Understanding and Reducing the Challenges Faced by Creators of Accessible Online Data Visualizations

Ather Sharif
Paul G. Allen School of Computer Science & Engineering | DUB Group, University of Washington
Seattle, Washington, USA
asharif@cs.washington.edu

Joo Gyeong Kim*
Paul G. Allen School of Computer Science & Engineering | DUB Group, University of Washington
Seattle, Washington, USA
kimjoogy@cs.washington.edu

Jessie Z. Xu*
Paul G. Allen School of Computer Science & Engineering | DUB Group, University of Washington
Seattle, Washington, USA
jessiexu@cs.washington.edu

Jacob O. Wobbrock
The Information School | DUB Group
University of Washington
Seattle, Washington, USA
wobbrock@uw.edu

ABSTRACT

We sought to understand and reduce the challenges creators face with making their data visualizations accessible. Specifically, we administered a formative survey of 57 creators to comprehend their challenges, perceived importance, knowledge, and prioritization of data visualization accessibility. Participants identified five interventions to minimize their challenges: *Workshops*, *Emulators*, *Evaluators*, *Feedback Collectors*, and *Multi-Modal Automated Tools*. Additionally, we report specifications and recommendations from 12 visualization creators for effective versions of each intervention, gathered via semi-structured interviews. Utilizing our findings, such as a "mini-survey" format that is effective for collecting accessibility-related feedback from screen-reader users, we implemented and integrated these interventions into VoxLens (Sharif *et al.*, 2022). We assessed our enhancements through a task-based user study with 10 visualization creators, finding 44%, 17%, and 12% improvements in their understanding of screen-reader users' challenges with data visualizations, knowledge of visualization accessibility, and perceived usefulness of the enhanced VoxLens, respectively.

CCS CONCEPTS

• **Human-centered computing → Information visualization**; **Accessibility systems and tools**; • **Social and professional topics → People with disabilities**.

KEYWORDS

visualization creators, challenges, data visualizations, accessibility, screen-reader users, blind

*These authors contributed equally to this work.

ASSETS '24, October 27–30, 2024, St. John's, NL, Canada

ACM ISBN 979-8-4007-0677-6/24/10
https://doi.org/10.1145/3663548.3675625

ACM Reference Format:
Ather Sharif, Joo Gyeong Kim, Jessie Z. Xu, and Jacob O. Wobbrock. 2024. Understanding and Reducing the Challenges Faced by Creators of Accessible Online Data Visualizations. In *The 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24), October 27–30, 2024, St. John's, NL, Canada.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3663548.3675625

1 INTRODUCTION

Lack of access to the underlying information in online data visualizations can be an inequity issue and severely disenfranchise screen-reader users[1] [16, 49, 52, 66], of which there are over 7.6 million in the United States, and who rely on visualization creators to make online data visualizations accessible. Empirical findings show that even the most commonly used technique in data visualization accessibility, alternative text ("alt-text"), is often missing or insufficient, causing screen-reader users to extract information 61% less accurately and spend 211% more time than non-screen-reader users [66, 90]. Recently, researchers have provided increased accessibility awareness for visualization creators, especially after inaccessible COVID-19 graphs, maps, and charts resulted in health concerns for the blind and low-vision community [25, 61]. Consequently, the visualization community has undertaken steps to address these accessibility concerns. For example, *The New York Times* recently created the new job position of an *Accessibility Visuals Editor*, which had never before existed [14, 20].

However, despite these efforts, accessibility[2] in data visualizations is often non-existent and significantly varies based on visualization creators' expertise and subject-matter familiarity. Consequently, this inaccessibility results in inconsistent and cumbersome interactions with online data visualizations for screen-reader users. Although researchers have proposed solutions to make data visualizations accessible, such as auto-generated alt-text [51, 67], sonification [4, 62], 3-D printing [11, 37], data tables [18], and multi-modality[3] [69, 79], investigating and *minimizing* the challenges

[1]People who use screen readers (*e.g.*, JAWS [65]) and might have complete or partial blindness, low vision, learning disabilities, or motion sensitivity [66, 70].
[2]We follow prior work's [7, 66, 69, 79, 90] definition of "accessibility" in this paper, *i.e.*, provision of access to information contained in data visualizations to screen-reader users in an efficient manner.
[3]We consider "multi-modality" as any combination of alt-text, sonification, tables, and voice-based question-and-answering, similar to prior work [7, 69, 79].

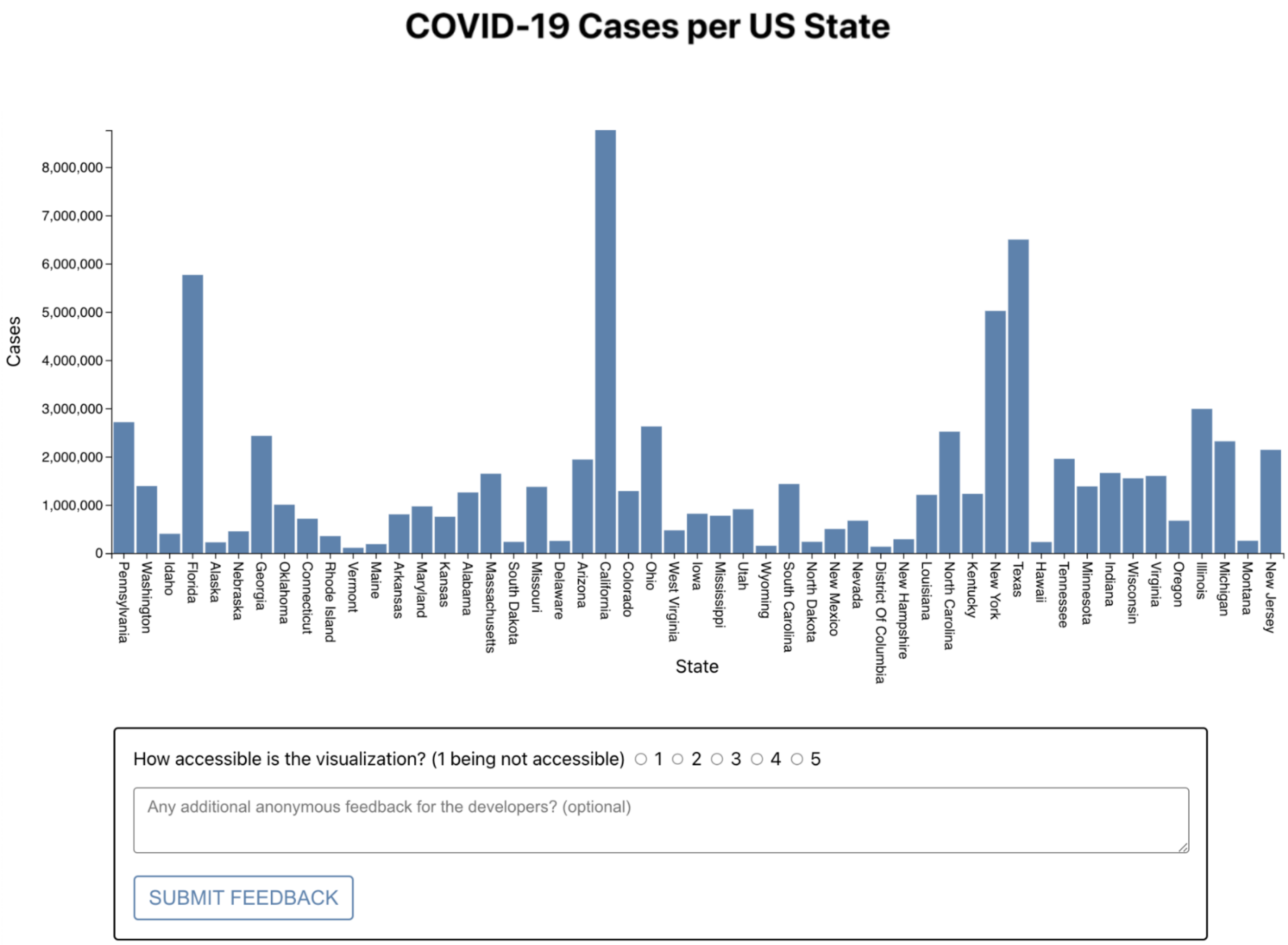

Figure 1: Feedback Collector, one of four new technology features added to VoxLens based on 12 semi-structured interviews with visualization creators.

for visualization creators when making online data visualizations accessible remains unexplored. While Joyner *et al.* [39] shed preliminary light on developers' challenges with accessibility, our work is the first to reduce those challenges by identifying interventions to: (1) *improve* developers' understanding of the experiences of screen-reader users, (2) *enhance* developers' knowledge of data visualization accessibility, and (3) *assist* developers in making data visualizations accessible to screen-reader users.

To this end, we surveyed 57 visualization creators to understand their perceived importance, knowledge, and prioritization of data visualization accessibility. Our findings indicate that their perceived importance of visualization accessibility impacts their knowledge of it, and together, these factors significantly affect their prioritization of it, which influences the challenges they encounter when making visualizations accessible. We also inquired about creators' challenges with accessibility and tools that could reduce those challenges. Our analysis revealed five interventions that, when used in combination with each other, elevate creators' understanding of screen-reader users' interactions with visualizations, strengthen their knowledge of data visualization accessibility, and aid them in making data visualizations accessible. These five interventions include one educational intervention (*Workshops*) and four technological interventions (*Emulators*, *Evaluators*, *Feedback Collectors*, and *Multi-Modal Automated Tools*).

To gather further insights into the features that make these interventions effective, we conducted semi-structured interviews with 12 visualization creators, inquiring about their perception of each identified intervention. Specifically, we asked about the distinctive features that would enhance each intervention's effectiveness to: (1) improve their knowledge of accessibility and screen-reader users' experiences, and (2) assist them in creating accessible visualizations. Our findings identify the nuanced features, perceived challenges, specifications, and recommendations from visualization creators for effective versions of each intervention.

Then, to assess the generalizability of our findings and the effectiveness of these interventions, we performed a task-based user study with 10 visualization creators. Specifically, we implemented each of the four technological interventions, integrating them into VoxLens, an open-source JavaScript plug-in that makes online data visualizations accessible by offering a multi-modal solution for screen-reader users [69–71]. We compared our enhanced version's performance with the original version of VoxLens. Our enhancements increased participants' understanding of screen-reader users' challenges with data visualizations, knowledge of data visualization accessibility, and perceived usefulness of the enhanced VoxLens by 43.8%, 16.7%, and 11.5%, respectively.

The contributions of this work are:

(1) Empirical results from our survey of 57 and interviews of 12 visualization creators, respectively. Emerging from these studies are five interventions (*Workshops*, *Emulators*, *Evaluators*, *Feedback Collectors*, and *Multi-Modal Automated Tools*) that enhance creators' experiences with data visualization accessibility.
(2) Implementations of each of the four technological interventions to extend the capabilities of VoxLens, an open-source JavaScript library to make data visualizations accessible to screen-reader users [69]. Our source code is available in our open-source GitHub repository.[4]
(3) Empirical results from our task-based user study with 10 visualization creators conducted to assess the generalizability of our findings and performance of our enhanced version compared to the original version of VoxLens.

