Mapping the Landscape of COVID-19 Crisis Visualizations

Yixuan Zhang[1], Yifan Sun[2], Sumit Barua[3], Enrico Bertini[4], and Andrea G. Parker[1]
[1] Georgia Institute of Technology, [2] William & Mary, [3] Northeastern University, [4] New York University

Abstract

A great number of visualizations have been created to communicate the constantly changing crisis of the COVID-19 pandemic. With the prevalence of these crisis visualizations, there is a critical need to organize and understand what and how visualizations have been produced and disseminated to the public, as information consumption can impact peoples' attitudes, responses to crisis and risk, behaviors, and thus ultimately the trajectory of the pandemic. We curated a list of 668 visualizations that communicate information about the pandemic. We performed a content analysis of these visualizations and derived six categories of intended messages in communication about the pandemic, including informing about severity; forecasting trends and influences; explaining the course of the disease; mirroring impact of the crisis; and communicating risk, vulnerability, and equity. We also identify issues and opportunities arising from COVID-19 crisis visualizations.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods

1 Introduction

The COVID-19 pandemic is touching many aspects of human life. To help people understand, analyze, and predict this constantly evolving public health crisis, an enormous number of *crisis visualizations* focusing on COVID-19 have been created (i.e., visual representations of information such as disease prevalence, epidemiological simulations, and economic and social changes). New visualizations are being circulated daily; they are produced by diverse content creators, such as scientists, government and healthcare officials, social media users, news media outlets, and the visualization community at large. It may be the first time in history that such a large proportion of the public is engaged with and responding to visualization works.

Given the massive number of crisis visualizations, we see both opportunities (e.g., novel and exquisite visualization design) as well as challenges (e.g., avoiding visual misinformation and bias) for visualization research and practice. With this proliferation of visualization production and consumption, there is a need to organize and understand what visualizations have been designed and put out into the public because information consumption can impact people's behaviors, attitudes, and thus ultimately the path of the pandemic. So far, there is no research examining how visualizations have been produced to communicate with the public amidst COVID-19 when uncertainty, misinformation, and anxiety are rampant.

In this study, we collected 668 crisis visualizations pertaining to COVID-19. Using content analysis, we derived a set of design dimensions characterizing the design space of crisis visualizations. Specifically, we mapped the visualization space on a set of strategic messages, including informing about severity; forecasting trends and influences; explaining the course of the disease; guiding risk mitigation; mirroring impact of the crisis; and communicating risk, vulnerability, and equity. We contribute an understanding of the landscape of COVID-19 crisis visualizations and what intended messages that these visualizations attempt to communicate, as well as issues and challenges in designing visualizations for each type of message. Examining these aspects will help us understand crisis visualizations evolve over time, and how to provide and spread effective messages to the public. Building upon our observations and analysis, we argue for future work that further explore and examine the design space of crisis visualizations, such as approaches to dealing with uncertainty and scaffolding crisis visualization literacy.

2 Methodology

This paper reports on our content analysis of a COVID-19 visualization collection that we assembled. To compile this collection, we used opportunistic sampling, also called emergent sampling, a non-probability sampling method for data collection. Opportunistic sampling "takes advantage of unforeseen opportunities after fieldwork has begun," and is particularly useful for synthesizing exploratory research areas [29]. The time sensitive and unpredictable evolution of the COVID-19 pandemic can make a priori sampling decisions challenging, thus motivating our use of opportunistic sampling, an approach that enabled us to flexibly collect visualizations as they were produced and disseminated. However, opportunistic sampling suffers from selection biases, necessitating care when interpreting the results.

Our collection includes information visualizations, data visualizations, and infographics focused on COVID-19. These terms can be defined distinctly depending on the meanings and approaches. Yet, the purpose of each is to visually present complex information in a planned and comprehensible manner [10]. Our goal is to capture and present the diversity and versatility of these crisis visualizations.

At the beginning of the data collection, early March 2020, we collected an initial set of visualization. Later in March, to expand the collection, we began compiling a corpus of crisis visualizations focusing on COVID-19 using image database searches (e.g., Google Images) and word-of-mouth for contribution. We also searched on the visualization blogs that contain special topics on the pandemic, such as `coronavirus` on FlowingData, and `Science & Health` on FiveThirtyEight. Whenever possible, we included the original visualization that appeared online in our database. Some visualizations in our collection have been updated regularly; therefore our analysis reflects the state of these visualizations at the time we visited the website.

