visual semantics through light perception (current partially functional vision making it possible to distinguish icons based on color for example), P9 mentioned late onset of blindness, and P5 mentioned knowledge of underlying interface code. *"One, I would almost always get the information by looking at the actual code. Or two, [I would] again limit myself to, for example, 'Oh, this heading structure doesn't make sense' or 'Hey, it makes sense to have a horizontal rule here to separate two sections' or you know something else, for example."* (P5). While understanding code requires a specific skillset, our participant's ability to understand visual semantics (like the relative location of elements) from linear representations is intriguing, if unusual.

4.1.4 Challenges When Learning Visual Semantics.

Despite benefiting from visual semantics, participants informed us about two key learning challenges across contexts: missing information (N=5) and inconsistent representations provided by screen readers (N=4). From observations, we learn from four participants that inconsistent information provided by screen readers was a major hurdle for their learning and interacting with spatial layouts. For example, some screen readers differ in whether the arrow keys map onto the direction in which elements are laid out: *"... if you are reading using the PC cursor voice, you are using the down arrow to go down, but if you do it with the JAWS cursor, the direction is going left to right. So the buttons may be in a line across the top but JAWS is reading as though they are down the page vertically."* (P6). This issue was specific to desktop interfaces due to the differences in the design of desktop and mobile screen readers.

In addition to inconsistent information, we observed that missing images and other graphic information by screen readers was a barrier for BLV users to understand visual semantics. Interestingly, P5 was under the impression, based on screen reader feedback, that there was an image on his personal webpage. Though the screen reader mentioned there being a graphic, what was visually displayed was just a placeholder. *"I actually don't know that it's actually a home icon over there and not actually home written in words there. So it's kind of a lot of gap ... "* (P9). Moreover, several participants mentioned missing information as a hurdle to understanding spatial layouts.

Multiple participants reflected on the effort necessary to learn visual semantics and how this affected their smartphone usage, both negatively and positively. For example, P10-LV said: *"When I first start using something, I don't know the placement and stuff, so it takes me a bit to play around [...] Once I am used to it, I am hesitant to try something new."* (P10-LV). Similarly, P2 described the trade-off between effort and payoff relative to interaction efficiency: *"If I'm using something frequently, I'll take the time to learn the spatial layout."* (P2).

To summarize, we observed challenges to learning due to the inconsistent interpretation and incomplete representation of visual elements, graphics and images by screen readers, and the inherent cost (in terms of time and complexity) of learning new interfaces.

4.2 Mental Models of Visual Semantics

To understand BLV users' perception of visual semantics and inform strategies to narrow the gap between non-visual representations and visual semantic understanding, we report findings on perceptions of visual semantics, notions of size and shape, and understanding of overall structure in BLV users. Insights into BLV mental models of visual semantics could better situate future work to support non-visual access to this information.

4.2.1 Perceptions of Visual Semantics. BLV users' perceptions were most acute for smartphone apps, followed by webpages on the smartphone, followed by webpages on the desktop. We now describe these perceptions of size and shape of UI elements, and their overall perceptions of UI layouts.

Participants associated size with functionality and context (app and web). For example, when reconstructing the layout of the Twitter app on smartphone, P4 said, *"These are tweets. Sometimes they can be small, depending on what people put. Sometimes, they put pictures and links."* (P4). While describing the layout of an app (when asked if they feel the need to know how UIs are arranged on screen), P5 said, *"There was this big... you know, like, these rich text editor fields. Usually those editors are like... big, right? I am just basing it on my general knowledge."* (P5).

4.2.2 Size and Shape. Four participants had some estimation of size when prototyping smartphone apps. While exploring the Google Maps app home screen before reconstruction, P3 said, *"There seems to be like a basketball game event suggestion, and that seems to take up like a lot of space."* (P3). She also said, *"I have like a skinny search bar on the top and then like two blobs of Play-Doh."* (P3). While not as precise as P3, P5 was aware of the misrepresentation of the size of elements. He said, *"I am giving a disclaimer right now: The size of elements does not correspond to the size of the buttons."* (P5). Relatedly, two participants had a notion of the size of elements on webpages as they appeared on the smartphone. *"I think this is ordered now... it seems a little wider than the Internet's websites."* (P3). While describing the layout of their reconstruction of the MOD Pizza website as it appears on a smartphone, P3 said, *"and then there is like a bigger thing like 'the great thing about MOD Pizza.' Seems to be bigger than the one above it. And then this heading is like two lines long."* (P3). Lastly, three participants had some notion of the size of elements as they appear on webpages when browsed from a desktop. *"Huge annoying part of the screen. To the right of the left list... I'm putting it two sticky notes wide... probably it's so annoying."* (P3). Similarly, for P8, the size of the canvas dictated where elements went in terms of columns. Participants constructed UIs by column and moved over to the right when they were out of vertical space.

While screen readers do not convey the meaning of shape, it was surprising to hear notions of shape described by BLV users. Interestingly, however, our findings do not inform comparison or contrast of these perceptions across device or contexts, since the shapes that participants created were not very distinct from each other across device platforms or usage scenarios. Participants used representative shapes instead: links were straight lines, buttons were blobs or circular shaped Play-Doh. While lo-fi prototyping the Facebook iOS app, P2 represented buttons for the camera and messenger icons as round circles. Some participants however, were very creative in physical representations of UI elements. For example, when describing her reconstructed model of a desktop website, P9 used a large rectangular piece of Play-Doh to denote a search results section and used a pen to make a horizontal indentation to denote a heading in the search results.