2 RELATED WORK

We spotlight prior research that has contributed to data visualization accessibility and provided insights into the challenges visualization creators experience in making data visualizations accessible to screen-reader users.

2.1 Importance of Data Visualization Accessibility

Several researchers have emphasized the importance of making data visualizations accessible by highlighting the inequities caused by inaccessible visualizations [24, 40, 44, 48, 50, 52, 66]. Most recently, Keilers *et al.* [40] surveyed 45 blind and low-vision adults to explore data visualization accessibility on computers, phones, tablets, paper, and TVs, finding that insufficient accessibility practices significantly impact these users to access the underlying data in visualizations. Similarly, Sharif *et al.* [66] conducted experiments with 72 users and reported that not only do screen-reader users spend 211% more time on data visualizations, they also extract information 61% less accurately from visualizations than non-screen-reader users.

As Marriott *et al.* [52] stated, the lack of access to the information contained in visualizations is a significant equity issue for screen-reader users. Therefore, recognizing the critical need for accessible visualizations, researchers have provided recommendations to visualization creators [13, 22, 50, 57, 66, 77], including auto-generating alternative text, multi-modality, participatory design, and appropriate use of Accessible Rich Internet Application (ARIA) attributes.

2.2 Tools to Make Data Visualizations Accessible

Following the recommendations in prior work, several researchers have created tools to make data visualizations accessible to screen-reader users. These tools include auto-generating alternative text [42, 51, 56, 67], sonification [2, 4, 26, 36, 54, 62, 69, 74, 89], summarization [41], tables [18], haptic graphs [81, 87], 3-D printing [11, 37, 72], and multi-modality [7, 69, 79]. Most recently, Thompson *et al.* [79] introduced ChartReader, an open-source prototype accessibility engine that renders accessible data visualizations. They created their tool following an iterative co-design study with 10 Microsoft employees and reported the evolution of the design of ChartReader during this five-month study. Blanco *et al.* [7] built an open-source library called Olli that converts visualizations into a keyboard-navigable structure accessible to screen-reader users, which enables visualization creators to easily create accessible visualizations across various toolkits, including Vega-Lite [64]. Similarly, Sharif *et al.* [69] developed VoxLens, an open-source JavaScript plug-in that enables screen-reader users to interact with online data visualizations using a multi-modal approach, assisting them in obtaining data through sonification, summary, and verbal querying (question and answer). They reported a 164% and 50% improvement in screen-reader users' information extraction and interaction times, respectively, compared to conventional methods, through multiple user studies with over 100 users over time.

2.3 Accessibility Challenges Faced by Visualization Creators

Several researchers have recognized the obstacles to making data visualizations accessible [43, 59, 73, 77, 86]. However, to our knowledge, only a few have studied challenges that visualization creators face in making data visualizations accessible to screen-reader users [39, 80]. Most relevant to our work is the exploration by Joyner *et al.* [39], in which they surveyed 144 developers and conducted follow-up interviews with 10 selected respondents to understand the rationale and context behind the design choices of visualization creators. Their findings provide insight into visualization creators' challenges, knowledge, and prioritization of making their data visualizations accessible.

In contrast, our work is the first to identify educational and technological interventions to *reduce* the challenges visualization creators experience with making online data visualizations accessible while additionally shedding light on their challenges with data visualization accessibility. Furthermore, we complement Joyner *et al.*'s work [39] by quantitatively analyzing our survey respondents' ratings of their accessibility challenges, knowledge, and prioritization. We also provide the results from our task-based user study that we conducted to assess and validate our findings.

3 ONLINE SURVEY

In this work, our goal was to investigate visualization creators' challenges with data visualization accessibility, identify the factors inducing these challenges, and determine solutions to minimize

[4]https://github.com/athersharif/voxlens

these challenges. Therefore, we surveyed 57 visualization creators. We used a mixed-methods approach.

3.1 Participants

We surveyed 57 visualization creators (M=35.3 years old, SD=8.4). We advertised our survey through social media platforms and various email distribution lists for visualization creators. Among our respondents, 28 identified as women, 24 as men, four as non-binary, and one did not disclose their gender identity. Thirty-one participants were affiliated with industry, 20 with academia, and six with both. Twenty-eight participants specified themselves as developers, 19 as researchers, and 10 as both. The highest level of education was a doctoral degree for 20 participants, a master's degree for 14, an undergraduate degree for 20, an associate's degree for two, and a high school diploma for one participant. For visualization accessibility knowledge, three participants identified as "beginner," seven as "moderate beginner," 10 as "advanced beginner," 10 as "competent," 19 as "proficient," six as "advanced proficient," and two as "expert."

3.2 Procedure

Our survey included five steps. Participants filled out each step without supervision. In the first step of the survey, we displayed the purpose of our study, eligibility criteria, and data anonymity clause. Additionally, we collected participants' demographic information, including their gender [76], preferred pronouns, age, education level, domain (*DO*; "industry," "academia," or "both"), and role (*RO*; "developer," "researcher," or "both"). For clarity, the term "developer" encompassed similar terms, including "programmer" and "designer."

Next, in the second step, we asked about their current practices and knowledge (*KR*) about visualization accessibility. We also asked them about their accessibility prioritization (*PR*) and perceived importance (*IR*), accessibility policy enforcement (*PO*; "no policies," "policies recommended," or "policies enforced") at their respective organizations, and their strategies to make data visualizations accessible to screen-reader users. Additionally, we asked our participants about their frequency of testing their visualizations with screen-reader users (*SF*). We used a Likert scale ranging from 1 (lowest; e.g., "not important") to 7 (highest; e.g., "extremely important") for collecting *KR*, *PR*, *IR*, and *SF*.

In the third step, we inquired about the challenges they experience in incorporating accessibility in visualizations and their Likert-scale rating for the challenge level (*CR*). Then, in survey step four, we asked the following open-ended questions to gather insights about interventions that could mitigate participants' challenges:

(1) Which resources or tools could enhance creators' understanding of screen-reader users' challenges?
(2) Which resources or tools could help enhance creators' knowledge regarding accessible online data visualizations?
(3) Which resources or tools could help make online data visualizations accessible to screen-reader users?

Finally, in survey step five, we asked for any optional comments and participants' interest in a follow-up interview.

3.3 Quantitative Evaluation

To quantitatively analyze our survey responses, we investigated the factors that contribute to visualization creators':

Table 1: Statistical results from our ordinal logistic regression analysis of N=57 survey responses. "IV" means the Independent Variable. "BIC" is the Bayesian Information Criterion (lower is better). "CR" = *Challenge Rating*, "PR" = *Prioritization Rating*, "DO" = *Domain*, "KR" = *Knowledge Rating*, "SF" = *Studies Frequency*, "IR" = *Importance Rating*, and "PO" = *Policies*. Results with p<.05 are statistically significant.

Model	IV	N	χ^2	p
$CR \sim PR + DO + age$ BIC = 190.76	*PR*	57	19.68	< .05
	DO	57	6.62	< .05
	age	57	3.18	.075
$KR \sim SF + IR + age$ BIC = 229.49	*SF*	57	22.45	< .001
	IR	57	19.68	< .05
	age	57	8.52	< .05
$PR \sim KR + IR + PO$ BIC = 232.67	*KR*	57	12.86	< .05
	IR	57	13.32	< .05
	PO	57	5.19	.075

- Challenges with implementing accessibility in data visualizations (RQ1)
- Knowledge of data visualization accessibility (RQ2)
- Prioritization of making data visualizations accessible (RQ3)

We used stepwise ordinal logistic regression (both directions) [19, 21, 35] to identify possible outcomes of our independent variables out of the following candidate variables: *SF*, *KR*, *IR*, *PR*, *DO*, *RO*, *PO*, and *age*. As a result, for RQ1, our dependent variable (DV) was *CR*, and our independent variables (IVs) were *PR*, *DO*, and *age*. For RQ2, our DV was *KR*, and our IVs were *SF*, *IR*, and *age*. For RQ3, our DV was *PR*, and our IVs were *KR*, *IR*, and *PO*.

Variables *CR*, *SF*, *KR*, *IR*, and *PR* had an ordinal representation (1 to 7 on a Likert scale), whereas *DO*, *RO*, and *PO* were trichotomous. We investigated the effect of our independent variables on our dependent variables using ordinal logistic regression [53, 55], a standard technique for analyzing ordinal responses. Table 1 shows our statistical results.

Variables *PR* and *DO* had a significant effect on *CR*. Specifically, 46.7% (N=21) of the respondents who regarded implementing accessibility as at least somewhat challenging (5–7 on the Likert scale) rated their prioritization as 5–7, and 60.0% (N=27) reported their affiliation with "industry." *Age* had only a marginal effect on *CR* ($p \approx$.075).

SF, *IR*, and *age* had a significant main effect on *KR*. Specifically, 3.7% (N=1) and 88.9% (N=24) of the respondents who considered themselves at least somewhat knowledgeable in accessibility (5–7 on the Likert scale) rated their frequency of conducting studies with screen-reader users and accessibility importance a 5–7. Additionally, 74.1% (N=20) of these respondents reported being between 30 and 50 years old.

KR and *IR* had a significant main effect on *PR*. Specifically, 63.3% (N=19) and 96.7% (N=29) of the respondents who at least somewhat prioritized accessibility (5–7 on the Likert scale) rated their accessibility knowledge and importance a 5–7. *PO* had only a marginal effect on *PR* ($p \approx .075$).

As a supplementary analysis, we performed Spearman's non-parametric rank correlation [75] to investigate the linear relationship between all our variables. We found a positive correlation between *KR* and *PR* ($p<.001$), *KR* and *PO* ($p<.05$), *KR* and *SF* ($p<.05$), *IR* and *PR* ($p<.001$), *PR* and *PO* ($p<.001$), *PO* and *SF* ($p<.05$), and *DO* and *RO* ($p<.05$). (See Appendix A for statistical results.) This analysis was supplementary to this work and these results do not imply causality. However, these findings provide avenues for future research to determine possible causal relationships among these variables to shed further light on visualization creators' challenges with data visualization accessibility.

3.4 Qualitative Evaluation

We qualitatively analyzed our participants' (N=57) open-ended responses to questions in Step 4 of our survey. We conducted a theoretical thematic analysis [9] using a semantic approach [58] and an essentialist paradigm [60, 84], following guidelines from Braun and Clarke [10]. Three co-authors coded each response. We calculated the inter-rater reliability using pairwise percentage agreement, reaching a high agreement percentage of 97%. Our analysis gave rise to one educational intervention (*Workshops*) and four technological interventions (*Emulators*, *Evaluators*, *Feedback Collectors*, and *Multi-Modal Automated Tools*) identified by creators that would reduce their challenges in making data visualizations accessible to screen-reader users. We discuss these below. (For a qualitative analysis of visualization creators' experiences with data visualization accessibility, we direct readers to work by Joyner *et al.* [39].)