To focus the scope of this work, we analyzed a subset of collected visualizations. The inclusion criteria for this work include visualizations in English that were published online and aimed to communicate with the general public. We did not include academic publications, which typically assume specialized audiences. With these criteria in mind, this paper reports on our analysis of 668 visualizations (published between January 22 and July 7, 2020).

We performed a content analysis on the collected visualizations. Two researchers in the team first went individually through the collection to develop an initial understanding of the visualizations, and then open-coded the first 100 visualizations in the collection. Then the two researchers discussed the coding scheme regularly to review the evolving codes to achieve a mutual understanding and to refine the coding scheme. After several iterations of the coding scheme, we ended up with a codebook to guide the analysis. Guided by the codebook, we recoded the previously-coded visualizations, and completed coding of the rest of the visualizations. At this point, we measured the inter-rater reliability (97% agreement between the coders). Then we discussed the items on which we had disagreement until we reached a consensus. The codebook, visualization examples,

[1] yixuan@gatech.edu, andrea@cc.gatech.edu
[2] ysun25@wm.edu
[3] barua.s@northeastern.edu
[4] enrico.bertini@nyu.edu

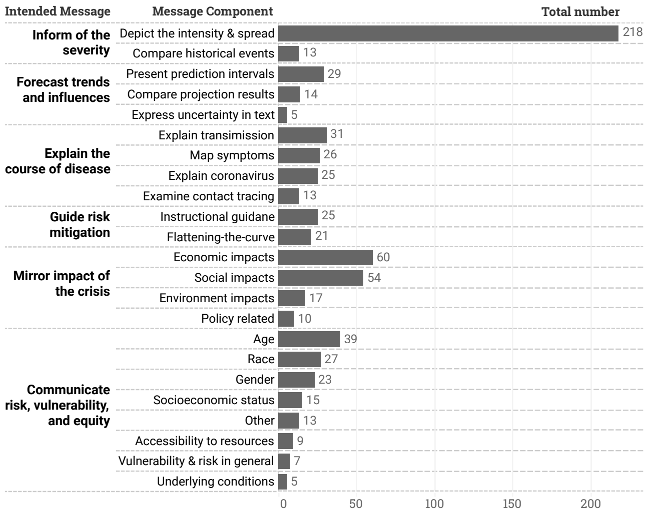


Figure 1: A summary of number of visualizations in our collection broken down by intended messages

and the collection used for this paper are provided as supplemental material at osf.io/turc3.

3 Visualizations by Intended Messages

Along with the COVID-19 pandemic came a big wave of visualizations that attempt to help the general public understand the development of the pandemic and manage its uncertainty. These visualizations touch on an unprecedented broad range of topics and convey a wide range of messages. We classify the intended messages into these purposes: (1) **informing** about severity; (2) **forecasting** trends and influences; (3) **explaining** the course of the disease including causes, symptoms, transmission, and contact tracing; (4) **guiding** risk mitigation; (5) **mirroring** impact of the crisis; and (6) **communicating** risk, vulnerability, and equity. Note that the messages are not intended to be mutually exclusive. It is common for one visualization to convey more than one message. Yet, our classification aims to provide an overview of crisis visualizations during the COVID-19 pandemic and help others identify how the intended messages and visualizations can be applied, as well as to guide future research to tackle the challenges in designing visualizations for each type of the messages.

3.1 Informing about Severity

Severity, as represented by the number of diagnoses, hospitalized patients, incubated patients, and deaths, is one essential aspect that people want to understand amidst COVID-19. The importance of informing about severity is reflected by the large proportion (231 out of 668) of the visualizations in our database. Most visualizations have been created to inform the public about the intensity and spread of the current situation (218 out of 231, 94%), while some others (13 out of 231, 6%) compare COVID-19 with historical events.

3.1.1 Depicting the Intensity and Spread

Visualizations to depict the intensity and spread of the pandemic (n=218) focus on presenting current state and conditions of the pandemic and tend to be data-driven (e.g., visualizing deaths, confirmed cases, testing, hospitalization and recovery, and hospital resource usage). The creation of visualizations is an iterative and dynamic process, even more so in times of crisis. Our analysis suggests that visualizations have evolved constantly from the initial to the maintenance stage of the COVID-19 pandemic.