We find BLV technology users have notions of different kinds of UIs (app and web) to varying degrees. They had notions of general layouts (N=2), layouts of webpages on desktops (N=6), and layouts of websites they own (N=1). Participants developed these layouts primarily based on screen reader representations (N=2). Six participants primarily perceived layouts of websites on the desktop as being vertical. Evident from P8's experience, *"Old habits, really more than anything. No reason why I couldn't when I learn; I've been using a computer since I was 11. So everything as far as [the] Internet was vertical, line-by-line."* (P8). Similarly, we observed P5's lo-fi prototype of a website as it appears on the desktop to predominantly have everything to the left, with the right side of the canvas remaining empty. Interestingly, participants' lo-fi prototypes of phone apps were more spread across the canvas as opposed to those of websites on desktops.

4.2.3 Understanding of Structure/Layout. Looking at a higher level than shape, P5 and P8 had some general awareness of news website structure. *"You know, for a news article. You know well there is probably going to be ads above and below the actual news content."* (P5). This was also evident in P1's lo-fi prototype of a webpage as it appears on the desktop. She was confident to construct the layout of a generic news website as opposed to the layout of a particular one.

Lastly, participants developed their mental models of layouts through screen reader feedback. As P6 said, *"Just the concept that,* *to me, spatial layout is based on the screen reader interface. It's not based on what it looks like. And that happens across accessibility formats. So my perception of how a page looks using a braille display is different than using JAWS. Because the braille display can go only 40 cells at a time. So the concept of spatial is very different."* (P6). To further understand this trend, we compare accuracy of participants' prototypes to whether or not they referred to the original interface.

Six participants referred to the original app interface and four did not. Of those that referred (*N*=6), three participants re-created fairly accurate prototypes. Two participants produced moderately accurate layouts, while one produced a prototype with low accuracy. Of those that chose not to refer to the interface (*N*=4), one produced a fairly accurate layout, one moderate and two low.

For websites on the smartphone, five participants reconstructed the layouts on referring, three chose not to refer and two did not prototype. Note that all the fairly accurate layouts were produced by participants who referred to the original web interface. Accuracy was overall moderate to low (*N*=8). For participants who chose not to refer, one participant produced a moderately accurate layout while two produced low accuracy prototypes.

For desktop web, six participants chose not to refer to the original website interface while prototyping. Of these six, five participants chose not to refer, and reconstructed prototypes with low accuracy. Of those that referred (*N*=3), one constructed a fairly accurate prototype, one constructed moderately accurate and one participant's prototype was less accurate. Surprisingly for personal websites, accuracy was moderate (*N*=4) both for participants who referred (*N*=2), and did not (*N*=2). One participant who did not refer produced a prototype of low accuracy.

Our findings reveal how participants represented UIs, what they struggled with, and their preferences while prototyping like reference points, strategies to describe layouts, and navigation order while constructing. Existing lo-fi prototyping techniques and software are not accessible to BLV individuals. However, to illustrate BLV conceptualization of visual semantics, we next present insights from the low-fi prototyping task.

4.3 Prototyping and Creation of Visual Semantics

To better understand BLV users' conception of visual semantics, we asked participants to prototype familiar UIs with lo-fi materials. We analyzed and report on our observations, their material designs, and their "think aloud" comments.

4.3.1 Prototyping Preferences and Observations. Interestingly, participant preference did not vary with device (smartphone vs. desktop) and usage scenario (app vs. web) while lo-fi prototyping.

Ten participants used reference points for smartphone app reconstructions. P4, for example, started construction on the bottom-left corner of the canvas, and P5 started on the top-left. Similarly, P9 started on the top and bottom, thereafter proceeding with placing straight lines of Wikki Stix from bottom to top. We observe that P3's notion of a reference point included static elements of the phone's UI. When asked to prototype the layout of a smartphone app, she clarified by asking, *"No status bar, right? Like the time and stuff."* (P3). Relatedly, we could observe reference points for five participants as they constructed layouts of websites they owned and built. P3, for example, started by constructing the browser's address bar. Similarly, P5 started from the top. P9, in addition to starting from the top, divided the canvas into sections. We assume that this division was due to her prior experience as a web developer.

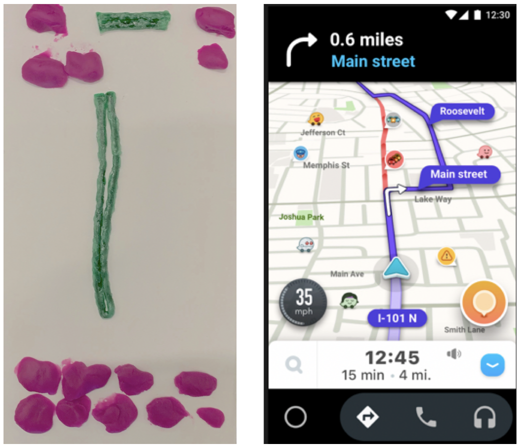


Figure 3: P6 smartphone app reconstruction (left) of the Waze navigation app (right).

While ten participants constructed lo-fi prototypes, we perceived four participants to be generally interested in these prototyping tasks. All participants prototyped the layout of the smartphone app. When asked to construct the layout of a phone app, P9 was very enthusiastic and accepted the task like a challenge. She said, *"I think I can do the phone app. It has more stuff. Feels like I've become a kid again. It's been a while since I played with all of these. Especially I used to play the fun school game."* (P9). Two participants expressed disinterest but continued to prototype the layout of a smartphone website, and two chose not to prototype. One participant chose to construct the layout of a generic webpage. For desktop web, one participant did not prototype, and one participant chose to prototype the layout of a generic news website, mentioning that she was not familiar with the layouts of any specific website that she browsed on the desktop. *"I don't have in my mind the layout of a particular website. I am trying to think [...] also the websites [...] they look different from one another to me."* (P1). All five participants who had a personal website prototyped its layout. Note that not all participants expressed a preference or sentiment about our prototyping task.

