Journal of the American Medical Informatics Association, 29(11), 2022, 1847–1858
https://doi.org/10.1093/jamia/ocac140
Advance Access Publication Date: 17 August 2022
Research and Applications

Research and Applications

Development of a usability checklist for public health dashboards to identify violations of usability principles

Bahareh Ansari [1,2] and Erika G. Martin [1,2,3]

[1]Center for Policy Research, Rockefeller College of Public Affairs and Policy, University at Albany, Albany, New York, USA, [2]Center for Collaborative HIV Research in Practice and Policy, School of Public Health, University at Albany, Albany, New York, USA, and [3]Department of Public Administration and Policy, Rockefeller College of Public Affairs and Policy, University at Albany, Albany, New York, USA

Corresponding Author: Bahareh Ansari, PhD, Center for Policy Research, Rockefeller College of Public Affairs and Policy, University at Albany, 300 Milne Hall, 135 Western Ave, Albany, NY 12203, USA; bansari@albany.edu

Received 17 May 2022; Revised 31 July 2022; Editorial Decision 1 August 2022; Accepted 12 August 2022

ABSTRACT

Objective: To develop a usability checklist for public health dashboards.
Materials and methods: This study systematically evaluated all publicly available dashboards for sexually transmitted infections on state health department websites in the United States ($N = 13$). A set of 11 principles derived from the information visualization literature were used to identify usability problems that violate critical usability principles: spatial organization, information coding, consistency, removal of extraneous ink, recognition rather than recall, minimal action, dataset reduction, flexibility to user experience, understandability of contents, scientific integrity, and readability. Three user groups were considered for public health dashboards: public health practitioners, academic researchers, and the general public. Six reviewers with usability knowledge and diverse domain expertise examined the dashboards using a rubric based on the 11 principles. Data analysis included quantitative analysis of experts' usability scores and qualitative synthesis of their textual comments.
Results: The dashboards had varying levels of complexity, and the usability scores were dependent on the dashboards' complexity. Overall, understandability of contents, flexibility, and scientific integrity were the areas with the most major usability problems. The usability problems informed a checklist to improve performance in the 11 areas.
Discussion: The varying complexity of the dashboards suggests a diversity of target audiences. However, the identified usability problems suggest that dashboards' effectiveness for different groups of users was limited.
Conclusions: The usability of public health data dashboards can be improved to accommodate different user groups. This checklist can guide the development of future public health dashboards to engage diverse audiences.

Key words: user-centered design, public health informatics, data visualization

INTRODUCTION

Increased data availability following open government data movements and enhanced visualization technology[1] have led to the proliferation of public health data dashboards. Data dashboards employ data visualization technology to provide information for non-technical users, encourage evidence-based practices, and expand the use of public health surveillance data beyond traditional audiences.[2] However, current practices for displaying information to diverse audiences have not been adopted fully.[3–5]

Previous evaluations of public health data visualizations found severe usability problems that could limit their usefulness for their target audiences. A review of several COVID-19 dashboards for people with low vision found that charts had incompatibility problems with keyboard interfaces, poor text alternatives, and insufficient color contrast.[6] An evaluation of the usability of coronavirus disease (COVID-19) contact tracing applications found that most applications were not usable for users who do not speak English or have limited digital literacy skills.[7]

Although the gold standard usability evaluation is user testing,[8] evaluating the usability of public health dashboards with end-users is challenging because of the diversity of users and use cases.[4,9] Therefore, usability checklists and principles can be useful to guide their development. For example, academic researchers might use dashboards for exploratory data analysis, while public health practitioners might use dashboards for routine and sometimes urgent tasks.[10] For common infectious diseases, such as influenza, audiences might expand to government officials, schools, hospital administrators, long-term care facilities, and the general public.[11] The wide audience of public health dashboards complicates identifying a representative sample of users and conducting user testing. Multiple studies found comparable usability problems identified by users or a group of evaluators using usability principles,[12–15] suggesting that usability principles can provide an inexpensive substitute for traditional user testing.

A recent study published a usability checklist for health information visualization; however, the checklist was tailored to clinical data dashboards.[16] Many usability principles for clinical data dashboards are transferrable to public health dashboards, such as employing best practices for creating charts, using familiar terminology, and organizing information logically.[16] However, public health dashboards provide population-level data and have more diverse use cases than clinical data dashboards, which require special considerations. For example, while clinical data dashboards may be used for diagnosis and care management of individual patients, public health dashboards have broader uses such as policy-making, education, or empowering communities.[2] Therefore, special usability considerations are required on public health dashboards to help diverse users navigate the dashboard and find relevant information for their specific use.

This study developed a usability checklist for public health dashboards via 2 specific aims: (1) to evaluate dashboards to find common violations of usability principles (hereafter, "usability problems" for consistency with the literature) using a set of principles from information visualization literature; and (2) to use evaluators' qualitative comments to develop a usability checklist for future dashboards. To compare different information displays on a similar topic in detail, this study focused on public health dashboards for sexually transmitted infections (STIs). STIs are an important public health issue in the United States, with rising infection rates and substantial healthcare-related costs.[17,18] Current gaps in provider communication about sexual health[19] make it particularly useful to have this information in a digestible format. The usability principles were chosen from the information visualization literature to avoid domain-specific principles and provide a checklist that could be generalizable to other diseases.

Three dashboard user groups were considered: public health practitioners, academic researchers, and the general public, each of whom include both domain experts and non-domain experts. Based on the authors' experience developing a state dashboard for STIs[20] and studying the development of open health data platforms,[21,22] these groups are frequently considered by developers of public health data dashboards. The evaluation rubric included items that addressed the perspectives of domain experts (eg, appropriate measures and sufficient granularity), and non-domain experts (eg, explanation of technical terms and guidance for interpretation). Furthermore, the evaluation rubric included items for persons with color vision deficiency, mobile users, and those with low-speed internet. This study did not evaluate the usability of the dashboards for persons with low vision, non-English language speakers, or people with limited digital literacy skills because prior research assessed the accessibility of public health dashboards and provided guidelines for considering accessibility for persons with disabilities.[6,7,23]

MATERIALS AND METHODS

Overview

First, a rubric was prepared by the first author (BA) and pilot-tested and finalized by both authors (BA and EGM) based on an existing set of principles in the information visualization literature. Second, 6 expert reviewers (BA, EGM, and 4 additional reviewers), all with knowledge of usability and information visualization and diverse domain expertise and public health experience, used the rubric to evaluate the STI dashboards. Third, their usability scores and textual comments were analyzed by the first author (BA) and reviewed by EGM to identify specific areas for improvement and prepare a general usability checklist for public health dashboards.