Temporal visualizations focus on depicting the trajectory of the pandemic over time. The most basic solution is to visualize daily and cumulative numbers over a time period, mostly using bar charts, line charts, and area charts. Two major variations in visualizing the temporal change of the pandemic include using logarithmic (log) scale and adding moving average. In our collection, 14 visualizations provided both log and linear scale (10 enabled users to toggle between modes), and 6 provided only the log scale. Using log scale may help better display the trend. However, one study found that people had less accurate understanding of the trajectory of the pandemic when showed the number of deaths on a log scale [31].

The moving average provides a more stable view of the trend than daily change and is an indicator for assessment of the effectiveness of surveillance and containment during the outbreak [25]. Most visualizations did not use moving average at the beginning of our data collection. To our knowledge, the earliest evidence of using moving averages was from Singapore's report on February 29, 2020 [25]. Later on, more visualizations started adding moving averages to better reflect the current state (n=30). For example, the New York Times added the 7-day average starting from April 8.

Novel visualization techniques have been introduced, such as Pez Charts, Growth Charts, and the use of event alignment. Though these more advanced methods have been studied and used before in academic research, they were not initially used for communicating information about the pandemic. The patterns conveyed in these novel visualizations are particularly relevant. The Pez Chart [40] shows case number changes in horizontally-placed blocks, with the x-axis representing time and the color of each block representing the values. When stacking multiple horizontal series together, it is easier to compare trends across multiple regions. Moreover, temporal event alignment is useful to reveal patterns that emerge by aligning the occurrence of events of interest over time [43, 44]. The trajectory of the pandemic can be seen through an alignment of days because cases or deaths across regions (on the x-axis) is usually presented in log scale. Comparing with the reference lines (i.e., slope of a curve) helps demonstrate doubling rate. Another emerging approach is using a type of Growth Chart that plots the total number of the confirmed cases as the x-axis and the weekly confirmed cases as the y-axis [3]. Applying the log scale to both axes and using the total case number as the x-axis (rather than time) helps reveal trends and patterns (e.g., demonstrating if cases are exponentially growing in a region or if a region is on the path to containing the virus).

Through the emerging visualization styles, we see that existing approaches may be insufficient to meet the challenges of visualizing the COVID-19 pandemic. There are still opportunities to develop new visualization methods to show trends and patterns of the development of the pandemic and to help people understand its severity. Yet, using new visualization methods may pose a challenge for people to interpret the message accurately. Therefore, care must be taken when using novel visualization techniques or approaches that the general public are less familiar with (e.g., explicitly explaining how to interpret visualizations for the public might be helpful).

Visualizing geospatial spread, another common approach, is to visually display variables of interest over a geographical map to demonstrate which regions are impacted and compare how the impact differs by regions. Among all the map-based visualizations we analyzed, choropleth maps are one of the most popular options (50% of the visualizations about intensity and spread), followed by proportional symbol maps (i.e., 29% bubble maps). Visualization creators typically need to decide the base map type, whether to normalize the data, and the shape for proportional symbol maps.

Both choropleth and proportional symbol maps overlay information on a base map; thus, the selection of a base map is an important decision. Most map-based visualisations use a map that represents the physical shapes of geographical regions. Traditional projections (e.g., Mercator, Robinson) are widely used, while a few other visualizations use other projections, such as the Cahill–Keyes projection  to enable symmetry of component maps [7], and the Armadillo projection  to provide a perspective to show most of the globe [36]. Tile grid maps are also commonly used. These maps abstract away the physical shape of geographical regions so that viewers can focus on the overlaying information.

One major challenge of map-based visualization is to effectively lay out the data and avoid overcrowding. Bubble maps, as the most common proportional symbol map, face a challenge of over cluttering the data. An interesting alternative proportional symbol map is to use of triangles (rather than circles) to show cumulative deaths and confirmed cases [17].

Another challenge with crisis visualization using choropleth maps is that color coding raw data may mislead viewers. For example, if a highly-populated state has the same case number as a less-populated state, the color will be the same on the map. A reader may see the map as the intensity are similar in the two states, while the actual intensity in the less-populated state is higher. Normalizing data may help remove this type of bias [8]. Typical normalization variations include normalizing by area, relevant population, a prior date, central tendency (e.g., mean, median, mode), and variability (e.g., standard deviation, above and below range) [15]. Approximately only 37% of the visualizations analyzed used some form of normalization.