4.3.2 Strategies to Describe and Construct Layouts. We summarize different strategies participants used to describe and construct interfaces to inform the readers of BLV notions of verbalizing visual semantics. Description strategies included the use of relative positions with some elements as anchor points (*e.g.*, "the button is below the title header"), use of corner elements, use of element counts and absolute positions (*e.g.*, "the button is at x,y location"), with relative positions being predominant (three for smartphone apps and three for personal webpages). We also present other strategies, including assumed location and absolute positions, and discuss participant strategies for constructing interfaces, *e.g.*, navigation order.

Three participants used relative positions to describe UIs and lo-fi prototypes. *"In my view, the settings app is having one kind of notification area on the top, which kind of shows me tips if my airplane mode is on; below that, I have a list of things network and Internet, display and sounds; I only go to network and Internet all the time."* (P9). Similarly, three participants used relative positions to describe layouts of web pages they owned or built. *"The way I think about this is on the top is my email, to the right is my phone number. Below that [is] essentially a 2 by 2 kind of thing, which has my social profile."* (P5).

Four participants used edge elements to describe layouts. *"Well, at the top of the screen is the buttons that tell you like the directions. There's also… when you first open the app, first there's search and then voice; here's a microphone. In the middle, there's the map and then, at the bottom, there's the buttons for menu and sound and speed and stuff like that."* (P6). Similarly, three participants used corner elements to describe smartphone layouts. *"And in the top right is the add contact button."* (P6). Interestingly, one participant, P9, used absolute position to describe the layout of the contacts app. *"On top is quick contact for blah blah blah. The one over here on the right hand side is the call button, say if I want to call them again."* (P9).

We next present findings on navigation order strategies that participants used to: prototype layouts, refer to the original app for construction, and to describe the constructed layouts. In the context of screen reader access, navigation order is determined by the UI builders. However, it was interesting to see participants using a variety of strategies, like going left-to-right, top-to-bottom, and bottom-to-top when constructing, referring to the original interface, and describing the reconstructions.

4.3.3 Navigation Order. We observe that participants preferred going left-to-right, top-to-bottom on the canvas when they were prototyping layouts (six for smartphone app, seven for websites on smartphones and desktops, and four for personal websites). The second most common observation was that participants placed UI elements on the canvas in a top-to-bottom fashion (two for smartphone app, one for smartphone web, two for desktop web, and one for personal website). We noticed that the constructions where participants went in a top-to-bottom fashion had less detail, likely resulting in their not going left-to-right. Lastly, it was interesting to see two participants go in no specific order when prototyping a smartphone app.

While describing their interfaces, participants went predominantly from left-to-right, top-to-bottom across contexts (eight for apps, seven for websites on smartphones and desktops, and three for personal websites). The second most commonly observed strategy was to only go top-to-bottom (two for apps, one for smartphone websites, two for desktops and one for personal websites). Interestingly, we observe that one of the participants lacked sufficient detail in their reconstruction to go left-to-right, top-to-bottom.

Four participants followed a left-to-right, top-to-bottom order when referring to the smartphone app during reconstruction. Three participants did not refer to the app, and we were unable to discern the order for one participant. For websites on the smartphone, four participants followed a left-to-right, top-to-bottom order, and two participants did not refer to the original webpage. For desktop websites, while three participants used screen reader ordering, it was interesting to see five participants not referring to the original layout. We observe similar behavior with respect to personal websites (two used screen readers, three did not refer to the original). Though an accuracy comparison between the reconstruction and actual interface would help clarify this, we do not report accuracy related findings; limited non-visual access to visual attributes of interfaces, the disconnect between visual ordering of elements and semantic structure provided to screen readers, and the lack of access to visual information may make such a comparison unfair.

4.3.4 Building Web Pages. Five of our ten participants had webpages that they built or maintained. These pages were also meant for visual consumption. We explore BLV developer strategies for making visual design decisions and building visual layouts.

Participants used WordPress and WordPress templates (three participants), sighted help (N=3), and HTML (N=3) as strategies to develop or maintain their webpages. WordPress was the back-end used by three of the five participants who had a webpage. Three participants perceived it as a tool, and two participants commented on the value of WordPress templates. When asked if they built their own website, P9 said, *"Yes and no. Like not kind of like I coded it. I just plugged in kind of a lot of WordPress plugins."* (P9). To the same question, P3 said, *"I used a template. I don't know what it is called […] I write the content. The layout is generated by the template."* (P3). Three other participants used their knowledge of HTML markup to develop parts of their personal webpages or blogs. *"I didn't use a platform as such, like WordPress or whatever. It's just plain HTML and CSS."* (P5). Lastly, three participants relied heavily on sighted assistance to develop their website. *"I literally copied someone's code… with their permission obviously and changed things."* (P5).

Surprisingly, based on our observations and interpretations, no participant independently selected the template for their webpage, and no participant independently pushed major updates to their webpage–major areas of contribution for future design tools. Participants used templates and settings that are known to be accessible (N=2) and assistance from friends and family members to make visual design decisions on their webpages. In P3's words:

> *"I used a template. I don't know what it is called. When I got a website, I had a friend set it up for me. He chose a template that is supposed to be more accessible. I don't really know what he did. He turned off the WYSIWYG editor…He made a couple of tweaks to make that easier for me. I hate that I don't know what he did."* (P3)

P3's words reveal the sense of dependency associated with making a visual layout choice, and the lack of autonomy and knowledge that result from this dependency.

From the observations, we see that BLV individuals find utility in visual semantics across context. While existing access technology provides this information, there are significant limitations that prevent BLV users from making the best use of this information. Furthermore, insufficient access to visual semantics increases the difficulty of non-visually building visually appealing layouts.