Sample

The sample comprised all available STI dashboards on the state health department websites ($N = 13$, as of June 1, 2021). To locate STI dashboards, the first author (BA) hand searched states' government websites and STI-related keywords (eg, "Massachusetts AND [sexually transmitted infections OR sexually transmitted diseases OR STD OR STI OR HIV OR AIDS]"). The hand search was done in May 2021, using the Google Chrome browser. All 50 states and the District of Columbia were included in the search. Websites were included if they had a data dashboard with STI statistics. The included states were: Arizona, Colorado, District of Columbia, Florida, Georgia, Hawaii, Iowa, Kansas, Maryland, New York, North Carolina, Tennessee, and Texas. Only state health department dashboards were considered because US state health departments receive data from local jurisdictions, allowing for more robust dashboards than single-county dashboards. State health departments typically have more expert staff than local health departments, with more capacity to develop complex data dashboards. Although variations exist, local health departments usually have limited staff, expertise, and IT infrastructure, making their websites less comparable to state departments for a systematic evaluation.[24,25]

Rubric

Table 1 displays the principles and considerations used in the rubric, which were adopted from the information visualization literature[26] with some revisions to this study's context. The information visualization literature offers different sets of principles.[26–31] This study started with Forsell and Johansson's heuristic set[26] because these principles incorporate those identified in past research. Forsett and Johansson's set comprises spatial organization, information coding, consistency, removal of extraneous ink, recognition rather than recall, prompting, dataset reduction, minimal action, flexibility, and help and documentation. In consultation with the second author (EGM), the first author (BA) made 3 revisions to these principles to adjust for this study's

Table 1. Usability evaluation rubric

Usability principles	Considerations
Spatial organization	• All information elements are clear and visible • The display follows a logical organization • There is a natural sequence to the menu choices for a data item • The section headings and subsections on each screen are ordered in a logical fashion • The system displays the reference context while showing details of specific data elements (eg, if a county is selected from a map, the statewide map goes to the background rather than disappearing)
Information coding	• It is easy to understand the mapping of data elements to visual elements • Labels and legends are clear • Size of elements is proportional to the data quantities they represent • Appropriate visual codes are used for different types of data (eg, color and shape for categorical variables, length, and position for continuous variables) • The connection made between visual elements (eg, with a line or with a similar shape or color) is appropriate for the data • Elements that are visually salient (eg, bold, large font, or very different color or orientation) are appropriately used to indicate higher importance
Consistency	• Data elements are always located in the same place • Similar screen formats and similar procedures to access menu options are used throughout the system • Similar data entry fields (eg, user input boxes or pull-down menus) are used and are displayed in a standard location • Color coding is consistent throughout the system • Consistent phrasing and punctuation are used throughout the system
Remove extraneous ink	• Labels are brief, familiar, and descriptive • Colors are used sparingly • Data are presented in a simple format • Unnecessary objects that could distract users' attention from relevant information (eg, background images or thick grid lines) are avoided
Recognition rather than recall	• Interaction-guiding messages are placed where the eye is likely to be looking on the screen (eg, a very small message in the bottom left corner of the screen is probably not visible and therefore not useful) • Visual cues are used to lead the eye in the appropriate direction (eg, appropriate sectioning with white space) • Information elements are grouped into logical zones, and headings are used to distinguish between zones • The system grays out the options that are currently inactive but may be used later • All data that users need is displayed at each step of a transition sequence • Legends and labels are available when using charts and maps
Minimal action	• The steps required to make a selection in menus are minimized • Users can easily move forward and backward between screens (eg, arrow buttons or icon for the home button) • Unnecessary data entry is avoided • For long, multi-page displays, it is possible to request a particular page directly, without having to go through all intermediary pages
Dataset reduction	• Users can filter the information to adjust the focus of interest (eg, selection of a time or a jurisdiction) • Users can adjust groups or clusters (eg, choice of different groups to be displayed with different colors) • User can cutoff irrelevant information from the visual representation (eg, accordion design to hide detailed information or tooltips to show details on demand)
Flexibility (consideration of user experience with similar websites)	• There is easy, step-by-step guidance available for inexperienced users that experienced users can skip • There is guidance available on how to read the charts • There are alternative modes of interaction available for experienced users (eg, download of data or change of data representation from a map to a chart to a table) • Different dialogue types are provided according to the experience of the various user groups (eg, annotation as an optional guidance feature that can be selected by inexperienced users but omitted by experienced users)

(continued)

Table 1. continued

Usability principles	Considerations
Scientific integrity (for domain experts)	• Data are available with enough granularity that can guide public health practice (eg, geographic distribution, time trend, distribution by age and sex, or by behavioral factors) • Data are available with enough granularity to examine health equity (eg, distribution by race and ethnicity and distribution by social determinants of health such as income) • A link is provided to access the underlying dataset, or an explanation is provided as to why the dataset cannot be publicly shared • An explanation is provided about data limitations (eg, lagging in data collection and reporting, missing or censored data) • Modeling choices are clearly explained (eg, use of log-scale, statistical models used for projection or handling uncertainties, and underlying assumptions)
Understandability of contents (for non-domain experts)	• The words, concepts, and phrases are familiar to lay users, and if jargon is used, it is accompanied by an explanation • Reference points are provided to help with assessing the provided measures (eg, it is clear what is normal, concerning, or crisis) • Easy explanations are provided for how to interpret the data • The scientific information is provided in a way that is available for domain experts but is not overwhelming for those without domain expertise (eg, a small footnote)
Readability	• The webpage loads in less than 3 s; following empirical research that 40% of visitors will not wait more than 3 s for a page to load • Persons with color vision deficiency can understand the contents • The webpage is mobile-friendly

Note: See Supplementary Appendix S1 for the instructions provided to evaluators, specific questions used in the rubric, and definitions of the usability principles.

intended user groups (public health practitioners, academic researchers, and the public, who include both domain experts and non-domain experts). First, help and documentation were separated into (1) scientific integrity (providing information about data sources and modeling assumptions for domain expert users) and (2) understandability (providing information about the context and interpretations for non-domain experts). Second, similar to a previous usability checklist developed for clinical dashboards,[16] each principle was placed into the context of public health data dashboards. The questions were not exhaustive, and evaluators were reminded that the questions' purpose was to guide their thinking rather than provide a checklist. Third, after completing the evaluation, the "prompting" principle was merged into "recognition rather than recall" because evaluators expressed confusion between these principles and the inapplicability of "prompting" in less interactive dashboards. Supplementary Appendix S1 shows the complete list of principles.