3.4.1 Workshops. For simplicity, we use the umbrella term "workshops" to represent synonymous terms that our participants used in their responses, including "courses," "trainings," and "tutorials." Our participants expressed an interest in attending workshops to gain "hands-on" experience (P52) to understand challenges that screen-reader users encounter with online data visualizations and to enhance their knowledge of data visualization accessibility. For example, P5 shared their enthusiasm for taking part in and effectively advertising a workshop:

> `I would LOVE to take a class taught by someone who uses these features. An effective way might be to raise awareness of such courses in U.S. federal programs.` (P5)

Additionally, P41 discussed the scarcity of such workshops at her organization:

> `Workshops and tutorials that specifically address the experiences of screen-reader users are helpful, but scarce at my institution.` (P41)

3.4.2 Emulators. In this paper, we define an "emulator" as software that mimics responses from screen readers to assist visualization creators in hearing the screen reader output without needing to install a screen reader on their computer. Our participants stated their desire to understand the interaction experiences of screen-reader users with data visualizations through emulating the output from screen readers. For example, P23 considered an emulator a "middle-ground task":

> `I wish there were a middle-ground task between running a screen reader myself and hiring someone to use a screen reader for me.` (P23)

Similarly, P47 and P51 identified an emulator's benefits as "quickly previewing the experience" and "instantly hearing the visualization from a screen reader perspective," respectively. Additionally, P21 highlighted its use case as a centralized tool that would assist in self-auditing data visualizations:

> `Centralized walk-throughs of different screen reader modalities would be cool. How blind and visually impaired users use the screen reader.` (P21)

3.4.3 Evaluators. Unsurprisingly, our participants identified the need for tools that "evaluate" their data visualizations for accessibility. Several tools currently exist that evaluate data visualizations' accessibility [3, 22, 78, 83] based on established guidelines, such as WCAG 2.1 [12]. However, as our participants identified in their responses, these solutions are not integrated into the visualization libraries and do not provide real-time evaluation of visualizations during the development process. For example, P49 communicated the need for automated evaluators by providing relatable examples of existing solutions:

> `Having a tool that checks your visualizations as you're working... Like, a Grammarly-type of resource, but to improve the accessibility of your visualization!` (P49)

Similarly, P48 shared his challenges with generating alternative text and identified a specific use case for evaluators:

> `I've been relying on alt-text and I feel that I have spent a lot of time and effort into this but I don't have a way to say, "okay, what you're doing is sufficient!" So a checker (e.g., alt-text checker) would be great!` (P48)

3.4.4 Feedback Collector. A common theme in our survey responses was visualization creators' desire to get feedback on the accessibility of their data visualizations from "real-life users" (P29) and "access to screen-reader users for input" (P46). (In this paper, we refer to this utility as "feedback collectors.") For example, P54 accentuated the need to connect with screen-reader users to receive feedback:

> `I'd like to connect with more people with disabilities to test our designs.` (P54)

Similarly, P35 highlighted the overall usefulness of a feedback collector, emphasizing protecting users' privacy:

> `I think writing to report a bug or enhancement should be easy. A tool to enable users to report without losing privacy or sensitive information would be valuable.` (P35)

3.4.5 Multi-Modal Automated Tool. A "multi-modal automated tool" in data visualization accessibility is software or a plug-in that provides multiple modalities for users to extract information,

Table 2: Gender, age, domain, role, and education level of our interview participants.

	Gender	Age	Domain	Role	Education
S1	Woman	29	Industry	Developer	Master's
S2	Woman	32	Industry	Developer	Bachelor's
S3	Woman	35	Academia	Both	Doctorate
S4	Man	40	Industry	Researcher	Doctorate
S5	Woman	27	Academia	Researcher	Master's
S6	Man	37	Industry	Developer	Bachelor's
S7	Woman	31	Industry	Developer	Bachelor's
S8	Man	33	Industry	Researcher	Doctorate
S9	Man	54	Both	Both	Bachelor's
S10	Man	49	Industry	Developer	Master's
S11	Man	36	Industry	Developer	Bachelor's
S12	Man	22	Academia	Researcher	Bachelor's

such as sonification, alt-text, and tables. A prevalent mention in our survey responses was that of an "easy to create" (P44) and "auto-implement" (P18, P46) multi-modal solution. Additionally, the visualization creators emphasized the need for such solutions to be "open-source" (P29) and "financially accessible" (P25). Notably, P53 classified an automated multi-modal solution as a "magic wand":

> `A magic wand... a charting tool where a chart is just made accessible in the background without the chart author needing to do anything special.` (P53)

P6 identified ways and specific domains where a multi-modal solution would be beneficial:

> `[It would be useful as] R/ggplot2 and python/seaborn extensions to automatically sonify data and/or generate alt-text, or a web-based library that can be built-in automatically.` (P6)

Additionally, our participants suggested using artificial intelligence (P2, P28, P42), computer vision (P52), and technologies such as ChatGPT (P26) and large-language models (P38) for creating multi-modal solutions.

Overall, our findings shed light on the challenges visualization creators experience with making data visualizations accessible, identifying factors contributing to these challenges. In particular, our qualitative results determine an educational intervention and four technological interventions that might assist visualization creators in minimizing these challenges. These findings motivated the need to conduct in-depth interviews with visualization creators to delve further into the specific features that improve the effectiveness of the recognized interventions. We present our methodology and findings from the interviews with visualization creators in the section below.

4 INTERVIEWS WITH VISUALIZATION CREATORS

To better understand the creators' perceived effectiveness of features for each intervention (*Workshops*, *Emulators*, *Evaluators*, *Feedback Collectors*, and *Multi-Modal Automated Tools*), we conducted semi-structured interviews with 12 visualization creators. We present the details and results of our qualitative analysis.

4.1 Participants

We randomly selected 12 survey respondents from those who indicated they would participate in a follow-up interview (see Table 2). Five participants identified their gender as women and seven as men. Their average age was 35.4 years (*SD*=9.0). Three participants had attained or were pursuing a doctoral degree, three a master's degree, and the remaining six a bachelor's degree. We discontinued recruiting participants once we reached saturation of insights. We compensated participants with a $25 Amazon gift card for participating in our hour-long interview.

4.2 Procedure

We interviewed our participants via Zoom and used its built-in features to record the calls, subsequently transcribing the interviews using Descript [17]. Our study sessions were semi-structured and lasted one hour. At least three authors administered each interview, one taking detailed notes during each session. We engaged our participants in a conversation to identify the utility, challenges, and suggestions to enhance the effectiveness of the interventions. Specifically, we explored each intervention's features that could assist visualization creators in: (1) understanding screen-reader users' challenges with data visualizations, (2) improving creators' knowledge about data visualization accessibility, and (3) making creators' data visualizations accessible to screen-reader users.

4.3 Analysis

To analyze our interviews, we used inductive thematic analysis [10]. Specifically, we followed a semantic approach [58] and an essentialist paradigm [60, 84]. As per prior work's guidelines on thematic analysis [10, 63], we used the first two interviews to develop an initial set of codes. We added codes to our codebook, when appropriate, during our analysis of the rest of the interviews, identifying 83 total codes. Three researchers coded each interview transcript independently, resolving disagreements through discussions. Our final analysis identified the utility, challenges, and suggestions for enhanced effectiveness for all five interventions, discussed in the section below.

We calculated inter-rater reliability using Krippendorff's α [45] as well as pairwise percentage agreement, following the recommendation by Landis *et al.* [47]. Computed using ReCal 3.0 [27], our Krippendorff's α was 0.88, indicating a high level of reliability [46]. The average pairwise percentage agreement between authors was also high, at 88.2% [28, 32].

4.4 Results

We present the findings from our qualitative analysis of our semi-structured interviews with 12 visualization creators. Specifically,

Table 3: Interventions and their effectiveness features identified during our survey of and interviews with visualization creators.

Intervention	Features
Workshops	*Synchronous*
	Co-instructed by screen-reader users and practitioners
	Materials available post-workshop
	Step-by-step instructions with examples
Emulators	*Output in textual format*
	Incorporate output from multiple screen readers
	Toggle activation (on/off)
Evaluators	*Score-based evaluation*
	Explanation of the underlying issue
	Evaluation of presence as well as quality of alt-text
	Evaluation of color contrast ratio
	Toggle activation (on/off)
	Post-creation evaluation (CI/CD pipelines, "linters")
Feedback Collectors	*Simple radio buttons (Likert scale)*
	Optional open-ended field for feedback
	Tracking feedback (GitHub tracker)
Multi-Modal Automated Tools	*Customization and personalization options*
	A.I. techniques for efficiency
	Low-effort avenues to test technical accuracy

we discuss each of the five identified interventions in turn. We show these findings in Table 3.

4.4.1 Workshops. Our participants preferred synchronous over asynchronous workshops, recognizing several benefits of in-person workshops, including the ability to ask questions in real-time, the value of discussions and collaborations with fellow attendees, and a desire to interact with screen-reader users. For example, S7 identified the advantages of real-time interactive workshops:

> `I find that to be pretty useful to have a person in real-time. You can ask your questions and have all of that happen in real-time. Because you wanna make sure those charts are accessible, it's nice if you can go ask someone.` (S7)

S3 agreed, saying:

> `[In] an in-person workshop, there's more motivation to be present and actually building collaborations and general interactions with other people that might also be interested in that space.` (S3)

Similarly, S6 discussed the importance of interacting with and learning from screen-reader users in workshops:

> `If there was a workshop given by someone who uses a screen reader and they give that firsthand perspective and show how they use their screen reader, challenges they run into... If I were to learn more about accessibility and understanding the challenges, having that sort of direct instruction is very valuable.` (S6)

Our participants also highlighted the importance of learning from the practitioners, emphasizing the need for "multiple expertise levels" (S5) and "relatability" (S3). For example:

> `The people I'd like to learn from are people who are sighted and trying to do the same things that I'm doing. Like practitioners, who know what the steps are.` (S1)

Additionally, some participants specified time prioritization as a challenge for attending in-person workshops, addressable by making the materials available post-workshop. S5 identified this challenge and also offered a solution:

> I think the struggle is always just finding time. My ideal format is having the materials be available asynchronously. It takes the burden off of having to be somewhere at a specific time and you feel like you're missing out on content. So there's something just really nice about that format that I feel like you kind of get the best of both worlds. (S5)

Additionally, P11 offered content suggestions and emphasized the need for step-by-step instructions with examples:

> Step-by-step... explain [it] like I'm five. "Accessible visualization for dummies." What is ARIA? What fields are available to me to use for screen readers? How can I actually make Tableau work with a screen reader? Give me viable examples of steps to follow. (P11)

Overall, our participants found workshops on data visualization accessibility effective in a synchronous format, in which they could collaborate with other attendees in person and ask questions. They preferred the workshops to be co-instructed by screen- and non-screen-reader users to leverage varying knowledge and expertise. Additionally, they suggested making content available after the workshops as a workaround for time prioritization challenges.