> *"I am really really bad at this is what I understood. […] Its also very interesting to understand and appreciate the more commonly used websites and know about them so that you can better design your website […] All that*

> *I really care about are things that visually might look awkward to people but how they exactly look, I don't care that much. If I think about it now, maybe its not the right thing. Maybe you want to understand how the layouts actually look and make some contribution there."* (P5)

Here, we find an introspective realization of importance of visual semantics for BLV users, further strengthening the need for non-visual visual semantic access. P5 did not care much about the appearance of user interfaces beyond they appearing awkward, but after answering our questions and prototyping four UIs, he reflects on the need to have access to this information and expresses desire to pay attention to these.

5 DISCUSSION

Our three-part study illustrates the potential of visual semantic information to enable more efficient use of and collaborative interactions with computers and smartphones for blind and low-vision users. We find that BLV users see value in access to spatial layouts in both desktops and smartphones, although this information is valued slightly less in the context of desktops. Furthermore, participants reported learning spatial layouts using a variety of strategies, including screen reader functionality and sighted assistance. Finally, users faced access gaps that limited their participation in prototyping and building visual layouts. While future work should explore the positive and negative impact of access to visual semantics on start-up costs associated with learning new, frequently used interfaces, this study demonstrates that there is both interest and value in exposing more of this information.

Below, we reflect on our findings and implications for design. Based on our findings, we suggest: (1) research into how to better describe semantics like the use of relative positioning when describing interface elements and integration of these descriptions into screen readers, and (2) support for BLV visual semantic prototyping. We discuss multi-modal access, techniques to verbalize positioning information, and adaptive methods to shorten descriptions. We end with the importance of visual semantics in prototyping GUIs, our study limitations, and ethical considerations.

5.1 Provide Multi-Modal Access to Visual Semantics

We find that knowledge of visual semantics helps BLV users with increased software usability, supports UI navigation, and helps BLV users collaborate with and/or guide other users who may be sighted or blind. At the same time, participants do not need access to this information at all times. New interaction designs aimed at conveying visual semantic information should consider *when* and *how* this information should be delivered. Participants valued visual semantic information to interact with interfaces and collaborate with sighted individuals, and learning took place predominantly on smartphones and to some degree on desktops. Direct access to UI elements on touchscreen smartphones, and the resulting familiarity with the spatial arrangement of interfaces over time, may have partly contributed to this predominance on smartphones. *How can technologies provide richer, meaningful access to visual semantics so BLV individuals can benefit across contexts, including the desktop?*

The full potential of visual semantic access of user interfaces could be realized if participants had access to visual semantics during regular use (*e.g.* using a screen reader), and when using devices prevalent in everyday computing (*e.g.* touchscreen devices). Given the understanding of visual semantics demonstrated in our study with respect to phones, it seems clear that this is possible without requiring costly special hardware or expertise to generate physical representations of visual semantics. This recommendation is in line with prior work exploring socio-technical considerations for accessible visualization design [48] and the cost associated with tactile graphics outlined in [31].

Given the ubiquity of smartphones relative to specialized hardware like embossers, we recommend emulating the sorts of exploratory features supported on phones for other devices. Smartphone or tablet touchscreens and/or laptop touchpads could be used as assistive input devices for desktops, with similar direct manipulation and gesture techniques for non-sighted exploration. We further recommend voice-based access to visual semantics. In a comparison of embossed visualizations to multi-modal touchscreen representations, Gorlewicz *et al.* [31] found that participant performance was equivalent, which suggests that tactile graphics, while promising, are not essential, at least in the domain of information visualization. Future work should verify whether the same is true for visual semantics of interfaces, and investigate translation of the guidelines developed by Gorlewicz *et al.* [31] to visual semantics.

5.2 Positioning Information and Verbal Descriptions

While existing screen readers present visual information by giving audio feedback on the mouse's location, conveying indentation and text formatting, and using three-dimensional sound to indicate the UI element location that is being activated [1, 54], this approach does not expose the full range of visual semantics *e.g.* size and shape of UI elements.

Our findings demonstrate that BLV users used a variety of strategies, *e.g.* sighted assistance and trial and error to familiarize themselves with visual semantics of interfaces. These findings show us that BLV users attempted to understand a wide range of visual semantic information whether or not specific features in screen readers to surface this information. These observations lead us to ask:

How can screen readers further help BLV users learn about the interfaces they use? Given that BLV users do not require visual semantic information at all times, we recommend that screen readers could provide a visual verbosity setting, where users can access visual information about interface elements, *i.e.*, an interface design mode similar to punctuation verbosity. Similarly, screen readers could provide visual descriptions of entire interfaces. For example, existing screen reader functionality to read screen content could be augmented to support descriptions of visual semantics. Our findings show BLV preference for description order of these semantics to be: left-to-right, top-to-bottom.

However, given that this preference may not be uniform (our participants showed variation in the element order when describing visual interfaces), we recommend that these orderings be customizable.

Our participants used a variety of vocabulary to describe interfaces *e.g.* indicating through relative positioning (above and below). Likewise, we recommend that computer generated descriptions use relative positioning, *e.g.*, next to, below and above, with one or multiple elements as a reference. Descriptions should also help users to collaborate with sighted and blind individuals.

5.3 Adaptive Visual Semantic Presentation

Given that describing visual semantics can be a lengthy process, adaptive automated description generation could make this more efficient for users. Our participants expressed familiarity with edge and corner elements of interfaces as well as familiarity with certain layouts, like news websites and frequently used smartphone apps. This suggests an opportunity to use data from UI repositories such as *Rico* [25] and *Webzeitgeist* [41] to compare a new interface to those known to be used frequently by the same user. A layout description could then be generated that compares the current UI to the familiar interface and uses AI methods to generate richer semantic descriptions.