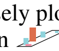
In addition to normalization, map classification—splitting data into aggregated classes and categories systematically—is another critical process as the choice of a classification scheme can influence the intended message of the map entirely. *Classification* can be done through a variety of approaches, including classed maps by splitting data with pre-defined classes by equal intervals, quantiles, natural breaks, standard deviation, and unclassed by leaving data unclassified [15]. Approximately 67% and 13% of choropleth maps used classed and unclassed methodologies, respectively. The remaining 20% were left with unknown classification. Unclassed maps may be more accurate to portray nuances in the distribution of data than classed maps, if designed carefully and properly. Yet, they may also become overwhelming for the viewers due to the potential information overload, and thus, making the map less effective to communicate an overall pattern [16].

In addition, maps have applied various color schemes, including diverging (28%), sequential (65%), and a mix of diverging and sequential methods (7%). Moreover, a majority of maps applied light background (e.g., white, light grey), but about 17% used dark background (e.g., black, dark blue). Prior work has examined color-related biases (e.g., dark-is-more, contrast-is-more bias) and the choice of background color [33]. These design choices, while nuanced, may also impact how people perceive the severity and risk.

Multivariate visualizations use both geospatial and temporal data to display two or more variables in a visualization. Approaches include superimposing small multiples with tile grid maps showing cumulative number of cases over time in different locations (e.g., sparkline maps), and using concentric bubble maps to show multiple metrics (e.g., positive and total tests) in absolute numbers.

3.1.2 Comparisons with Historical Events

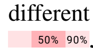
History serves as a mirror to learn from the past. To depict the severity of the crisis, some visualizations compare between the current pandemic with other diseases and other major outbreaks in history (n=13). Most of these visualizations (8 out of 13, 62%)in our collection were published in the early stage of the outbreak (January 1 to March 31, 2020). This might be because there has been a greater level of uncertainty during the initial phase of the crisis.

Visualizations in this category use side-by-side table-like representations to compare characteristics between the current pandemic and other events, such as basic reproduction number, incubation time, hospitalization rate, case fatality rate, and cumulative deaths. In addition to merely comparing these characteristics, other visualizations embedded a timeline of the historical public health crisis events and/or other diseases to provide a holistic overview of the data along the timeline, such as using horizontal timeline and 3D views  along the Z-axis (the depth) [22], while still keeping data comparable. Having criticized of not being able to precisely plot the overall death toll, an alternative "distorted" 3D version  has been proposed to allow for easy comparison of the death toll [32].

The message behind these visualizations attempts to distinguish the current pandemic from previous ones and to further inform about severity. It is challenging to compare current events with historical events due to potential discrepancies between COVID-19 and previous pandemics. Potentially relevant difference range from the course of the disease (e.g., incubation periods, transmission rate) to public health context. To address the difficulty of comparing historical events, visualizations have to incorporate explanations about how and whether past events can be compared to current events [1].

3.2 Forecasting Trends and Influences

Visualizations have been created to estimate and forecast trends and influences in terms of estimation of current transmissibility, effective reproduction rate, and projections of the pandemic (57 out of 668, 9%). Forecasting models always have a certain level of uncertainty therefore visualizing uncertainty is essential. Many visualizations in this category use graphical annotations to quantify uncertainty (48 out of 57, 84%). This type of uncertainty refers to direct uncertainty [38], the type of uncertainty that focuses on fact, numbers, or scientific hypotheses and that typically can be communicated quantitatively (e.g., probability distributions, confidence intervals, and ratios). Among the collected visualizations for forecasting and projection, major approaches to communicate uncertainty include: presenting prediction intervals (29 out of 48, 60%), comparing projection results of multiple models or different scenarios (14 out of 48, 29%), and communicating uncertainty in text (5 out of 48, 10%).

Visualizing uncertainty interval for projections (n=29) includes using solid line charts with shade  (16 out of 29, 55%), solid line indicating current and prior situation followed by dashed line with shade indicating projection  (6 out of 29, 21%), superimposing multiple levels of confidence intervals (4 out of 29, 14%) shown with different opacity using interval funnels  and interval range bar 

, as well as showing range (3 out of 29, 10% for visualizing effective reproduction rate).