The scoring system, adapted from previous studies,[32] had 5 levels: (1) no usability problem, (2) cosmetic problem (ie, it need not be fixed unless extra time is available), (3) minor usability problem (ie, low priority usability problem), (4) major usability problem (ie, high priority usability problem), and (5) usability catastrophe (ie, critical problems to fix). In the analysis, the categories of cosmetic and minor problems, and major problems and catastrophes, were combined to facilitate the interpretation of the findings. Cosmetic problems and usability catastrophes were scored infrequently, and combining them with other levels did not impact findings.

Evaluators

The evaluators comprised 3 university faculty members (EGM, XY, and LL-R), a state health department employee (RH-M), and 2 doctoral students (BA and PL); see the acknowledgments for reviewer names. Three evaluators had STI and public health domain expertise (BA, EGM, and RH-M), of whom one works primarily in public health practice (RH-M) and a second has over a decade of experience leading collaborative academic research projects with public health practice partners (EGM). Reviewers' highest degrees were in public health (EGM and RH-M) and information science (BA, LL-R, PL, and XY). Three evaluators represented non-domain experts (PL, LL-R, and XY) and had other domain expertise, including digital government (LL-R) and human–computer interaction (PL and XY). An evaluator with direct clinical care experience was not included because this evaluation was focused on public health dashboards rather than clinical dashboards. Furthermore, all evaluators were knowledgeable in usability and data visualization. A layperson representing the general public was not included because the evaluation was an expert review requiring evaluators with usability knowledge.[33] The usability literature suggests that persons with different expertise would likely find different usability problems depending on their past experiences with similar products.[29,33] This group of experts with diverse domain expertise and practitioner versus researcher experience could evaluate the dashboards from different perspectives and provide a comprehensive usability checklist for varied audiences.

Data collection

The rubric was pilot-tested by both authors for understandability and ease of use. Each evaluator independently applied the finalized rubric to the dashboards with all reviews completed during July and August 2021. To ensure consistency in coding experiences and that diverging opinions were due to evaluators' unique perspectives and

not their method of viewing the dashboards, evaluators used the Chrome browser and a PC or laptop. To increase the generalizability of the identified usability problems, evaluators were asked to check whether the dashboard is mobile-friendly using their phones or Chrome's "toggle device" toolbar to simulate how web pages look on mobile devices.

Data analysis

Evaluation of STI dashboards: The intensity of usability problems, by principle, was quantified by calculating the percentage of major, minor, and no usability problems among all scores given by the 6 evaluators across the 13 dashboards. Consistent with the usability literature, inter-rater reliability analysis was not conducted because the coding exercise's purpose was to find diverse and numerous usability problems rather than reach a consensus between evaluators on a limited number of problems.[29,33]

Development of the usability checklist: Evaluators' qualitative comments about the 11 principles for the 13 dashboards were organized into an Excel spreadsheet, with each comment in a separate row and additional columns to denote the associated principle, dashboard, and whether the evaluator scored that principle as having a major, minor, or no usability problem. The first author (BA) filtered comments by principle and major versus minor usability problems to synthesize occasional or frequent comments related to different principles. The usability principles remained unchanged from the evaluation rubric, which was prepared based on the established literature and pilot testing. The final usability checklist was developed using the synthesized comments to refine specific considerations within each principle and to classify them as major versus minor usability problems.

RESULTS

Evaluation of STI dashboards

Table 2 displays dashboard characteristics. Dashboard names are suppressed to avoid comparison between state programs and remain respectful of the voluntary nature of these dashboards and competing time demands of public health agencies, following recommendations from several public health practitioners. Five dashboards were general data dashboards with information on different diseases, and 8 were specific to HIV/STI information. There were no major differences between the specific and generic dashboards regarding visualizations or interactive features. The organization of information varied, with some dashboards presenting all information on a single page divided into sections and others spreading information across multiple pages. The most common visualizations were thematic maps of counties for geographical comparisons, bar charts to compare males versus females or different racial and ethnic groups, line charts for annual trends, and tables. The visualizations were accompanied by filters and other interactive features such as hovering and show/hide effects. The complexity of data dashboards was determined by the available visualizations, filters, and other interactive features excluding filters. Four dashboards provide downloadable data.

Figure 1 displays the relative frequency of experts' scores. For each principle, experts identified major and minor usability problems. The 3 principles with the most frequent major usability problems were: flexibility of user experience (39.0% of the given scores were major usability problems), understandability of contents (38.5% major usability problems), and scientific integrity (35.1% major usability problems). The 3 principles with the fewest usability problems were: removing extraneous ink (57.9% of the given scores were no usability problem), minimal action (57.1% no usability problem), and consistency (54.5% no usability problem).

The evaluated dashboards in Table 2 are ranked based on the frequency of usability problems classified as major. Dashboards 1 through 5 were the most usable; however, they had limited complexity and were less suitable for public health practice because they either had very few visualizations or no demographic filtering for population comparisons. Dashboards 6 through 9 had fair usability and appropriate complexity that was achieved through multiple visualizations, filtering, or other features such as the hover effect. Dashboards 10 through 13 had appropriate complexity but the lowest usability. Unlike Dashboards 6 through 9, their complexity was achieved through the availability of multiple static visualizations or filters rather than interactive features.