5.4 Increasing Prototyping Ability

Though our primary focus was to understand how BLV users perceived visual semantics and the impact of this visual semantic understanding among BLV users on accessibility, it is worth noting that these visual semantics become important in GUI prototyping. Our observations of lo-fi prototyping during the study provide helpful preliminary insights that could better support nonvisual prototyping of user interfaces. Prototyping a visual interface for a BLV user should be as simple as drawing a design on a piece of paper (lo-fi sketching) or dragging, dropping and drawing in a What You See Is What You Get (WYSIWYG) interface (mid-fi or high-fi prototyping) [20], both methods are available to sighted users. For BLV individuals, however, the lack of access to prototyping methods and software makes it extremely hard to prototype, with the only way often being to write code.

In our study, we adapted lo-fi prototyping for accessibility by using tactile material (Wikki Stix and Play-Doh). This study design decision was made to offer maximum flexibility for participants to tactually construct their prototypes. Future work should study effectiveness of different tactual modalities for BLV technology users to provide data to experiments aimed at examining visual understanding of user interfaces, and eventually prototype these interfaces in the real world. Our findings inform us that the shapes of UI elements that our participants created were not very distinct from each other, and that participants used representative shapes *e.g.* straight lines for links. While it was interesting to observe some understanding related to shape and size of elements, and correlations of different elements to representative shapes, these findings also surface the limited understanding of shape and size of UI elements among BLV users. Relatedly, with limited perceptions of shape, size and visual appeal, future work should explore methods to help BLV individuals meaningfully prototype. To increase BLV prototyping capabilities, future tools should estimate visual information and appeal of prototypes, provide modalities to spatially organize visual interface elements on a canvas, and analyze conformance of designs to guidelines. These capabilities could make UI design more accessible for BLV designers, allowing them to make informed visual design choices and create visually appealing prototypes.

All of these are based on technical advances that have been shown to be feasible in other domains. For example, Microsoft Office PowerPoint Designer [23] suggests visual designs. Data from [41] could be used to build recommenders for accessible design recommendations for a very wide variety of web and mobile interfaces. Similarly, declarative UI languages like SwiftUI [6] that encode visual semantics nonvisually, could enable BLV creators to declaratively create visual layout prototypes. Alternatively, new interaction techniques similar to those explored in [62] should be developed to support nonvisual spatial placement of visual interface elements. Finally, prototyping tools should inform BLV designers if their changes to prototypes adhere to visual design guidelines. Program verification approaches such as work to formally verify webpage accessibility [56] could be extended to verify visual design guidelines.

5.5 Limitations

Our ten-person, exploratory qualitative study to examine how BLV users understand visual semantics had samples drawn from a single geographic area in the US, and had five users who previously built visual interfaces such as web pages, making our sample skewed toward expertise. This bias is appropriate to determining *what* information to convey, a primary goal of this work, but may not fully explain novice perspectives. Our findings demonstrate that even with technical ability, BLV participants did not find visual interfaces fully understandable. Less proficient BLV users may have even larger hurdles to overcome. Future work should examine visual semantic understanding across diverse groups of users.

5.6 Deviating From Normative Notions of Visual Semantic Access

While it may seem like the goals of this work are to norm visual semantic access, our objectives are contrary. We are not proposing that BLV users *should* have the same experiences with visual semantics as sighted users, but rather our work questions ableist assumption that BLV users can not or do not want to understand visual semantics. At the same time, designers should have a deeper discussion about placing responsibility on blind people to understand visual semantics to have similar experiences as sighted people.

Our work also deviates from a problematic corollary assumption, that BLV users cannot design visually pleasing interfaces that sighted users could use. Our findings show promise for a larger conversation about BLV users as capable and interested visual designers. In an inherently normative world where people form impressions in a split second [45], an understanding of visual semantics is critical for BLV users who wish to present themselves online. Our work provides a start at understanding this important problem.

5.7 Ethical Considerations

It is important to consider the disability context when designing methods to introduce accessible visual semantics, particularly about the expectations relating to knowledge that BLV users may have about visual aesthetics. While this could open up new possibilities, it also has the potential to put undue burden on BLV users to learn new tools, to produce good quality output in a modality that they do not have complete access to, and open them up to additional, potentially unfair criticisms. We could observe some of these tensions while running our user studies: though not quantifiable, a few participants were not very comfortable with the prototypes they produced, and were highly critical of their skill in prototyping. While we reiterated that we *are not* testing their skills or ability to do a certain activity, the effect this experience may have had on their confidence is unknown. Likewise, if a BLV user creates an accessible UI tool that provides feedback on what is appealing, and if the output is not appealing to an end user or a person evaluating the interface generated by this tool, is it fair to criticize the BLV creator? Contrastingly, given that the BLV creator is responsible for decisions related to the output, *what could constructive criticism look like?* It is important to be explicit about the limitations of future tools that provide nonvisual access to visual semantics.

6 CONCLUSION

In this work, we performed a three-part study to identify how BLV users understand and use visual semantics across device platforms (desktop and smartphone) and common usage tasks (apps and websites). We find that these users see value in access to visual semantics in both desktops and smartphones, although this information is valued slightly less in the context of desktops. Furthermore, participants reported learning visual semantics using a variety of strategies, including screen reader functionality and sighted assistance. We closed by discussing ideas for implementing visual semantic access in non-visual interfaces.

With increasing prevalence of visually dominant modalities, and given that visual representations of interfaces are of value and utility for non-visual access, it is critical to rethink screen reader design and explore deviations from prevalent linear representations without breaking existing functionality or introducing a steep learning curve.