4.4.2 Emulators. Our participants found emulators essential for understanding the experiences of screen-reader users. They identified several use cases for emulators, including assisting visualization creators in comprehending how screen readers "read" their visualizations, minimizing the need to install and learn several screen readers, and enhancing their accessibility testing experience. For example, S9 said:

> The simple fact that there is more than one [screen reader] out there and that they would need to be installed is a barrier to somebody like me. I mean, obviously, I can just go and install lots of different ones, but then I think if I knew I had the option to not install and simply go to one place, which is going to emulate the experience, I would start there for sure. (S9)

S2 expressed a similar sentiment:

> I'll say testing with screen readers has always just been a little difficult. For somebody just trying to quickly see in general what's happening with my data through a screen reader, you could use it for that case. (S2)

S12 expressed a suggestion to emulate the output of a screen reader in textual format:

> I would appreciate if it could just generate text so I can just look at the output of a screen reader. So you kind of wanna first take away the narration part of it and [just] have it in a transcript kind of form. (S12)

Unsurprisingly, some of our participants preferred using the actual screen readers for a better understanding of the experiences of screen-reader users while also identifying use cases where emulators could be beneficial. For example, S11 advocated for using the screen reader itself while also recognizing situations where an emulator could assist him:

> This is really hard because there is a little bit of me that thinks, "Why not enable the screen reader in the first place?" But I could definitely see a use case for having an emulator because at the moment I've [got] VoiceOver running and I wanna see how it works in NVDA. And I can't really do that very easily. So yeah, actually, to be honest, there's definitely a use case. (S11)

Our participants also emphasized that emulators would be increasingly useful if they incorporated the ability to emulate the responses from various screen readers, including users' customizations and preferences. For example, S12 shared his enthusiasm for this feature:

> There are different screen readers on the market, and visually impaired groups, they tend to use different tools according to their preferences. So if this emulator can account for all different screen readers, I think that would be great. (S12)

S2 said she also wanted an emulator to combine screen readers:

> So, obviously if there's a way to kind of combine all of those screen readers into, you know, this emulator, and you were able to flip to different screen readers to test the nuances, that might be cool. (S2)

In summary, our participants considered emulators a beneficial intervention to understand screen-reader users' experiences. They also noted that learning a screen reader rather than using an emulator could have potential advantages. However, they recognized that due to the technical and learning challenges involved with installing and using several screen readers, emulators would be a plausible alternative to test visualization accessibility. We note that our findings are particularly directed toward emulating the interactions of screen-reader users with data visualizations; these findings might not be generalizable to their interactions with other digital content.

4.4.3 Evaluators. In our interviews, creators identified the usage of evaluators as an educational toolkit. For example, S1 recognized the benefit of evaluators to improve knowledge of accessibility:

> That's a common mistake that people don't put alt-texts, right? So, say it starts to catch that and show that, hey, well no alt-text. So then it's slowly, in a way, teaching you a little bit and improving your knowledge of accessibility. (S1)

Additionally, our participants highlighted essential features to enhance the effectiveness of evaluators, including a score-based evaluation and an explanation of the underlying accessibility issues with visualization elements. S9 discussed the benefits of a score-based evaluation:

> Well, definitely having an overall score is a good thing. Scoring maybe each element or assessing each element so you can make design choices. (S9)

Our participants also provided recommendations to include nuanced evaluations of accessibility measures such as alternative text ("alt-text"), color contrast ratio, and the ability to toggle the activation of automated checkers. For example, S3 shared her thoughts about alternative text:

> Appropriate alt-text... You have to be able to recognize what it is and for it to be useful, prompt somebody to actually include meaningful elements in the alt-text. (S3)

S1 shared the importance of toggling features for the evaluator:

> Depending on where you don't want a certain behavior and if it keeps giving you an error, then you just get annoyed. Then you don't want to use it anymore. So being able to either dismiss it or [have] a "linter" where you can turn off rules. (S1)

Furthermore, some participants were interested in evaluating visualizations post-creation in addition to *during* visualization development, such as through Continuous Integration/Continuous Delivery ("CI/CD") pipelines and "linters" (code analysis tools that report programming errors):

> I would want it as part of a test system or a "linting" system. Cause I always see those run on most CI/CD projects before it gets to deployment. I think it'd be good to give that red flag before anyone gets a chance to see and encounter it. (S7)

In summary, our participants identified the use case for evaluators as an educational toolkit for visualization creators, particularly novice creators, to learn about the underlying accessibility issues with their visualizations. They also noted several features of an effective evaluator, including evaluating alt-text quality and color contrast ratio and using score-based methods during and after visualization creation, such as in CI/CD pipelines.

4.4.4 Feedback Collectors. Our participants expressed a keen interest in "welcoming feedback" (S5) on the accessibility of their visualizations from screen-reader users. S9 shared the enthusiasm of the visualization community for this idea:

> It's a fantastic idea! I don't know anybody who's working in this field who doesn't welcome feedback. It helps you refine, improve your work. In the visualization world, everybody wants to get better at making their data visualizations accessible. (S9)

As expected, some participants acknowledged and noted the undue burden screen-reader users face in reporting accessibility issues online. For example, S5 said:

> We think a lot about user burnout and burdens around constantly having to provide feedback. So, I want to be cognizant of that and not overwhelm screen-reader users who are wanting to provide feedback. (S5)

To make the feedback collector effective and minimize the burden on screen-reader users providing feedback, our participants suggested using simple radio-button options with an optional open-ended text field to collect the accessibility rating of a visualization. S5 shared some solutions to her concerns:

> Something simple, something easy with options to provide more feedback if they choose to, give them the agency to how much and not put the burden and onus on them to do a lot. So, as much as they want to provide is effective for you. (S5)

S11 shared his ideas for capturing feedback efficiently:

> Maybe you can make it a little easier, like, quick "yes" or "no." For example, was the visualization appropriate? You just want to keep it simple, but at the same time, you want an open-ended [option], like a mini-survey that they can fill out. (S11)

Additionally, our participants highlighted that tracking feedback would be a beneficial feature for screen-reader users. For example, S8 used GitHub's issue tracker as an example to identify the benefits of tracking feedback:

> [By] going to GitHub, you can do issue trackers or report a bug and then it gets posted. So there should potentially be some way for a dialogue with the creator and also notification of resolution of the feedback. (S8)

Overall, our participants emphasized the potential usefulness of feedback collectors, typified by the enthusiastic quote, "[E]verybody wants to get better at making their data visualizations accessible" (S9). Our participants suggested making feedback collectors simple, similar to a mini-survey, with an optional choice to provide open-ended responses. Additionally, they identified feedback tracking as a potentially beneficial feature.

4.4.5 Multi-Modal Automated Tools. Our participants expressed enthusiasm for multi-modal automated tools, identifying their inherent characteristic of being an "easy sell" (S11) for decision-makers and helpful for minimizing visualization creators' challenges with making online data visualizations accessible to screen-reader users. As S11 said:

> So the idea that, well, we have a solution, where we can make [visualizations] more accessible rather than rewriting them. I can't even begin to tell you how appealing that is and how much of an easier sell that would be. (S11)

S12 was similarly enthusiastic:

> It would be great. It would be absolutely awesome to have a tool that's sort of in the middle between the developer and screen reader; it would just be awesome. (S12)

As identified by our participants, customization and personalization options for screen-reader users would make multi-modal automated tools more effective. In particular, our participants recommended utilizing the default screen reader settings of the screen-reader users to aid in developing the personalized tool. S11 discussed his concerns with customization options and offered a solution:

> I worry about too many options and sensible defaults; are they gonna be too prescriptive? So instead, what if we took their preferences that they already have and used them? I definitely think that would certainly help. (S11)

Our participants noted that multi-modal automated tools offer limited opportunities to learn about accessibility, but also argued that the goal should be about screen-reader users being able to extract the information from visualizations:

> It'd be great if people knew about accessibility, but if everything was just magically accessible in the background and you didn't have to know, that's the best possible outcome for people who need the stuff to be accessible. So, I believe that having an automated tool that just fixes it for you is the most effective way. The outcome of accessibility should be [that] screen-reader users are able to do stuff. (S7)

Additionally, our participants recommended using research studies with screen-reader users to develop such tools:

> So, if it's a person who does accessibility research and they're the ones building that tool... So, research done with screen-reader users, and something that has been evaluated with them, assessed with them, taken their preferences into account, that would give you more credibility for the tool. (S1)

S6 suggested providing low-effort avenues for visualization creators to test for technical accuracy to build trust:

> I think if the tool was research-based and reacting to direct feedback from screen reader users, [and] best practices from developers, that makes it feel much more valuable. You know... research-based, and has been validated and vetted. (S6)

Altogether, our participants expressed their exhilaration for multi-modal automated tools, classifying them as an "easy sell" for decision-makers. They offered suggestions to customize and personalize these tools to enhance the experiences of screen-reader users. Additionally, they discussed the importance of using research methods and co-design approaches in building such tools with screen-reader users, considering it essential for building trust in the tool's accuracy.

Utilizing our findings, we implemented and integrated the four technological interventions into `VoxLens` [69]. Subsequently, we conducted a task-based user study to assess the utility of our findings in enhancing creators' work to make data visualizations accessible. We discuss our integration and assessment below.

5 INTEGRATION OF TECHNOLOGICAL INTERVENTIONS INTO VOXLENS

We extended the the open-source tool `VoxLens` [69] by implementing the four technological interventions identified by visualization creators in our aforementioned studies. Our goal in implementing and integrating these interventions into `VoxLens` was to assess the generalizability of our findings. Although these enhancements were open source and can assist in creating future solutions, we developed them within the context of an existing system—`VoxLens`. Therefore, these enhancements may not be generalizable to other platforms and tools, the exploration of which we leave to future work. Furthermore, as `VoxLens` is a technological solution, our implementation and integration focused on the identified technological interventions. (We plan to carry out the educational intervention, namely *Workshops*, as future work.) We discuss the four new technological interventions and their implementations below.

5.1 VoxLens Overview

We chose `VoxLens` [69] for our integration because `VoxLens` is (1) a multi-modal solution, (2) open-source, and (3) easy to integrate into existing data visualizations. Additionally, while commercial versions of some of our interventions exist, such as the WebAIM Contrast Checker [82], `VoxLens` provides us with the ability to *collectively* use these interventions *during* the development of online data visualizations via JavaScript libraries.