Development of the usability checklist

Table 3 displays the usability checklist derived from synthesizing the evaluators' qualitative comments regarding identified usability problems on the examined dashboards. For example, under the consistency principle, consideration of consistent periods was based on an evaluator's comment, "I was expecting that the interactive dashboard would have the same information as the static ones, but interactive. The static, however, have information until 2018, and the newer information is in the interactive one, which is like having two dashboards instead of one." As another example, consideration of the consistent arrangement of data entry fields resulted from multiple comments from different evaluators: "Filters are located above the visuals, but impact each visual differently, if at all" and "There is no consistency in terms of how menus are structured or where they are placed." A selection of the evaluators' qualitative comments is available in Supplementary Appendix S2.

The checklist's 11 categories correspond to the 11 principles in the rubric. The checklist under each principle is separated into major and minor considerations. For example, major spatial organization considerations include unique formats for headings and titles, hierarchical relationships between headings and subheadings, consistency between headings and content, the spread of textual information and visualizations throughout the dashboard, and navigation between the views in dashboards with multiple views. Minor spatial organization considerations include avoiding long layouts, presenting similar visuals on the same side of the screen, clear positions for drop-down menus, checking visualizations for potential occlusion or blockage of information, and clarifying the navigation and the total number of pages in a multi-page dashboard. These considerations are presented in a question format to facilitate their use. The checklist is intended for finding and fixing potential usability problems with a public health dashboard. It is not intended for quantitative usability evaluation and comparison of dashboards because the included usability considerations originate from qualitative comments. A user-friendly version of the checklist with instructions is available in Supplementary Appendix S3.

DISCUSSION

This study systematically evaluated the usability of STI dashboards on US state health department websites. The evaluated dashboards had varying complexity, which impacted their usability scores. The most major usability problems were in 3 areas: understandability of

Table 2. Characteristics of 13 state-level data dashboards for sexually transmitted infections, ranked from the fewest to the most identified major usability problems

Usability rank (1 = fewest major problems)	Generic or STI/HIV specific[a]	Organization of information	Complexity			Downloadable data
			Available visualizations	Available filters	Interactive features (except filter)	
Dashboard 1	Specific	Multiple pages, each with a map for a specific STI	• Map of counties with a filter	• Year	Hover effect	No
Dashboard 2	Specific	Single page with one interactive chart	• Stacked bar chart with filters	• Sex or transmission mode or age • Sex or race and ethnicity or transmission mode or age	Hover effect	No
Dashboard 3	Generic	Single page with multiple tabs showing different visualizations of a selected disease	• Line chart (year) • Map (counties) • Bubble chart (disease burden) • Table	• Year • County • Disease • Measure	Hover effect	Yes
Dashboard 4	Generic	Single page with 2 sections including different visualizations for the selected disease	• Line chart (year) • Map and bar chart (counties) • Table	• County • Year • Disease	Show or hide	Yes
Dashboard 5	Generic	Single page with 2 sections including different visualizations for the selected disease	• Line chart (year) • Map and bar chart for counties • Table	• County • Year • Disease	Show or hide	Yes
Dashboard 6	Specific	Single page with an interactive section and multiple tabs with infographics on different STIs	• Line chart (year and month) • Pyramid chart (age and sex) • Heat map (race and ethnicity) • Infographics with maps, dot chart, and dumbbell chart	• County • Year • Disease	Show or hide, hover effect	No
Dashboard 7	Generic	Single page with multiple tabs (introduction, overview, table, map)	• Line chart (year) • Bar charts (sex, race and ethnicity, age) with filters #1 • Map (counties) • Table with filters #2	Filter #1: • Disease • Year Filter #2: • Disease • Year • Sex • Race and ethnicity • Age • County	Hover effect	No
Dashboard 8	Specific	Single page with different visualizations of HIV data	• Dumbbell chart with filters #1 • Stacked bar chart for engagement in HIV care with filter #2	Filters #1: • Sex or race and ethnicity or transmission mode or age • Sex or race and ethnicity or transmission mode or age Filters #2: • Region	Sort, hover effect	No

(continued)

Table 2. continued

Usability rank (1 = fewest major problems)	Generic or STI/HIV specific[a]	Organization of information	Complexity			Downloadable data
			Available visualizations	Available filters	Interactive features (except filter)	
Dashboard 9	Specific	Single page with 2 sections: a chart and a table for the selected county; a map and table for the state	• Line chart with filters #1 • Map (counties) with filters #2 • Table	Filters #1: • County • Year • Sex/race and ethnicity Filters #2: • Year • Sex/race and ethnicity	Hover effect	Yes
Dashboard 10	Specific	Single page with 3 sections including different visualizations of HIV, PrEP,[b] and services	• Line and area chart (year) • Map and bar chart (counties) • Bar chart (age) • Bar chart (race and ethnicity)	—	Show or hide, hover effect	No
Dashboard 11	Specific	Multiple pages for different HIV statistics (persons living with HIV or newly diagnosed)	• Line charts (race and ethnicity, gender, age, transmission mode) • Stacked bar chart (stages of disease) • Map (counties) with a filter	• Year	—	No
Dashboard 12	Specific	Multiple pages with different visualizations of STI data	• Line chart (year) • Pyramid chart (age and sex) • Map (counties) • Table	• Sex • Race • Ethnicity • Age • County • Year • Disease • Measure	—	Yes
Dashboard 13	Generic	Multiple pages for state or county data, each containing multiple tabs for map, table, and charts	• Map (counties) • Table (counties and region) • Bar chart (counties and region)	• Selected indicators (eg, disease rate per 100 000 males aged 15–44 years old)	—	Yes

[a]Generic dashboard refers to the ones that contain information about different diseases, while specific dashboards only contain information about STIs or HIV.
[b]PrEP = Pre-Exposure Prophylaxis, a medication to protect people at risk of HIV.

contents, the flexibility of user experience, and scientific integrity. The fewest usability problems were in 3 areas: removing extraneous ink, minimal action, and consistency. The usability problems informed a checklist to help future designers avoid common usability problems on public health dashboards.