ACKNOWLEDGMENTS

We thank our BLV participants for their expertise to help us examine visual semantic understanding. We thank Google, UW CREATE and the National Science Foundation (NSF 1836813) for supporting this work. We also are grateful for valuable input from Dhruv Jain and Kelly Mack that helped us execute this work.

REFERENCES

[1] NV Access. 2019. NVDA 2019.3.1. https://www.nvaccess.org/files/nvda/documentation/userGuide.html

[2] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: A navigational cognitive assistant for the blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 90–99. https://doi.org/10.1145/2935334.2935361

[3] Khaled Albusays, Stephanie Ludi, and Matt Huenerfauth. 2017. Interviews and Observation of Blind Software Developers at Work to Understand Code Navigation Challenges. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) *(ASSETS '17)*. Association for Computing Machinery, New York, NY, USA, 91–100. https://doi.org/10.1145/3132525.3132550

[4] Apple. 2019. iPhone User Guide. https://support.apple.com/guide/iphone/welcome/ios

[5] Apple. 2019. macOS User Guide. https://support.apple.com/guide/mac-help/welcome/mac

[6] Apple. 2019. SwiftUI. (2019). https://developer.apple.com/xcode/swiftui/

[7] Apple, INC. 2019. *VoiceOver*. Apple, Cupertino, CA. https://www.apple.com/accessibility/mac/vision/

[8] Mark S. Baldwin, Gillian R. Hayes, Oliver L. Haimson, Jennifer Mankoff, and Scott E. Hudson. 2017. The Tangible Desktop: A Multimodal Approach to Nonvisual Computing. *ACM Trans. Access. Comput.* 10, 3, Article 9 (Aug. 2017), 28 pages. https://doi.org/10.1145/3075222

[9] Nikola Banovic, Rachel L. Franz, Khai N. Truong, Jennifer Mankoff, and Anind K. Dey. 2013. Uncovering Information Needs for Independent Spatial Learning for Users Who Are Visually Impaired. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. Association for Computing Machinery, New York, NY, USA, Article 24, 8 pages. https://doi.org/10.1145/2513383.2513445

[10] Cynthia L. Bennett. 2018. A Toolkit for Facilitating Accessible Design with Blind People. *SIGACCESS Access. Comput.* 120 (Jan. 2018), 16–19. https://doi.org/10.1145/3178412.3178415

[11] Cynthia L. Bennett, Kristen Shinohara, Brianna Blaser, Andrew Davidson, and Kat M. Steele. 2016. Using a Design Workshop To Explore Accessible Ideation. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) *(ASSETS '16)*. Association for Computing Machinery, New York, NY, USA, 303–304. https://doi.org/10.1145/2982142.2982209

[12] Jeffrey P. Bigham, Anna C. Cavender, Jeremy T. Brudvik, Jacob O. Wobbrock, and Richard E. Ladner. 2007. WebinSitu: A Comparative Analysis of Blind and Sighted Browsing Behavior. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility* (Tempe, Arizona, USA) *(Assets '07)*. Association for Computing Machinery, New York, NY, USA, 51–58. https://doi.org/10.1145/1296843.1296854

[13] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and Tom Yeh. 2010. VizWiz: Nearly Real-time Answers to Visual Questions. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. ACM, New York, NY, USA, 333–342. https://doi.org/10.1145/1866029.1866080

[14] Jeffrey P. Bigham, Tessa Lau, and Jeffrey Nichols. 2009. Trailblazer: Enabling Blind Users to Blaze Trails through the Web. In *Proceedings of the 14th International Conference on Intelligent User Interfaces* (Sanibel Island, Florida, USA) *(IUI '09)*. Association for Computing Machinery, New York, NY, USA, 177–186. https://doi.org/10.1145/1502650.1502677

[15] BlindSquare. 2020. BlindSquare: Pioneering accessible navigation - indoors and outdoors. https://www.blindsquare.com/

[16] Andy Borka. 2019. Developer Toolkit. https://addons.nvda-project.org/addons/developerToolkit.en.html

[17] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[18] Robin N. Brewer. 2018. Facilitating Discussion and Shared Meaning: Rethinking Co-Design Sessions with People with Vision Impairments. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare* (New York, NY, USA) *(PervasiveHealth '18)*. Association for Computing Machinery, New York, NY, USA, 258–262. https://doi.org/10.1145/3240925.3240981

[19] Craig Brown and Amy Hurst. 2012. VizTouch: Automatically Generated Tactile Visualizations of Coordinate Spaces. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* (Kingston, Ontario, Canada) *(TEI '12)*. Association for Computing Machinery, New York, NY, USA, 131–138. https://doi.org/10.1145/2148131.2148160

[20] Bill Buxton. 2010. *Sketching user experiences: getting the design right and the right design.* Morgan kaufmann.

[21] Zaira Cattaneo, Tomaso Vecchi, Cesare Cornoldi, Irene Mammarella, Daniela Bonino, Emiliano Ricciardi, and Pietro Pietrini. 2008. Imagery and spatial processes in blindness and visual impairment. *Neuroscience & Biobehavioral Reviews* 32, 8 (2008), 1346–1360.