`VoxLens` is a JavaScript plug-in that improves the accessibility of online data visualizations for screen-reader users using a multi-modal approach [69–71]. It requires only a single line of code for integration into visualizations created using D3 [8], Google Charts [18], or ChartJS [15]. There are three modes of `VoxLens`: (1) *Question-and-Answer*, where the user verbally interacts with the visualization; (2) *Summary*, where `VoxLens` gives a summary of the underlying information; and (3) *Sonification*, where `VoxLens` enables listeners to interpret data by mapping it to a musical scale using the `Sonifier` library [68, 69].

5.2 Video Tutorial

We created a video tutorial for implementing `VoxLens`, further explaining the usage of configuration options available in both versions. We uploaded the tutorial to YouTube and provided its link to participants in our step-by-step instructions in the second part of our study. (Although a video tutorial can be an educational intervention, we do not consider it a replacement for *Workshops*.)

5.3 Debug Mode as Emulator

We introduced a "debug" mode in our `VoxLens` enhancements to emulate the responses from screen readers. This mode enables visualization creators to hear `VoxLens`'s responses *without* using

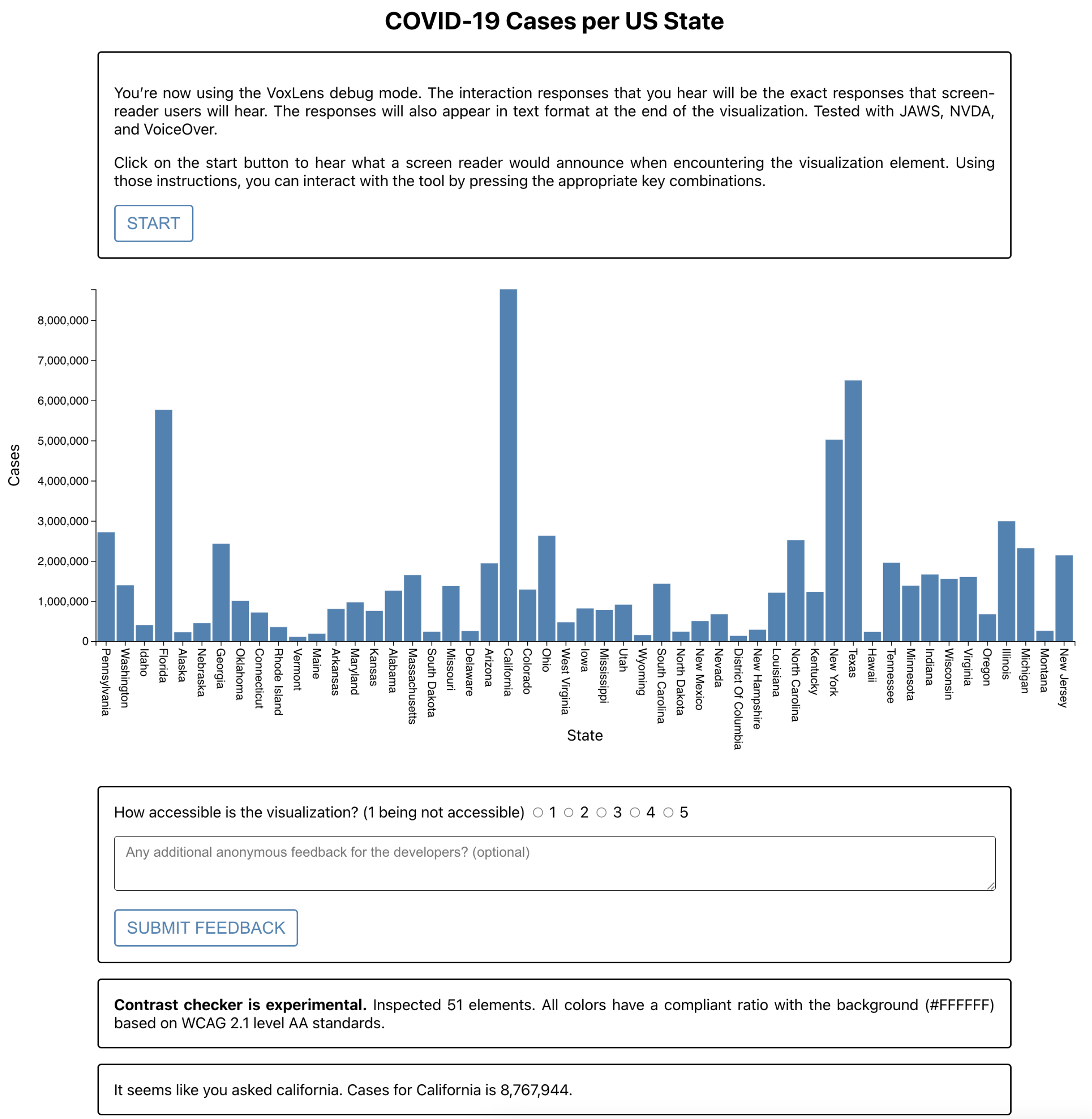

Figure 2: Screen capture of a sample COVID-19 data visualization with our VoxLens enhancements. Without our enhancements, only the visualization and its title would be shown to the visualization creators.

a screen reader, similar to prior work [6]. We relayed the audio output using the Web Speech API, commonly available in modern browsers. As VoxLens responses are consistent across various screen readers, including JAWS [65], NVDA [1], and VoiceOver [38], we implemented the following features from Table 3: (1) output in textual format, and (2) toggle activation (on/off). Specifically, visualization creators could deactivate the emulator (speech and text) or turn off the voice output to view the responses only in textual format. With these features, visualization creators can hear and see every piece of information that VoxLens relays to a screen reader, an additional capability that adds emulation and improves multi-modality of VoxLens, which it did not possess before.

5.4 Color Contrast Evaluator

To uphold our interview findings of an evaluator, we built a color contrast evaluator for visualizations that automatically extracts the background color of the visualization and assesses it against the colors of the visualization elements, including the lines, bars, shapes, and text, using the Level AA standards specified in WCAG 2.1 [12]. To extract the background colors from SVG-based visualizations, we traverse the elements recursively. For visualizations created using HTML Canvases, we convert the canvas element into an image and fetch the color palette using the `Color Thief` library.[5] For background-to-foreground ratios not in compliance with the Level AA standard, we provide the visualization creators with details of each element for easier debugging with an overall score. When in compliance, we show the number of elements inspected and confirmation of compliance (see Figure 2). Additionally, per our participants' preferences, we added the ability to disable the evaluator by setting the "contrastChecker" configuration option to "false" (see Appendix B), which turns this feature off.

As `VoxLens` automatically generates a summary, we did not implement an evaluator for checking the presence and quality of alt-text (see Table 3). Similarly, we did not implement post-creation tools, as these solutions are internal to the deployment environment of developers and outside of the technical scope of `VoxLens`.

5.5 Feedback Collector

We developed a feedback collector that enables screen-reader users to provide feedback to the visualization creator regarding the accessibility of the visualization. Based on the features identified by our participants (see Table 3), we limited the feedback to a simple rating of the visualization accessibility using radio buttons with an optional text field for further comments. By default, the rating is on a 5-point scale ranging from 1 ("not accessible") to 5 ("fully accessible"). Per their preferences, visualization creators can modify the scale and hide the feedback collector in debug mode. A valid email address is required from the visualization creators to activate the feedback collector. We collaborated with and utilized the services from `EmailJS`[6] to enable client-side email sending. We note that screen-reader users provide feedback anonymously to the visualization creators using our feedback collector without any additional risk to their privacy. We did not implement the tracking feature that enables users to get the real-time status of their reported issues, as it requires additional setup from visualization creators, which is outside of the technical scope of our exploration.

5.6 Algorithm Improvements for Multi-Modality

Our respondents suggested using advanced and contemporary algorithms for the multi-modal automated tools (see Table 3). Specifically, they recommended the use of artificial intelligence (A.I.). Therefore, we improved the functionality of `VoxLens` by using fuzzy logic [29, 33, 88], a natural language processing (NLP) technique, to replace the current keyword matching algorithm in use by the original `VoxLens`. Our internal testing revealed that fuzzy logic performed queries faster and more accurately than the keyword matching algorithm, improving the performance of `VoxLens`'s *Question and Answer* mode. We did not implement customization and personalization options, as this feature warrants a separate exploration, requiring additional empirical studies. We also did not implement avenues to test technical accuracy, as `VoxLens` already provides automated unit and functionality testing with 100% coverage.

Figure 2 displays a screen capture of our `VoxLens` enhancements, and Appendix B shows a comparison of the configuration options between the original `VoxLens` and our enhanced version.

Table 4: Gender, age, domain, role, and education level of our `VoxLens` task-based experiment participants.

	Gender	Age	Domain	Role	Education
V1	Woman	29	Industry	Developer	Master's
V2	Woman	19	Academia	Developer	Bachelor's
V3	Man	20	Academia	Researcher	Bachelor's
V4	Man	19	Both	Developer	Bachelor's
V5	Woman	29	Industry	Developer	Master's
V6	Man	33	Industry	Researcher	Doctorate
V7	Woman	49	Industry	Developer	Master's
V8	Woman	33	Industry	Researcher	Master's
V9	Man	54	Industry	Developer	Bachelor's
V10	Woman	23	Academia	Researcher	Doctorate

6 TASK-BASED USER STUDY USING VOXLENS

To evaluate the usability and usefulness of our `VoxLens` enhancements, which were based on our survey and interview findings, we conducted a task-based user study with 10 visualization creators using the original `VoxLens` and our enhanced version. Our goal from this study was to investigate the effect of our enhancements on our participants' understanding of screen-reader users' challenges, knowledge of data visualization accessibility, and perceived usefulness of the enhanced `VoxLens`. We present our methodology and results below.

6.1 Participants

Our participants were 10 visualization creators recruited via social media platforms and email distribution lists for data visualization creators (see Table 4). Six participants identified as women and four as men. Their average age was 30.8 years (*SD*=12.2). The highest level of education was a doctoral degree for two participants; four had a master's degree, and the remaining four had a bachelor's degree. We compensated them with a $100 Amazon gift card for their participation in our task-based user study.

[5]https://lokeshdhakar.com/projects/color-thief/
[6]https://www.emailjs.com/

6.2 Procedure

Our study was a two-part unsupervised task-based study. Each part included a step-by-step instruction for our participants to complete the task. The first part involved using the original VoxLens [69] without any modifications, following the instructions, documentation, and examples from its open-source GitHub repository. We provided instructions for participants to familiarize themselves with VoxLens before the study to account for learning. In the second part, our participants used our enhanced VoxLens version. Our goal was to assess the generalizability of our empirical findings as instantiated in VoxLens, not to assess the accessibility of visualizations themselves. Therefore, we used the same visualization in both parts and there was no difference in the accessibility of the visualization between the original and enhanced versions.