The varying complexity of the evaluated dashboards suggests the dashboards might have been developed for a diversity of target audiences and strategic orientations. Most of the major usability problems were found in the 3 areas related to the diversity of targeted users (ie, understandability of contents for non-domain experts, scientific integrity for domain expert users, and the flexibility of user experience), which suggests that dashboards' effectiveness for multiple end-users was limited. Evaluators' scores and comments identified trade-offs between the usability dimensions. The complex dashboards containing many visualizations and interactivity received high scientific integrity scores because they presented information important for domain experts interested in specific populations but received lower scores for understandability for non-domain experts and readability for mobile users. In contrast, simple dashboards with few visualizations or interactivity received high understandability scores, but their low scientific integrity scores indicated their information is less useful for domain experts. These findings suggest that the evaluated dashboards have major usability problems for the diverse targeted audience of public health dashboards identified in previous studies.[3,10,11] To balance complexity and other usability dimensions, such as readability and understandability, future dashboards should be evaluated with different user groups, including both domain experts and non-domain experts.

Three groups of dashboard users were considered in this study: public health practitioners, academic researchers, and the general pub-

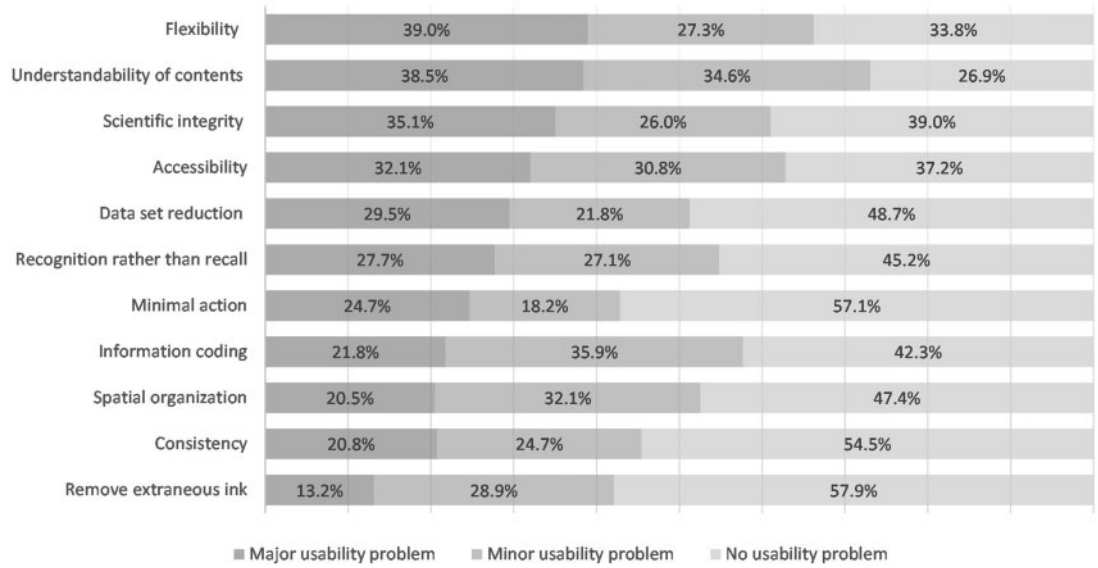

Figure 1. Relative frequency of usability problems on sexually transmitted infections dashboards in the United States. Some problems are double-counted if mentioned by multiple evaluators. The percentages are calculated across 6 evaluators and 13 dashboards (maximum = 78). For example, 39% major usability problem for flexibility means that flexibility of dashboards had major usability problems in 30 out of 78 given scores.

lic, each of whom may have varying levels of domain expertise. These users have different purposes, including exploratory data analyses (academic researchers), data-driven decision-making (public health practitioners),[10] and education or personal risk assessment (general public).[11] This study considered users' different needs in 2 ways. First, the expert reviewers had varying domain expertise and professional experience in public health practice versus academic research. Ensuring these viewpoints and experiences were present among experts was important because domain experts are accustomed to viewing similar charts and, therefore, may find different usability problems than non-domain experts. Second, the usability checklist included items that addressed the perspectives of multiple users. The scientific integrity principle, critical for many public health practitioners and academic researchers, included items, such as the use of appropriate measures and sufficient granularity. Anticipating that non-domain experts and the general public would require more information to interpret each chart,[34] the checklist included items, such as the explanation of technical terms and guidance for interpretation under the understandability of contents principle.

In contrast to previous usability evaluations,[6,7,11] this study evaluated public health data dashboards with different levels of complexity which yielded interesting findings regarding the relationship between complexity and usability of data dashboards. The identified usability problems depended on the level of dashboard complexity. Some items were identified as usability problems on complex dashboards but not on simple dashboards. For example, improving the understandability of contents was more urgent on complex dashboards but improving scientific integrity was a more critical problem on simple dashboards. These findings suggest that specific considerations are needed to avoid losing complexity while developing usable dashboards.

The usability checklist extends the existing guidelines for creating usable dashboards and systematically evaluating the existing dashboards. For example, a well-known set of dashboard design guidelines[35] comprise 4 rules: (1) information should be organized based on its meaning and use (eg, business functions for a business dashboard), (2) dashboard sections should be visually consistent to help users gain a quick interpretation, (3) viewing experiences should be aesthetically pleasing to communicate messages simply and clearly (eg, using less saturated colors and readable text), and (4) dashboards should provide a high-level insight which is supplemented with additional information through drill-down or filtering abilities. In addition to these rules, the current study's usability checklist includes specific items for each rule and additional rules that might be more relevant to public health dashboards (eg, understandability of contents and flexibility).

The usability checklist can guide the development of future public health data dashboards; however, some additional points need to be considered when making complex public health data dashboards. The ideal public health dashboard should have few usability problems and extensive complexity. The need for more complexity in public health dashboards was documented in previous user-centered design studies, which found that public health dashboards need to provide detailed information to support the analysis tasks required by their users.[20,36–38] However, users without domain expertise have been absent in previous requirement assessment studies. Moreover, the information visualization literature recommends filtering or faceting methods to deal with complexity, while retaining usability.[39] The current study found that neither method alone could achieve a usable and complex dashboard. Instead, dashboards that used both methods achieved a balance of usability and complexity, which might optimize the usefulness of public health dashboards for non-domain experts and domain experts. Another consideration for using the checklist is attention to new techniques and principles. General usability principles, such as consistency and flexibility, have remained unchanged over time.[8] However, some specific principles might change when new techniques are introduced, or users get accustomed to new features on common products. For example, guiding users to navigate the system and find interaction possibilities might become less important as users get accustomed to interactive dashboards. Future dashboard designers should consider these changes when using this checklist.