[22] Aira Tech Corp. 2020. Aira. https://aira.io/

[23] Microsoft Corporation. 2020. *Create professional slide layouts with PowerPoint Designer - Office Support.* https://support.microsoft.com/en-us/office/create-professional-slide-layouts-with-powerpoint-designer-53c77d7b-dc40-45c2-b684-81415eac0617

[24] Microsoft Corporation. 2021. Seeing AI. https://www.microsoft.com/en-us/ai/seeing-ai

[25] Biplab Deka, Zifeng Huang, Chad Franzen, Joshua Hibschman, Daniel Afergan, Yang Li, Jeffrey Nichols, and Ranjitha Kumar. 2017. Rico: A Mobile App Dataset for Building Data-Driven Design Applications. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 845–854. https://doi.org/10.1145/3126594.3126651

[26] LightHouse for the Blind and Visually Impaired. 2020. TMAP: Tactile Maps Automated Production. https://lighthouse-sf.org/tmap/

[27] Stéphanie Giraud, Anke M. Brock, Marc J.-M. Macé, and Christophe Jouffrais. 2017. Map Learning with a 3D Printed Interactive Small-Scale Model: Improvement of Space and Text Memorization in Visually Impaired Students. *Frontiers in Psychology* 8 (2017), 930. https://doi.org/10.3389/fpsyg.2017.00930

[28] Google. 2019. Android Accessibility Help. https://support.google.com/accessibility/android/answer/9031274?hl=en

[29] Google. 2019. Android Help. https://support.google.com/android/?hl=en#topic=7313011

[30] INC Google. 2019. *TalkBack*. Google, Mountainview, Ca. https://support.google.com/accessibility/android/answer/6283677?hl=en

[31] Jenna L. Gorlewicz, Jennifer L. Tennison, P. Merlin Uesbeck, Margaret E. Richard, Hari P. Palani, Andreas Stefik, Derrick W. Smith, and Nicholas A. Giudice. 2020. Design Guidelines and Recommendations for Multimodal, Touchscreen-Based Graphics. *ACM Trans. Access. Comput.* 13, 3, Article 10 (Aug. 2020), 30 pages. https://doi.org/10.1145/3403933

[32] Darren Guinness, Annika Muehlbradt, Daniel Szafir, and Shaun K. Kane. 2019. RoboGraphics: Using Mobile Robots to Create Dynamic Tactile Graphics. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 673–675. https://doi.org/10.1145/3308561.3354597

[33] Waqar Haider and Yeliz Yesilada. 2020. Tables on the Web Accessible? Unfortunately Not!. In *Proceedings of the 17th International Web for All Conference* (Taipei, Taiwan) *(W4A '20)*. Association for Computing Machinery, New York, NY, USA, Article 7, 5 pages. https://doi.org/10.1145/3371300.3383349

[34] Web Accessibility initiative. 1999. Understanding Success Criterion 1.3.1: Info and Relationships. url. Retrieved Jan 16, 2020 from https://www.w3.org/WAI/WCAG21/Understanding/info-and-relationships.html.

[35] Web Accessibility initiative. 2019. Tables Consepts. Retrieved January 16, 2020 from https://www.w3.org/WAI/tutorials/tables/.

[36] Shaun K. Kane, Jeffrey P. Bigham, and Jacob O. Wobbrock. 2008. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-touch Interaction Techniques. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility* (Halifax, Nova Scotia, Canada) *(ASSETS '08)*. ACM, New York, NY, USA, 73–80. https://doi.org/10.1145/1414471.1414487

[37] Shaun K. Kane, Meredith Ringel Morris, Annuska Z. Perkins, Daniel Wigdor, Richard E. Ladner, and Jacob O. Wobbrock. 2011. Access Overlays: Improving Non-visual Access to Large Touch Screens for Blind Users. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. ACM, New York, NY, USA, 273–282. https://doi.org/10.1145/2047196.2047232

[38] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 413–422. https://doi.org/10.1145/1978942.1979001

[39] Rushil Khurana, Duncan McIsaac, Elliot Lockerman, and Jennifer Mankoff. 2018. Nonvisual Interaction Techniques at the Keyboard Surface. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 11, 12 pages. https://doi.org/10.1145/3173574.3173585

[40] Ravi Kuber, Wai Yu, and Graham McAllister. 2007. A Non-visual Approach to Improving Collaboration Between Blind and Sighted Internet Users. In *Universal Access in Human-Computer Interaction. Applications and Services*, Constantine Stephanidis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 913–922.

[41] Ranjitha Kumar, Arvind Satyanarayan, Cesar Torres, Maxine Lim, Salman Ahmad, Scott R. Klemmer, and Jerry O. Talton. 2013. Webzeitgeist: Design Mining the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3083–3092. https://doi.org/10.1145/2470654.2466420

[42] Richard E. Ladner, Melody Y. Ivory, Rajesh Rao, Sheryl Burgstahler, Dan Comden, Sangyun Hahn, Matthew Renzelmann, Satria Krisnandi, Mahalakshmi Ramasamy, Beverly Slabosky, Andrew Martin, Amelia Lacenski, Stuart Olsen, and Dmitri Groce. 2005. Automating Tactile Graphics Translation. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, MD, USA) *(Assets '05)*. Association for Computing Machinery, New York, NY, USA, 150–157. https://doi.org/10.1145/1090785.1090814

[43] Eric Lawrence. 2020. Fiddler-Free Web Debugging Proxy. https://www.telerik.com/fiddler

[44] Jingyi Li, Son Kim, Joshua A. Miele, Maneesh Agrawala, and Sean Follmer. 2019. Editing spatial layouts through tactile templates for people with visual impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 206, 11 pages. https://doi.org/10.1145/3290605.3300436

[45] Gitte Lindgaard, Gary Fernandes, Cathy Dudek, and J. Brown. 2006. Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology* 25, 2 (2006), 115–126. https://doi.org/10.1080/01449290500330448 arXiv:https://doi.org/10.1080/01449290500330448

[46] Jack M Loomis, Roberta L Klatzky, Reginald G Golledge, Joseph G Cicinelli, James W Pellegrino, and Phyllis A Fry. 1993. Nonvisual navigation by blind and sighted: assessment of path integration ability. *Journal of Experimental Psychology: General* 122, 1 (1993), 73.