As more and more forecasting models are generated, visualizations are more commonly used to compare between different models or scenarios, e.g., using ensemble plots [23] (n=9) and multiple hypothetical outcomes (n=5). The goal is to present differences in the assumptions and corresponding confidence intervals as well as to help the public understand potential limitations of the forecasts. Multiple hypothetical outcomes can be useful to understand various scenarios regarding what might happen in the future. For example, visualizations that juxtapose small multiple choropleth maps in the same view  show how different infection prevention and control measures might influence the outbreak's spread [18]. This type of visualization can be more intuitive to help people understand the probability amidst this complex pandemic.

We also observed some issues in existing visualizations, including *not visualizing uncertainty*, *lack of clarity* of visualizing uncertainty (e.g., no legend or labeling indicating what the uncertainty is about), as well as *inconsistency of communicating uncertainty* (e.g., description in text but no visual display). Prior work has found that presentation of uncertainty in text increases level of anxiety and perceived risk, whereas visually depicting uncertainty may decrease anxiety and risk [38]. More research is needed to further examine how communicating uncertainty shapes people's reactions and actions in crisis situations. Understanding the nuances of the effect of presentation methods will help design more effective crisis visualizations.

3.3 Explaining the Course of the Disease

Many visualizations aim to help people understand the nature of the virus (25 out of 95, 26%), symptoms of illness, the process and timeline of diagnosis (26 out of 95, 27%), disease transmission (31 out of 95, 33%), and contact tracing (13 out of 95, 14%). In the following we describe these categories in more detail.

3.3.1 Explaining Causes, Symptoms, and Transmission

This body of medical visualizations focuses on describing the virus and its structure including presenting information about its genomic epidemiology and genetic sequences. Communicating scientific research to the general public is important for scientists, yet it is also challenging to effectively communicate with diverse audiences [5]. Visualization designers have made efforts to help bring science-based information closer to the public. For example, the medical illustrations team from CDC attempted to make the abstract medical concepts more approachable and comprehensible: *"The colours were chosen for visual impact. The bold red of the S proteins contrasted by the gray of the viral wall, adds a feeling of alarm... Shadows add to the realism"* [13].

Visualizations have been created to help people understand the symptoms and to compare symptoms between COVID-19 and other conditions (e.g., allergies, cold, and, influenza). Side-by-side table-like visualizations have been widely used for symptom comparison. Some attributes of this type of visualizations include using categorical color encoding to represent various frequencies of symptoms (like rare, sometimes, and common), and embedding pictorial aids next to the text. While side-by-side tables are easy for comparison, *symptom-body maps* that map associated symptoms on the body shape help visualize the invisible to become visible. Incorporating timelines to the type of symptom-body map helps people understand how symptoms, infections, and complications develop according to the incubation time. Using pictorial aids in health communication has been shown to be an effective way to facilitate understanding and improve recall of medical instructions [2]. Though it is promising to incorporate pictorial forms to help interpreting textual explanations, it is also important to keep in mind that these visual aids might be too complex to understand or they may fail in reflect viewers' expectations [19].

Messages that aim at explaining transmission focus on explaining how the virus spreads or the various way in which it can spread. For example, a 3D simulation visualization [28] vividly explains the possible transmission routes and attempts to persuade people to keep social distance. To that end, these visualizations also aim to help people *understand the reasonableness* of recommended actions (e.g., social distancing) and reinforce personal responsibility to reduce harm to other individuals and the community [30].

3.3.2 Examining Contact Tracing

Contact tracing is an approach to facilitate the understanding of disease transmission, the epidemiology of a disease within particular populations, and helping the control of a disease [11]. Visualizing contact tracing can be classified into explaining the process behind contact tracing and visualizing individual-level or group-level tracing. At an individual level, visualizations illustrate patient movement paths and timelines to explain how an individual might have come into contact with a person infected with the virus.

At the group-level, node-link networks explore clusters of core groups or super spreaders.

Contact tracing is often connected with surveillance and privacy issues. Visualization also plays an important role in privacy preservation for anonymization and fine-tuning parameters for protecting patients' privacy. More work is needed to examine how algorithms, taxonomies, and visualization techniques [4] have been applied to preserve or not preserve privacy while combating COVID-19.

3.4 Guiding Risk Mitigation

Visualizations have been also created to guide risk mitigation, including practical instructional guidance to mitigate risks (n=44), and conceptual visual representations to flatten the curve (n=21).