Table 3. Usability checklist for public health dashboards

Spatial organization	
Major considerations	• Are headings and titles presented in a unique and bold format to provide an overview of the content? (Yes, No, NA) • Is there a clear hierarchical relationship between headings and subheadings? (Yes, No, NA) • Does each section only include the information that is implied by its heading? (Yes, No, NA) • Are the textual information and visualizations spread out throughout the dashboard? (Yes, No, NA) • In dashboards with multiple views, is it clear how users should navigate between views? (Yes, No, NA)
Minor considerations	• Are long layouts that require too much scrolling avoided? (Yes, No, NA) • Are similar visuals organized on the same side of the screen to make the layout more organized and easier for users to navigate? (Yes, No, NA) • Are drop-down menus positioned in a location to clearly identify for which chart they are available? (Yes, No, NA) • Are visualizations checked for potential occlusion or blockage of information in the charts, labels, or titles at different zoom levels? (Yes, No, NA) • Is it clear how to navigate to the next page and the total number of pages in a multi-page dashboard? (Yes, No, NA)
Information coding	
Major considerations	• Is the use of the same color for different concepts, or else different colors for the same concept, avoided? (Yes, No, NA) • Do all visual elements (eg, color, shape, and size) have supporting legends and labels to help users understand the meaning of visualizations? (Yes, No, NA) • Are familiar visuals (eg, icons and chart elements) used to capitalize on users' prior knowledge? (Yes, No, NA) • Are bar charts sorted to be understood more quickly and more easily? (Yes, No, NA) • When side-by-side charts are provided for comparison, are the elements aligned, and are the same scales used for the charts? (Yes, No, NA) • If an object has a contrasting size, color, and orientation, compared to other similar objects, does it have a different meaning? (Yes, No, NA)
Minor considerations	• When statements accompany charts, is the evidence for the statement (eg, the relevant value) highlighted on visualizations to help users view the statement in the data? (Yes, No, NA) • For color coding, are the colors easily distinguishable? (Yes, No, NA) • Are drastic changes in the axis avoided if users select different variables on interactive charts? (Yes, No, NA) • Are clear titles and labels used to guide users in reading the charts? (Yes, No, NA) • Are acronyms and jargon avoided in titles and labels unless they are clearly defined? (Yes, No, NA) • Are explanations provided for less common visualization formats (eg, dumbbell charts or Sankey diagrams)? (Yes, No, NA) • Is bold font restricted to important text such as titles? (Yes, No, NA)
Consistency	
Major considerations	• Are consistent time periods presented in different sections of a dashboard? (Yes, No, NA) • Are the arrangement and position of data entry fields consistent throughout the system so that once users learn where they are expected to input values, they can easily do that in other panels? (Yes, No, NA) • Is consistent color coding used across different views (eg, red means male, and yellow means female across all pages)? (Yes, No, NA) • Are consistent font styles used across different views make different types of information readily identifiable to users? (Yes, No, NA)
Minor considerations	• Are interactive and static charts readily identifiable by the users? (Yes, No, NA) • Is the use of different charts for the same concept (eg, using both line charts and area charts for trends) avoided? (Yes, No, NA) • Is the inconsistent horizontal versus vertical orientation of bar charts avoided? (Yes, No, NA) • Are legends kept in consistent places to facilitate users' chart reading efforts? (Yes, No, NA)
Remove extraneous ink	
Major considerations	• Do data tables provide information that is not available on the charts? (Yes, No, NA) • If not, can they be made available for download? (Yes, No, NA) • Can extra information be added as an "expand extra info" button? (Yes, No, NA) • Are too many instructions avoided, as they might signal that the tool is hard to use? (Yes, No, NA) • Are too many colors avoided? (Yes, No, NA) • Are long and overwhelming titles avoided? (Yes, No, NA)
Minor considerations	• Are colorful images avoided that could distract users from the main content? (Yes, No, NA) • Does the textual information change to only fit the user's selection? (Yes, No, NA) • Are background images and non-essential borders removed from the charts? (Yes, No, NA) • Is it possible to make some of the labels only available on-demand? (Yes, No, NA)
Recognition rather than recall	
Major considerations	• Are visual cues, in addition to instructions, provided for users to guide them on interacting? (Yes, No, NA) • Are the data entry fields immediately visible? (Yes, No, NA) • Is it clear what users should expect upon each interaction? (Yes, No, NA) • Are users' selections shown in the data entry box? (Yes, No, NA)

(continued)