At the end of each part of the study, our participants filled out a questionnaire rating their understanding of the challenges screen-reader users experience with online data visualizations (UL), knowledge of data visualization accessibility (KL), and perceived usefulness of the enhanced VoxLens (PU) on a scale of 1 to 7 on a Likert scale (with "1" being the lowest and "7" being the highest). The questionnaire also included questions from the NASA-TLX workload instrument [30, 31] and open-ended queries to collect our participants' liked or disliked features, ideas for improvement, and general comments. We also recorded their demographic information.

6.3 Design & Analysis

We conducted a task-based user study to investigate the effect of our VoxLens enhancements on our participants' subjective responses. Our independent variable was VoxLens *Version* (VX; within-Ss.), having two levels ("original," "enhanced"). Our dependent variables were *Understanding Level* (UL), *Knowledge Level* (KL), and *Perceived Usefulness* (PU); all dependent variables were ordinal (1 to 7 on a Likert scale; "1" being the lowest and "7" being the highest). To analyze the effect of VX on each of these variables, we used mixed ordinal logistic regression [53, 55], identical to our quantitative survey analysis. This model is a type of generalized linear mixed model (GLMM). Our focus being a task-based user study and not a controlled experiment, we deliberately had all participants interact with the original version first, and then the enhanced version, since the latter contains a strict superset of functionality of the first.

To qualitatively assess the open-ended questionnaire responses, we followed the same protocol for thematic analysis [10] that we used for analyzing our survey. To quantitatively assess how VX impacts NASA TLX perceived workload ratings, we performed the nonparametric aligned rank transform (ART) procedure [23, 34, 85]; the 1-21 scales were for *mental demand, physical demand, temporal demand, performance, effort,* and *frustration.*

6.4 Results

We present the findings from our quantitative, qualitative, and perceived workload assessment of our VoxLens enhancements.

6.4.1 Quantitative Assessment. VoxLens *Version* (VX) had a significant effect on all three ordinal outcomes, indicating significant differences between the original VoxLens and our enhanced version (see Table 5). Overall, there was a 43.8%, 16.7%, and 11.5% increase

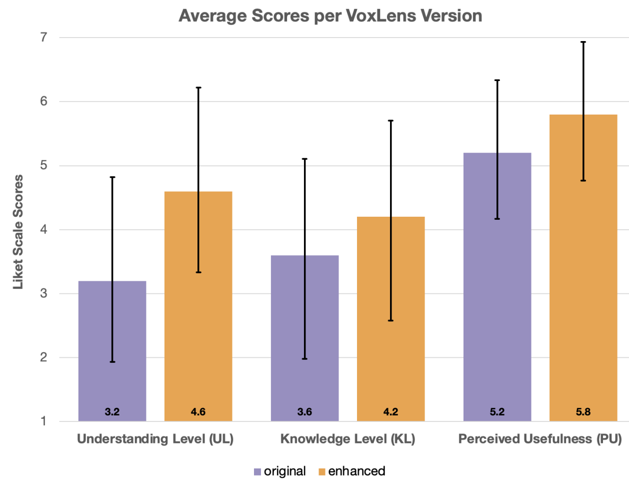

Figure 3: Visualization showing the Likert scale scores (with "1" being the lowest and "7" being the highest) for the original and our enhanced VoxLens version. Higher scores are better. Error bars represent ±1 standard deviation.

Table 5: Statistical results from our mixed ordinal logistic regression analysis from N=10 visualization creators with VoxLens *Version* (VX) as the independent variable ("original" *vs.* "enhanced"). "DV" means dependent variable. All results are statistically significant (p < .05).

DV	χ^2	p
Understanding Level (UL)	7.78	< .05
Knowledge Level (KL)	5.88	< .05
Perceived Usefulness (PU)	4.07	< .05

in *Understanding Level* (UL), *Knowledge Level* (KL), and *Perceived Usefulness* (PU), respectively, for our enhanced version compared to the original VoxLens (see Figure 3). These results confirm that the interventions and features revealed by our interviews provided measurable improvements to visualization creators' understanding, knowledge, and utility of visualization accessibility.

6.4.2 Qualitative Assessment. We assessed our enhancements' usefulness by analyzing the open-ended questionnaire responses from all participants. Specifically, we asked about their liked and disliked features and improvement areas.

Overall, in comparison to the original VoxLens, participants found our enhanced version to be *valuable for researchers in the field of data visualization; more user-friendly* (V10), *a great inclusion for developers; makes it easier to create accessible visualizations* (V3), and *a good step forward to helping with inclusive data visualization development* (V6). Additionally, participants expressed their excitement by stating *[the enhanced version] provided better accessibility* (V2) and *[is] really useful, [I] can't get over it!* (V5).

Our participants also emphasized the usefulness of the emulator feature by classifying it as a good *sanity check that testing was working as expected; an entry point for novice creators* (V1), *definitely enhanced the experience* (V3), *much more convenient than opening and closing a Narrator* (V4), and *ease of testing screen reader use without having a screen reader* (V5). Additionally, they found the evaluator feature *handy* (V1) and *a good addition that could save developers time versus having separate tools to do that* (V9). They also shared their liking for the other features, including customization options to toggle these features through configuration preferences.

We also asked our participants to share areas for improvement. They suggested having more working examples and customization settings for the emulator feature. For example, our participants wished to have the capability to increase and decrease the response speed of the emulator. At the time of this writing, we have started the process of implementing this suggestion to improve our enhanced `VoxLens` version.

6.4.3 Workload Assessment. We collected perceived workload ratings for both `VoxLens` versions (VX) using the NASA Task Load Index survey (NASA-TLX) [31], recording participants' mental demand, physical demand, temporal demand, performance, effort, and frustration. Our analysis used the nonparametric aligned rank transform procedure [23, 34, 85], but did not reveal a significant effect of VX on any of the six workload scales, indicating similar workload levels for each version of `VoxLens`. (See Appendix C for the workload scores.) This result indicates that the additional features and tools in the enhanced version of `VoxLens` did not detectably increase the perceived workload of the original `VoxLens`.

7 DISCUSSION

In this work, we sought to understand and reduce the challenges visualization creators experience in making data visualizations accessible to screen-reader users. We surveyed 57 visualization creators to comprehend their visualization accessibility challenges, identifying one educational and four technological interventions that can, when used together, assist them in enhancing their understanding of the challenges screen-reader users experience with data visualizations, improving their accessibility knowledge, and making data visualizations accessible. We conducted interviews with 12 visualization creators to determine the effectiveness of these interventions. Additionally, we implemented the four technological interventions into an existing open-source library, `VoxLens`, and conducted a task-based user study with 10 participants to validate the generalizability of our findings.

7.1 Uniqueness of Each Intervention

Our participants identified and discussed the effectiveness of five interventions: *Workshops* (WS), *Emulators* (EM), *Evaluators* (EV), *Feedback Collectors* (FC), and *Multi-Modal Automated Tools* (MM). Assessing against our three objectives of enhancing visualization creators' (1) understanding of screen-reader users' challenges with data visualizations, (2) knowledge of data visualization accessibility, and (3) ease of making data visualizations accessible, our participants determined that each intervention, when used separately, was beneficial in fulfilling the three objectives only to a certain degree. For example, *MM* would drastically improve the ease of making visualizations accessible but would be less effective in understanding the challenges screen-reader users experience with data visualizations. In contrast, these benefits and limitations would be vice versa for *EM*. (For clarity, in this work, we do not explore if these interventions build empathy for screen-reader users, a notion criticized by Bennett *et al.* [5]. Instead, our work examines the utility of these interventions to understand the *technology*.) Therefore, to achieve all three objectives, we recommend future work to utilize these interventions conjointly, leveraging the unique benefits they offer visualization creators and screen-reader users.

7.2 The Paradox of a "Magic-Wand" Solution

Our participants referred to the *Multi-Modal Automated Tools* (MM) as a "magic-wand" solution (S2, S3, S7). Although they communicated volumes about the usefulness of *MM*, some participants also identified how *MM* could restrict visualization creators from fully understanding the challenges of screen-reader users and enhancing their knowledge of accessibility. While S7 provided clarity on the matter by acknowledging the importance of gaining accessibility knowledge (*it'd be great if people knew about accessibility*) and sharing her opinions on the outcome (*the outcome of accessibility should be people who use screen readers are able to do stuff*), we encourage researchers and practitioners to carry forward this critical discussion, taking into account the perspectives of both visualization creators and screen-reader users. (We note that all authors remained neutral on this subject when communicating with our participants and remain similarly so in this paper.)

7.3 User Agency for Visualization Creators

A prominent observation from our interviews was the individualized usage needs and preferences of visualization creators in determining the effectiveness of each intervention. For example, S3 conveyed an essential component for the evaluators to *turn off rules* so people wouldn't get *annoyed* and, in turn, *wouldn't wanna use it anymore*. Prior work on data visualization accessibility has identified the need for customization and personalization options for screen-reader users [69, 70, 90]. Based on our findings, we recognize that the same is true for visualization creators, especially given that our participants appreciated the customization options to toggle the features in our enhanced version of `VoxLens`. Therefore, we recommend future work to investigate and provide more customization and personalization options in tools created to support visualization creators in making data visualizations accessible. Future work can further examine the differences in these options based on granular factors, including their domain (e.g., industry, academia), role (e.g., developer, researcher), or experience with data visualization accessibility.

7.4 Limitations and Future Work

Our survey findings revealed a significant correlation between some variables used in our study, such as visualization creators' knowledge of data visualization accessibility and their prioritization of making data visualizations accessible to screen-reader users. Our data was observational and not experimental; therefore, we did not explore causal relationships among the variables we analyzed for correlations. We encourage future work to investigate these correlations further, particularly via experiments to determine possible causal relationships. Furthermore, we did not implement our educational intervention (*Workshops*) due to logistical constraints. Utilizing the findings from our studies, we intend to organize a synchronous in-person workshop in the future with screen-reader users and practitioners as co-leads to assess the usefulness of *Workshops* for visualization creators. Additionally, future work can conduct studies with screen-reader users to assess further competencies of visualization creators.

8 CONCLUSION

In this work, we sought to understand and improve visualization creators' (1) understanding of screen-reader users' challenges with data visualizations, (2) knowledge of visualization accessibility, and (3) ease of making data visualizations accessible to screen-reader users. To this end, we surveyed visualization creators to determine and identify interventions to minimize their challenges with visualization accessibility. These interventions were: *Workshops*, *Emulators*, *Evaluators*, *Feedback Collectors*, and *Multi-Modal Automated Tools*. Then, we conducted semi-structured interviews to examine the effective versions of each of the five interventions and reported the specific features that make these interventions effective. To validate our findings, we implemented and integrated the technological interventions into `VoxLens` [69] and performed a task-based user study with visualization creators to assess the generalizability of our findings and the usefulness of our open-source enhancements. Using our enhancements, visualization creators improved their understanding of screen-reader users' challenges with data visualizations, knowledge of visualization accessibility, and perceived usefulness of `VoxLens`. Overall, we offer the effective versions of the interventions we used to achieve our three objectives as generalizable knowledge for researchers and practitioners, particularly in the data visualization community. We hope our work will inspire future efforts in further supporting visualization creators to make data visualizations accessible to screen-reader users.