[47] Christina Low, Emma McCamey, Cole Gleason, Patrick Carrington, Jeffrey P. Bigham, and Amy Pavel. 2019. Twitter A11y: A Browser Extension to Describe Images. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 551–553. https://doi.org/10.1145/3308561.3354629

[48] A. Lundgard, C. Lee, and A. Satyanarayan. 2019. Sociotechnical Considerations for Accessible Visualization Design. In *2019 IEEE Visualization Conference (VIS)*. 16–20. https://doi.org/10.1109/VISUAL.2019.8933762

[49] Microsoft. 2019. *Windows 10 Manual User Guide 2019*. https://windows10-guide.com/

[50] Microsoft. 2020. Microsoft Soundscape: a map delivered in 3D sound. https://www.microsoft.com/en-us/research/product/soundscape/

[51] Microsoft. 2020. Office 365. https://www.microsoft.com/en-us/microsoft-365

[52] Meredith Ringel Morris, Jazette Johnson, Cynthia L. Bennett, and Edward Cutrell. 2018. Rich Representations of Visual Content for Screen Reader Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 59, 11 pages. https://doi.org/10.1145/3173574.3173633

[53] NV Access. 2019. *Nonvisual Desktop Access*. NV Access. https://www.nvaccess.org/

[54] Musharraf Omer. 2019. *Audio Themes*. https://addons.nvda-project.org/addons/AudioThemes.uk.html

[55] Steve Oney, Alan Lundgard, Rebecca Krosnick, Michael Nebeling, and Walter S. Lasecki. 2018. Arboretum and Arbility: Improving Web Accessibility Through a Shared Browsing Architecture. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 937–949. https://doi.org/10.1145/3242587.3242649

[56] Pavel Panchekha, Adam T. Geller, Michael D. Ernst, Zachary Tatlock, and Shoaib Kamil. 2018. Verifying That Web Pages Have Accessible Layout. *SIGPLAN Not.* 53, 4 (June 2018), 1–14. https://doi.org/10.1145/3296979.3192407

[57] Sujeath Pareddy, Anhong Guo, and Jeffrey P. Bigham. 2019. X-Ray: Screenshot Accessibility via Embedded Metadata. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 389–395. https://doi.org/10.1145/3308561.3353808

[58] Venkatesh Potluri, Tad Grindeland, Jon E. Froehlich, and Jennifer Mankoff. 2019. AI-Assisted UI Design for Blind and Low-Vision Creators.

[59] Venkatesh Potluri, Liang He, Christine Chen, Jon E. Froehlich, and Jennifer Mankoff. 2019. A Multi-Modal Approach for Blind and Visually Impaired Developers to Edit Webpage Designs. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 612–614. https://doi.org/10.1145/3308561.3354626

[60] Hrishikesh V. Rao and Sile O'Modhrain. 2020. 2Across: A Comparison of Audio-Tactile and Screen-Reader Based Representations of a Crossword Puzzle. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376207

[61] Eva-Lotta Sallnäs, Kajsa Bjerstedt-Blom, Fredrik Winberg, and Kerstin Severinson Eklundh. 2006. Navigation and Control in Haptic Applications Shared by Blind and Sighted Users. In *Haptic and Audio Interaction Design*, David McGookin and Stephen Brewster (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 68–80.

[62] Ashrith Shetty, Ebrima Jarjue, and Huaishu Peng. 2020. Tangible Web Layout Design for Blind and Visually Impaired People: An Initial Investigation. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 37–39. https://doi.org/10.1145/3379350.3416178

[63] Lei Shi, Yuhang Zhao, Elizabeth Kupferstein, and Shiri Azenkot. 2019. A Demonstration of Molder: An Accessible Design Tool for Tactile Maps. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 664–666. https://doi.org/10.1145/3308561.3354594

[64] Alexa F. Siu, Son Kim, Joshua A. Miele, and Sean Follmer. 2019. ShapeCAD: An Accessible 3D Modelling Workflow for the Blind and Visually-Impaired Via 2.5D Shape Displays. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 342–354. https://doi.org/10.1145/3308561.3353782

[65] Abigale Stangl, Jeeeun Kim, and Tom Yeh. 2014. 3D Printed Tactile Picture Books for Children with Visual Impairments: A Design Probe. In *Proceedings of the 2014 Conference on Interaction Design and Children* (Aarhus, Denmark) *(IDC '14)*. Association for Computing Machinery, New York, NY, USA, 321–324.

https://doi.org/10.1145/2593968.2610482

[66] Markel Vigo and Simon Harper. 2013. Coping tactics employed by visually disabled users on the web. *International Journal of Human-Computer Studies* 71, 11 (2013), 1013 – 1025. https://doi.org/10.1016/j.ijhcs.2013.08.002

[67] Thorsten Völkel and Gerhard Weber. 2008. RouteCheckr: Personalized Multicriteria Routing for Mobility Impaired Pedestrians. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility* (Halifax, Nova Scotia, Canada) *(Assets '08)*. Association for Computing Machinery, New York, NY, USA, 185–192. https://doi.org/10.1145/1414471.1414506

[68] Kristin Williams, Taylor Clarke, Steve Gardiner, John Zimmerman, and Anthony Tomasic. 2019. Find and Seek: Assessing the Impact of Table Navigation on Information Look-up with a Screen Reader. *ACM Trans. Access. Comput.* 12, 3, Article 11 (Aug. 2019), 23 pages. https://doi.org/10.1145/3342282

[69] Yeliz Yesilada, Robert Stevens, Carole Goble, and Shazad Hussein. 2004. Rendering Tables in Audio: The Interaction of Structure and Reading Styles. In *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility* (Atlanta, GA, USA) *(Assets '04)*. ACM, New York, NY, USA, 16–23. https://doi.org/10.1145/1028630.1028635