	• Is it clear what changes upon clicking on each selectable item (eg, the change could be incorporated in the title or color-coding of interactive charts)? (Yes, No, NA)
	• For each chart, are complete titles and legible legends and labels used? (Yes, No, NA)
	• Do prompts have different colors and sizes to distinguish them from other items? (Yes, No, NA)
	• Is it clear how to get back to the homepage at any time during user interaction? (Yes, No, NA)
Minor considerations	• Do pop-up windows provide additional and not repetitive information? (Yes, No, NA)
	• Are they shown in a way that does not block other information? (Yes, No, NA)
	• Do the chart titles change as appropriate upon user selection? (Yes, No, NA)
	• Are instructions placed where eyes are most likely to look for them (eg, at the top of the chart rather than in footnotes or the introduction text)? (Yes, No, NA)
	• Is there sufficient space between charts, with appropriate sectioning of the information with white space to guide the eye in the right direction? (Yes, No, NA)
	• Is it possible for users to compare their multiple selections? (Yes, No, NA)
	• Is there a loading animation or other "waiting" signal to indicate the page is actively working while users wait for the system's response? (Yes, No, NA)
	• Is jargon avoided in the instructions and prompts? (Yes, No, NA)
Minimal action	
Major considerations	• Are unwanted movements prevented (eg, interferences between map and page scrolls)? (Yes, No, NA)
	• Are long tables and visuals that require scrolling to view all contents avoided? (Yes, No, NA)
	• Is the data already populated to prevent requiring users to complete steps before showing any data? (Yes, No, NA)
	• Are important instructions buttons (eg, "get data" or "get chart") readily available? (Yes, No, NA)
	• Is the information provided upon interaction only limited to users' selection? (Yes, No, NA)
	• Do users' selections in previous pages carry over to subsequent pages in multi-page dashboards? (Yes, No, NA)
Minor considerations	• Is it possible to jump between pages so experienced can skip parts of the dashboard? (Yes, No, NA)
	• Are too many interaction options and actions avoided? (Yes, No, NA)
	• Instead of asking users to hit the "submit" button, are user selections automatically applied? (Yes, No, NA)
Dataset reduction	
Major consideration	• Can users select time, place, or other variables to easily digest the information? (Yes, No, NA)
	• Can users make groups or clusters on the presented charts? (Yes, No, NA)
	• Can the user cut the irrelevant information to hide detailed information (eg, using an accordion design)? (Yes, No, NA)
	• Is it clear how to bring back the filtered data? (Yes, No, NA)
Flexibility	
Major considerations	• Is there enough guidance available to guide inexperienced users in the interaction and navigation? (Yes, No, NA)
	• Are there options for experienced users to download the data, make their charts, and analyze and interpret information? (Yes, No, NA)
Scientific integrity	
Major considerations	• Are both counts and population rates displayed to guide public health practice? (Yes, No, NA)
	• Are thematic maps shaded based on population rates rather than counts to prevent misleading messages about geographic areas with larger populations? (Yes, No, NA)
	• Are demographic and behavioral measures provided to allow identification of the most impacted populations? (Yes, No, NA)
	• Are data provided with a reasonable granularity (eg, counties for less populated areas and zip codes for more populated areas)? (Yes, No, NA)
	• Are data sources and data limitations clear and easy to find on the dashboard? (Yes, No, NA)
	• If percentages are the primary display, are numerators and denominators also presented for experienced users? (Yes, No, NA)
Understandability	
Major considerations	• Are technical terms, such as the name of infections, measures, and indicators briefly explained? (Yes, No, NA)
	• Is guidance provided for non-domain experts on how to interpret each chart? (Yes, No, NA)
	• Are reference points (eg, US rate or the state rate) provided to help interpret values (ie, is the rate of 75 per 100 000 low or high?) (Yes, No, NA)
	• Are lengthy data notes avoided not to overwhelm users with technical information? (Yes, No, NA)
	• Are data tables provided only as an option for advanced users to not overwhelm non-domain experts? (Yes, No, NA)
Readability	
Major considerations	• Does it take more than 3 s for the page to load? (Yes, No, NA)
	• Is color coding interpretable by users with color vision deficiency? (Yes, No, NA)
	• Is the dashboard mobile-friendly to be usable on smartphones and tablets? (Yes, No, NA)

Note: Major considerations have high priority for fixing, and minor considerations have low priority for fixing. NA: not applicable.

This study had several limitations. First, STI dashboards might have different strategic orientations and target audiences. This evaluation did not consider the differences between strategic goals and target user groups. Second, although the evaluators comprised a diverse group (with respect to discipline, domain expertise, and academic research versus public health practice experience) and used established principles for evaluation, their perspectives may not be representative of all stakeholders. Third, this study's focus on STIs in the United States might limit generalizability. However, the usability principles were not domain-specific and selected from the information visualization literature to limit the generalizability concern. Fourth, consistent with the usability literature, an inter-rater reliability analysis was not conducted on the evaluators' given scores.[29,33] It is common in quantitatively-oriented studies (eg, quality assessments for systematic reviews, developing new measures, or assessing diagnostic tools) to have larger sample sizes and test inter-rater reliability. However, these practices are not common in usability evaluations because the focus is on identifying as many usability problems as possible rather than reaching a consensus. Prior published usability evaluations examined only a few websites and 3–5 evaluators with similar expertise.[6,7,11] Compared to those usability evaluations, the current study used a larger sample of dashboards, evaluators with more diverse expertise, and detailed qualitative comments in addition to usability scores to expand the generalizability of findings. Fifth, we did not evaluate the accessibility of STI dashboards for persons with disabilities, and the final checklist does not include an accessibility dimension, which could be explored in more detail in future research.

CONCLUSION

Public health data dashboards should consider usability principles to provide a positive user experience for different audiences, including domain experts and non-domain experts. This study systematically evaluated data dashboards of STIs on US state health department websites to provide a usability checklist. The checklist can guide the development of future public health data dashboards. Furthermore, these findings can be used to guide future designers of public health data dashboards to balance usability and complexity.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS

Both authors conceived the study, finalized the rubric, and conducted usability evaluations. BA directed the research, conducted the analysis, and wrote the initial draft. EGM reviewed the results and revised the manuscript for intellectual content.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

ACKNOWLEDGMENTS

We are grateful to Rachel Hart-Malloy (New York State Department of Health), Xiaojun Yuan, Luis Luna-Reyes, and Ping Li (University at Albany), who generously took the time to evaluate the dashboards. Their scores and qualitative comments were integrated to develop the usability checklist.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