ACKNOWLEDGMENTS

This work was supported in part by the University of Washington Center for Research and Education on Accessible Technology and Experiences (CREATE). We thank `EmailJS` for offering to utilize their email sending platform for our feedback collector intervention. Lastly, we extend our warmest gratitude to Harry and Gemma for their *purr*vasive feline support and cuddles.

REFERENCES

[1] NV Access. 2006. NV Access | Download. https://www.nvaccess.org/download/. (Accessed on 08/08/2021).

[2] Dragan Ahmetovic, Niccolò Cantù, Cristian Bernareggi, João Guerreiro, Sergio Mascetti, and Anna Capietto. 2019. Multimodal Exploration of Mathematical Function Graphs with AudioFunctions.Web. In *Proceedings of the 16th International Web for All Conference* (San Francisco, CA, USA) *(W4A '19)*. Association for Computing Machinery, New York, NY, USA, Article 8, 2 pages. https://doi.org/10.1145/3315002.3332438

[3] Ruben Alcaraz Martinez, Mireia Ribera Turró, and Toni Granollers Saltiveri. 2022. Methodology for heuristic evaluation of the accessibility of statistical charts for people with low vision and color vision deficiency. *Universal access in the information society* 21, 4 (2022), 863–894.

[4] David Austin and Volker Sorge. 2023. Authoring Web-Accessible Mathematical Diagrams. In *Proceedings of the 20th International Web for All Conference* (Austin, TX, USA) *(W4A '23)*. Association for Computing Machinery, New York, NY, USA, 148–152. https://doi.org/10.1145/3587281.3587297

[5] Cynthia L Bennett and Daniela K Rosner. 2019. The promise of empathy: Design, disability, and knowing the" other". In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.

[6] Jeffrey P Bigham, Craig M Prince, and Richard E Ladner. 2008. WebAnywhere: a screen reader on-the-go. In *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*. 73–82.

[7] Matthew Blanco, Jonathan Zong, and Arvind Satyanarayan. 2022. Olli: An extensible visualization library for screen reader accessibility. *IEEE VIS Posters* 6 (2022).

[8] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2301–2309.

[9] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development*. sage.

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

[11] 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

[12] Ben Caldwell, Loretta Guarino Reid, Gregg Vanderheiden, Wendy Chisholm, John Slatin, and Jason White. 2018. Web content accessibility guidelines (WCAG) 2.1.

[13] Dustin Carroll, Suranjan Chakraborty, and Jonathan Lazar. 2013. Designing accessible visualizations: The case of designing a weather map for blind users. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, Springer Berlin Heidelberg, Berlin, Heidelberg, 436–445.

[14] Johny Cassidy. 2023. Blind news audiences are being left behind in the data visualisation revolution: here's how we fix that | Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/blind-news-audiences-are-being-left-behind-data-visualisation-revolution-heres-how-we-fix. (Accessed on 08/21/2023).

[15] ChartJS. 2015. Accessibility | Chart.js. https://www.chartjs.org/docs/3.5.1/general/accessibility.html. (Accessed on 01/08/2022).

[16] Joel J Davis. 2002. Disenfranchising the Disabled: The Inaccessibility of Internet-Based Health Information. *Journal of Health Communication* 7, 4 (2002), 355–367. https://doi.org/10.1080/10810730290001701

[17] Descript. 2017. Descript | All-in-one video & podcast editing, easy as a doc. https://www.descript.com/. (Accessed on 08/09/2023).

[18] Google Developers. 2014. Charts. https://developers.google.com/chart/

[19] Norman R Draper and Harry Smith. 1981. Applied Regression Analysis, John Wiley and Sons. *New York* 407 (1981).

[20] Steve Duenes. 2022. Jaime Tanner Joins as Accessibility Visuals Editor | The New York Times Company. https://www.nytco.com/press/jaime-tanner-joins-as-accessibility-visuals-editor/. (Accessed on 08/21/2023).

[21] Michael Alin Efroymson. 1960. Multiple regression analysis. *Mathematical methods for digital computers* (1960), 191–203.

[22] Frank Elavsky, Cynthia Bennett, and Dominik Moritz. 2022. How accessible is my visualization? Evaluating visualization accessibility with Chartability. *Computer Graphics Forum* 41, 3 (2022), 57–70. https://doi.org/10.1111/cgf.14522

[23] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 754–768. https://doi.org/10.1145/3472749.3474784

[24] Niklas Elmqvist. 2023. Visualization for the Blind. *Interactions* 30, 1 (2023), 52–56.

[25] Danyang Fan, Alexa Fay Siu, Hrishikesh Rao, Gene Sung-Ho Kim, Xavier Vazquez, Lucy Greco, Sile O'Modhrain, and Sean Follmer. 2023. The accessibility of data visualizations on the web for screen reader users: Practices and experiences during covid-19. *ACM Transactions on Accessible Computing* 16, 1 (2023), 1–29.

[26] Danyang Fan, Alexa Fay Siu, Wing-Sum Adrienne Law, Raymond Ruihong Zhen, Sile O'Modhrain, and Sean Follmer. 2022. Slide-Tone and Tilt-Tone: 1-DOF Haptic Techniques for Conveying Shape Characteristics of Graphs to Blind Users. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 477, 19 pages. https://doi.org/10.1145/3491102.3517790

[27] Deen G. Freelon. 2010. ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science* 5, 1 (2010), 20–33.

[28] Matthew Graham, Anthony Milanowski, and Jackson Miller. 2012. Measuring and Promoting Inter-Rater Agreement of Teacher and Principal Performance Ratings. *Online Submission* (2012).

[29] Charu Gupta, Amita Jain, and Nisheeth Joshi. 2018. Fuzzy logic in natural language processing–a closer view. *Procedia computer science* 132 (2018), 1375–1384.

[30] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

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

[32] Donald P Hartmann. 1977. Considerations in the choice of interobserver reliability estimates. *Journal of applied behavior analysis* 10, 1 (1977), 103–116.

[33] Martin Hellmann. 2001. Fuzzy logic introduction. *Université de Rennes* 1, 1 (2001).

[34] James J Higgins and Suleiman Tashtoush. 1994. An aligned rank transform test for interaction. *Nonlinear World* 1, 2 (1994), 201–211.

[35] Ronald R Hocking. 1976. A Biometrics invited paper. The analysis and selection of variables in linear regression. *Biometrics* (1976), 1–49.

[36] Leona M Holloway, Cagatay Goncu, Alon Ilsar, Matthew Butler, and Kim Marriott. 2022. Infosonics: Accessible Infographics for People Who Are Blind Using Sonification and Voice. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 480, 13 pages. https://doi.org/10.1145/3491102.3517465

[37] Amy Hurst. 2018. Making "Making" Accessible. In *Proceedings of the 15th International Web for All Conference* (Lyon, France) *(W4A '18)*. Association for Computing Machinery, New York, NY, USA, Article 1, 1 pages. https://doi.org/10.1145/3192714.3192715

[38] Apple Inc. 2005. Accessibility - Vision - Apple. https://www.apple.com/accessibility/vision/.

[39] Shakila Cherise S Joyner, Amalia Riegelhuth, Kathleen Garrity, Yea-Seul Kim, and Nam Wook Kim. 2022. Visualization Accessibility in the Wild: Challenges Faced by Visualization Designers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 83, 19 pages. https://doi.org/10.1145/3491102.3517630

[40] Chloe Keilers, Garreth W Tigwell, and Roshan L Peiris. 2023. Data Visualization Accessibility for Blind and Low Vision Audiences. In *International Conference on Human-Computer Interaction*. Springer, 399–413.

[41] Edward Kim and Kathleen F McCoy. 2018. Multimodal deep learning using images and text for information graphic classification. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 143–148.

[42] Edward Kim, Connor Onweller, and Kathleen F McCoy. 2021. Information Graphic Summarization using a Collection of Multimodal Deep Neural Networks. In *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, Institute of Electrical and Electronics Engineers, New York, NY, USA, 10188–10195.

[43] N. W. Kim, S. C. Joyner, A. Riegelhuth, and Y. Kim. 2021. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum* 40, 3 (2021), 173–188. https://doi.org/10.1111/cgf.14298 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14298

[44] Mario Konecki, Charles LaPierre, and Keith Jervis. 2018. Accessible data visualization in higher education. In *2018 41st international convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, Institute of Electrical and Electronics Engineers, New York, NY, USA, 0733–0737.

[45] Klaus Krippendorff. 2011. Computing Krippendorff's Alpha-Reliability. https://repository.upenn.edu/asc_papers/43 Retrieved from.

[46] Klaus Krippendorff. 2018. *Content analysis: An introduction to its methodology*. SAGE Publications Inc., Pennsylvania, USA.

[47] Richard J. Landis and Gary G. Kock. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174. http://www.jstor.org/stable/2529310

[48] Bongshin Lee, Eun Kyoung Choe, Petra Isenberg, Kim Marriott, and John Stasko. 2020. Reaching broader audiences with data visualization. *IEEE Computer Graphics and Applications* 40, 2 (2020), 82–90.

[49] Bongshin Lee, Arjun Srinivasan, Petra Isenberg, John Stasko, et al. 2021. Post-WIMP Interaction for Information Visualization. *Foundations and Trends® in Human-Computer Interaction* 14, 1 (2021), 1–95.

[50] Alan Lundgard, Crystal Lee, and Arvind Satyanarayan. 2019. Sociotechnical considerations for accessible visualization design. In *2019 IEEE Visualization Conference (VIS)*. IEEE, Institute of Electrical and Electronics Engineers, New York, NY, USA, 16–20.

[51] Alan Lundgard and Arvind Satyanarayan. 2021. Accessible visualization via natural language descriptions: A four-level model of semantic content. *IEEE transactions on visualization and computer graphics* 28, 1 (2021), 1073–1083.

[52] Kim Marriott, Bongshin Lee, Matthew Butler, Ed Cutrell, Kirsten Ellis, Cagatay Goncu, Marti Hearst, Kathleen McCoy, and Danielle Albers Szafir. 2021. Inclusive data visualization for people with disabilities: a call to action. *Interactions* 28, 3 (2021), 47–51.

[53] Peter McCullagh. 1980. Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)* 42, 2 (1980), 109–127.