1. Ansari B, Barati M, Martin EG. Enhancing the usability and usefulness of open government data: a comprehensive review of the state of open government data visualization research. *Gov Inf Q* 2022; 39 (1): 101657.
2. Valdiserri R, Sullivan PS. Data visualization in public health promotes Sound public health practices: the AIDSVU example. *AIDS Educ Prev* 2018; 30 (1): 26–34.
3. Carroll LN, Au AP, Detwiler LT, *et al*. Visualization and analytics tools for infectious disease epidemiology: a systematic review. *J Biomed Inform* 2014; 51: 287–98.
4. Thorpe LE, Gourevitch MN. Data dashboards for advancing health and equity: proving their promise? *Am J Public Health* 2022; 112 (6): 889–92.
5. Dasgupta N, Kapadia F. The future of the public health data dashboard. *Am J Public Health* 2022; 112 (6): 886–8.
6. Alcaraz-Martinez R, Ribera-Turró M. An evaluation of accessibility of Covid-19 statistical charts of governments and health organisations for people with low vision. *Prof Inform* 2020; 29 (5): e290514.
7. Blacklow SO, Lisker S, Ng MY, *et al*. Usability, inclusivity, and content evaluation of COVID-19 contact tracing apps in the United States. *J Am Med Inform Assoc* 2021; 28 (9): 1982–9.
8. Sharp H, Preece J, Rogers Y. *Interaction Design: Beyond Human-Computer Interaction*. 5th ed. Indianapolis, IN: John Wiley & Sons, Inc.; 2019.
9. Crisan A. The importance of data visualization in combating a pandemic. *Am J Public Health* 2022; 112 (6): 893–5.
10. Preim B, Lawonn K. A survey of visual analytics for public health. *Comput Graph Forum* 2020; 39 (1): 543–80.
11. Charbonneau DH, James LN. FluView and FluNet: tools for influenza activity and surveillance. *Med Ref Serv Q* 2019; 38 (4): 358–68.
12. Tan W, Liu D, Bishu R. Web evaluation: heuristic evaluation vs. user testing. *Int J Ind Ergon* 2009; 39 (4): 621–7.
13. Tory M, Moller T. Evaluating visualizations: do expert reviews work? *IEEE Comput Graph Appl* 2005; 25 (5): 8–11.
14. Maguire M, Isherwood P. A comparison of user testing and heuristic evaluation methods for identifying website usability problems. In: Marcus A, Wang W, eds. *Design, User Experience, and Usability: Theory and Practice*. Cham: Springer; 2018: 429–38.
15. Santos BS, Silva S, Ferreira BQ. *An Exploratory Study on the Predictive Capacity of Heuristic Evaluation in Visualization Applications*. Cham: Springer; 2017: 369–83.
16. Dowding D, Merrill JA. The development of heuristics for evaluation of dashboard visualizations. *Appl Clin Inform* 2018; 9 (3): 511–8.
17. Chesson HW, Spicknall IH, Bingham A, *et al*. The estimated direct lifetime medical costs of sexually transmitted infections acquired in the United States in 2018. *Sex Transm Dis* 2021; 48 (4): 215–21.
18. Kreisel KM, Spicknall IH, Gargano JW, *et al*. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2018. *Sex Transm Dis* 2021; 48 (4): 208–14.
19. Zhang X, Sherman L, Foster M. Patients' and providers' perspectives on sexual health discussion in the United States: a scoping review. *Patient Educ Couns* 2020; 103 (11): 2205–13.

20. Ansari B. *Taking a User-Centered Design Approach to Develop a Data Dashboard for New York State Department of Health and Implications for Improving the Usability of Public Data Dashboards* [ProQuest Dissertations and Theses]. University at Albany; 2022. https://www.proquest.com/dissertations-theses/taking-user-centered-design-approach-develop-data/docview/2666596595/se-2?accountid=14166.
21. Martin EG, Begany GM. Opening government health data to the public: benefits, challenges, and lessons learned from early innovators. *J Am Med Inform Assoc* 2017; 24 (2): 345–51.
22. Martin EG, Helbig N, Shah NR. Liberating data to transform health care: New York's open data experience. *JAMA* 2014; 311 (24): 2481–2.
23. Jo G, Habib D, Varadaraj V, *et al*. COVID-19 vaccine website accessibility dashboard. *Disabil Health J* 2022; 15 (3): 101325.
24. Hyde JK, Shortell SM. The structure and organization of local and state public health agencies in the U.S.: a systematic review. *Am J Prev Med* 2012; 42 (5 Suppl 1): S29–41.
25. Singh SR, Bekemeier B, Leider JP. Local Health Departments' spending on the foundational capabilities. *J Public Health Manag Pract* 2020; 26 (1): 52–6.
26. Forsell C, Johansson J. An heuristic set for evaluation in information visualization. In: *Proceedings of the International Conference on Advanced Visual Interfaces*. ACM; 2010: 199–206; Roma, Italy.
27. Freitas CMDS, Luzzardi PRG, Cava RA, *et al*. On evaluating information visualization techniques. In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. New York, NY, USA: Association for Computing Machinery; 2002: 373–4. doi:10.1145/1556262.1556326.
28. Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behav Inf Technol* 1997; 16 (4–5): 220–31.
29. Zuk T, Schlesier L, Neumann P, *et al*. Heuristics for information visualization evaluation. In: *Proceedings of the 2006 AVI Workshop on Beyond Time and Errors*. ACM; 2006: 1–6; Venice, Italy.
30. Amar R, Stasko J. *A Knowledge Task-Based Framework for Design and Evaluation of Information Visualizations*. IEEE Computer Society Press; 2004: 143–9.
31. Shneiderman B. *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations*. IEEE Computer Society; 1996: 336–43.
32. Nielsen J, Mack RL. *Usability Inspection Methods*. New York, NY: John Wiley & Sons; 1994.
33. Nielsen J. Finding usability problems through heuristic evaluation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery; 1992: 373–80. doi:10.1145/142750.142834.
34. Ancker JS, Senathirajah Y, Kukafka R, *et al*. Design features of graphs in health risk communication: a systematic review. *J Am Med Inform Assoc* 2006; 13 (6): 608–18. doi:10.1197/jamia.M2115
35. Few S. *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, CA: Oreilly & Associates Incorporated; 2006.
36. Sutcliffe A, de Bruijn O, Thew S, *et al*. Developing visualization-based decision support tools for epidemiology. *Inf Vis* 2014; 13 (1): 3–17.
37. Robinson AC, MacEachren AM, Roth RE. Designing a web-based learning portal for geographic visualization and analysis in public health. *Health Informatics J* 2011; 17 (3): 191–208.
38. Livnat Y, Rhyne T-M, Samore MH. Epinome: a visual-analytics workbench for epidemiology data. *IEEE Comput Graph Appl* 2012; 32 (2): 89–95.
39. Munzner T. *Visualization Analysis and Design*. Boca Raton, FL: A K Peters/CRC Press; 2014.