[54] David K. McGookin and Stephen A. Brewster. 2006. SoundBar: Exploiting Multiple Views in Multimodal Graph Browsing. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles* (Oslo, Norway) *(NordiCHI '06)*. Association for Computing Machinery, New York, NY, USA, 145–154. https://doi.org/10.1145/1182475.1182491

[55] Richard D McKelvey and William Zavoina. 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology* 4, 1 (1975), 103–120.

[56] Silvia Mirri, Silvio Peroni, Paola Salomoni, Fabio Vitali, and Vincenzo Rubano. 2017. Towards accessible graphs in HTML-based scientific articles. In *2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)*. IEEE, Las Vegas, NV, USA, 1067–1072. https://doi.org/10.1109/CCNC.2017.7983287

[57] Manuel M Oliveira. 2013. Towards more accessible visualizations for color-vision-deficient individuals. *Computing in Science & Engineering* 15, 5 (2013), 80–87.

[58] Michael Quinn Patton. 1990. *Qualitative evaluation and research methods*. SAGE Publications, inc.

[59] Catherine Plaisant. 2004. The Challenge of Information Visualization Evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (Gallipoli, Italy) *(AVI '04)*. Association for Computing Machinery, New York, NY, USA, 109–116. https://doi.org/10.1145/989863.989880

[60] Jonathan Potter and Margaret Wetherell. 1987. *Discourse and social psychology: Beyond attitudes and behaviour*. Sage.

[61] Sarbeswar Praharaj, Patricia Solis, and Elizabeth A Wentz. 2023. Deploying geospatial visualization dashboards to combat the socioeconomic impacts of COVID-19. *Environment and Planning B: Urban Analytics and City Science* 50, 5 (2023), 1262–1279.

[62] Tanumon Roy and Lakshmi Boppana. 2022. Interactive web-based image and graph analysis using Sonification for the Blind. In *2022 IEEE Region 10 Symposium (TENSYMP)*. 1–6. https://doi.org/10.1109/TENSYMP54529.2022.9864411

[63] Gery W Ryan, H Russell Bernard, et al. 2000. Data management and analysis methods. *Handbook of qualitative research* 2, 1 (2000), 769–802.

[64] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2016. Vega-lite: A grammar of interactive graphics. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 341–350.

[65] Freedom Scientific. n.d.. JAWS® – Freedom Scientific. https://www.freedomscientific.com/products/software/jaws/. (Accessed on 08/08/2021).

[66] Ather Sharif, Sanjana Shivani Chintalapati, Jacob O. Wobbrock, and Katharina Reinecke. 2021. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 14, 16 pages. https://doi.org/10.1145/3441852.3471202

[67] Ather Sharif and Babak Forouraghi. 2018. evoGraphs — A jQuery plugin to create web accessible graphs. In *2018 15th IEEE Annual Consumer Communications Networking Conference (CCNC)*. IEEE, Las Vegas, NV, USA, 1–4. https://doi.org/10.1109/CCNC.2018.8319239

[68] Ather Sharif, Olivia H. Wang, and Alida T. Muongchan. 2022. "What Makes Sonification User-Friendly?" Exploring Usability and User-Friendliness of Sonified Responses. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 45, 5 pages. https://doi.org/10.1145/3517428.3550360

[69] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 478, 19 pages. https://doi.org/10.1145/3491102.3517431

[70] Ather Sharif, Andrew M. Zhang, Katharina Reinecke, and Jacob O. Wobbrock. 2023. Understanding and Improving Drilled-Down Information Extraction from Online Data Visualizations for Screen-Reader Users. In *Proceedings of the 20th International Web for All Conference* (Austin, TX, USA) *(W4A '23)*. Association for Computing Machinery, New York, NY, USA, 18–31. https://doi.org/10.1145/3587281.3587284

[71] Ather Sharif, Ruican Zhong, and Yadi Wang. 2023. Conveying Uncertainty in Data Visualizations to Screen-Reader Users Through Non-Visual Means. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, To Appear pages. https://doi.org/10.1145/3597638.3614502

[72] Lei Shi, Idan Zelzer, Catherine Feng, and Shiri Azenkot. 2016. Tickers and Talker: An Accessible Labeling Toolkit for 3D Printed Models. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 4896–4907. https://doi.org/10.1145/2858036.2858507

[73] Mandhatya Singh, Muhammad Suhaib Kanroo, Hadia Showkat Kawoosa, and Puneet Goyal. 2023. Towards accessible chart visualizations for the non-visuals: Research, applications and gaps. *Computer Science Review* 48 (2023), 100555.

[74] Alexa Siu, Gene S-H Kim, Sile O'Modhrain, and Sean Follmer. 2022. Supporting Accessible Data Visualization Through Audio Data Narratives. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 476, 19 pages. https://doi.org/10.1145/3491102.3517678

[75] C. Spearman. 1904. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology* 15, 1 (1904), 72–101. http://www.jstor.org/stable/1412159

[76] Katta Spiel, Oliver L Haimson, and Danielle Lottridge. 2019. How to do better with gender on surveys: a guide for HCI researchers. *Interactions* 26, 4 (2019), 62–65.

[77] Adam Strantz. 2021. Using Web Standards to Design Accessible Data Visualizations in Professional Communication. *IEEE Transactions on Professional Communication* 64, 3 (2021), 288–301.

[78] Hironobu Takagi, Chieko Asakawa, Kentarou Fukuda, and Junji Maeda. 2003. Accessibility Designer: Visualizing Usability for the Blind. In *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility* (Atlanta, GA, USA) *(Assets '04)*. Association for Computing Machinery, New York, NY, USA, 177–184. https://doi.org/10.1145/1028630.1028662

[79] John R Thompson, Jesse J Martinez, Alper Sarikaya, Edward Cutrell, and Bongshin Lee. 2023. Chart Reader: Accessible Visualization Experiences Designed with Screen Reader Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 802, 18 pages. https://doi.org/10.1145/3544548.3581186

[80] Shari Trewin, Brian Cragun, Cal Swart, Jonathan Brezin, and John Richards. 2010. Accessibility Challenges and Tool Features: An IBM Web Developer Perspective. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)* (Raleigh, North Carolina) *(W4A '10)*. Association for Computing Machinery, New York, NY, USA, Article 32, 10 pages. https://doi.org/10.1145/1805986.1806029

[81] Frances Van Scoy, Don McLaughlin, and Angela Fullmer. 2005. Auditory augmentation of haptic graphs: Developing a graphic tool for teaching precalculus skill to blind students.

[82] WebAIM. [n. d.]. WebAIM: Contrast Checker. https://webaim.org/resources/contrastchecker/. (Accessed on 07/24/2024).

[83] WebAIM. 2016. WAVE Web Accessibility Evaluation Tools. https://wave.webaim.org/. (Accessed on 08/21/2023).

[84] Sue Widdicombe and Robin Wooffitt. 1995. *The language of youth subcultures: Social identity in action.* Harvester/Wheatsheaf.

[85] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 143–146.

[86] Keke Wu, Emma Petersen, Tahmina Ahmad, David Burlinson, Shea Tanis, and Danielle Albers Szafir. 2021. Understanding Data Accessibility for People with Intellectual and Developmental Disabilities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 606, 16 pages. https://doi.org/10.1145/3411764.3445743

[87] Wai Yu, Rameshsharma Ramloll, and Stephen A. Brewster. 2000. Haptic Graphs for Blind Computer Users. In *Proceedings of the First International Workshop on Haptic Human-Computer Interaction*. Springer-Verlag, Berlin, Heidelberg, 41–51.

[88] Lotfi A Zadeh. 1988. Fuzzy logic. *Computer* 21, 4 (1988), 83–93.

[89] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Jonathan Lazar. 2008. Data sonification for users with visual impairment: a case study with georeferenced data. *ACM Transactions on Computer-Human Interaction (TOCHI)* 15, 1 (2008), 1–28.

[90] Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. 2022. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum* 41, 3 (2022), 15–27. https://doi.org/10.1111/cgf.14519 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14519

A SPEARMAN'S RANK CORRELATION RESULTS

Table 6: Statistical results from our Spearman's rank correlation analysis of all our variables. "V1" means Variable 1 and "V2" means Variable 2. ρ is the Spearman's rank correlation coefficient. All results with p <.05 are statistically significant.

V1	V2	ρ	p
CR	*KR*	-.05	.719
CR	*IR*	.00	.980
CR	*PO*	-.05	.725
CR	*SF*	.09	.528
CR	*PR*	-.08	.561
CR	*DO*	.16	.247
CR	*RO*	-.08	.541
KR	*IR*	.20	.145
KR	*PR*	.42	< .001
KR	*PO*	-.39	< .05
KR	*SF*	.29	< .05
KR	*DO*	.23	.088
KR	*RO*	-.08	.548
IR	*PR*	.49	< .001
IR	*PO*	-.24	.068
IR	*SF*	.21	.119
IR	*DO*	.66	.659
IR	*RO*	.09	.499
PR	*PO*	-.43	< .001
PR	*SF*	.12	.373
PR	*DO*	.14	.304
PR	*RO*	-.06	.655
PO	*SF*	-.34	< .05
PO	*DO*	-.07	.605
PO	*RO*	-.03	.803
SF	*DO*	-.15	.257
SF	*RO*	-.01	.962
DO	*RO*	-.25	< .05

B VOXLENS CONFIGURATIONS OPTIONS

```
// original voxlens config

const voxlensOptions = {
  x: 'state',
  y: 'cases',
  title: 'COVID-19 Cases per US State',
  chartType: 'bar',
};

// enhanced version config with feedbackCollector and all debug options

const voxlensOptions = {
  x: 'state',
  y: 'cases',
  title: 'COVID-19 Cases per US State',
  chartType: 'bar',
  feedbackCollector: {
    scales: 5,
    email: 'visualization_creator@example.com'
  },
  debug: true,
};

// enhanced version config with granular debug options

const voxlensOptions = {
  x: 'state',
  y: 'cases',
  title: 'COVID-19 Cases per US State',
  chartType: 'bar',
  feedbackCollector: {
    scales: 5,
    email: 'visualization_creator@example.com'
  },
  debug: {
    instructions: {
      onlyMain: false,
    },
    hideFeedbackCollector: false,
    contrastChecker: true,
    response: {
      onlyText: false,
    },
  },
};
```

Figure 4: VoxLens configuration for the original version, our enhanced version with default debug options, and our enhanced version with granular debug options.

C WORKLOAD ASSESSMENT ANALYSIS RESULTS

Table 7: Statistical results from our nonparametric aligned rank transformation procedure used in the workload assessment of our task-based user study findings. "VO" means Original Version and "VE" means Enhanced Version. Scores are mean values. No results are statistically significant (p <.05).

Scale	VO Score	VE Score	p
Mental Demand	3.3	3.7	.552
Physical Demand	1.4	2.3	.188
Temporal Demand	2.4	3.5	.104
Performance	3.6	3.4	.824
Effort	3.4	3.3	.739
Frustration	2.8	2.7	1.000