Providing Instructional Guidance: Infographics, especially instructional graphics, have been produced and distributed widely to provide instructional guidance on mitigating strategies, such as how to maintain social distancing, prepare for personal protective equipment (PPE), and keep hygiene. Visual metaphors are heavily used (95%) to help people understand things that people are unfamiliar with by comparing it with things that people are familiar with to match people's mental model. People tend to pay attention to new information if it is associated with a mental model that is meaningful to an individual [9]. For example, guidance to social distancing designed by a Japan organization adopted the metaphor of tatami (a type of mat in traditional Japanese-style rooms) to guide people to "stay one tatami apart" [27], while a Florida county reminded people to "stay one alligator apart" [12].

Conceptualizing Flattening-the-curve: The rhetorical force of the *flattening-the-curve* chart, particularly with the addition of a horizontal line marking "healthcare system capacity" was prevalent. Flattening-the-curve charts aim to conceptually depict the management of the healthcare systems' capacity. A variety of versions have been created and adapted, such as using animated and cartoon-style [42]. They may be effective to emphasize personal responsibilities to minimize the unprecedented strain on the health system [20, 34]. This type of visuals have become controversial amidst the pandemic as they simplified the complex pandemic situations (e.g., it is good enough to keep the patient number within the healthcare system capacity) but can be also misinterpreted by the public (e.g., fast rising leads to fast dropping). More work is needed to further examine the effectiveness of data-driven empirical visualizations and conceptual visuals and the effect of these graphs.

3.5 Mirroring Impact of the Crisis:

Another trend is about examining the initial impact of the crisis (n=141), including *government response and interventions*; *economic disruptions* such as unemployment, GDP, S&P ratings, and business sales; *social disruptions* due to school closures, quarantine, and lockdown in response to government policies; and *environmental impact* such as the change of urban pollution and vibration of the Earth's surface. The impact of the pandemic can be multi-faceted and interconnected. This body of work shows the shift in daily life due to the crisis and it also serves as evidence of the effectiveness of interventions to risk mitigation.

There are two major challenges for the visualizations that aim to demonstrate impacts of the crisis. The first challenge is how to vividly tell a story about how certain aspects of daily life are vastly different from regular times. To make distinguishable comparisons, for example, index charts (charts that use percentage as y-axis to show relative change) are commonly used (n=10). However, these charts can be also misleading as they do not use absolute values.

The second challenge is building the connection between the theme of the visualization and the major events amidst the crisis. Techniques that help building such connections include event alignment (e.g., aligning by Lunar New Year to show the change of coal consumption in China), and event annotations that indicate when key events occurred, including *point-event indicators without duration* (e.g., first death in Wuhan) and *interval-event indicators* (e.g., duration shutdown order in effect).

3.6 Communicating Risk, Vulnerability, and Equity

Previously identified potential risk factors for severe diseases include age, race/ethnicity, gender, some medical conditions, the use of certain medications, poverty, crowding, certain occupations, and pregnancy [6]. These factors are also influenced by the social determinants of health focusing on social and economic conditions that influence the health of individuals and communities, such as income, education, employment, social support, and access to health care [24]. Visualizations have been created to communicate risk, unpack vulnerability, and health equity amidst the pandemic (124 out of 668, 19%).

Unpacking Vulnerability and Risk in General: This body of visualizations (7 out of 81, 9%) typically use aggregated scores or scales, such as CDC's Social Vulnerability Index (SVI) and Behavioral Risk Factor Surveillance System (BRFSS). Choropleth maps are commonly used to show geospatial distribution of vulnerable populations. Visualization has been created to reveal health disparities using multiple metrics including majority-minority, highly vulnerable, overcrowded households, and most uninsured [41].

Unpacking Potential Risk Factors An increasing number of visualizations have added the disaggregated demographic factors in their existing visualizations as governmental authorities began to release demographic information of patients. Bar charts are used to *display demographic distribution* of affected population. Age distribution was the most commonly visualized (n=39), followed by race (n=27) and gender (n=23). Some visualizations (n=5) *unpack the influence of underlying health conditions*, one risk factor [24], to understand the characteristics of affected population, and the likelihood of people seeking intensive care treatment. Limited access to healthcare is another risk factor for severe disease [24]. People who have limited access to health care resources are likely to be more vulnerable during a public health crisis [30]. To help understand these factors, visualizations have been created to show *accessibility and allocation of resources* (n=4), such as visualizing the travel distance to nearest hospitals to get treatment and displaying the geospatial spread of tangible social support (e.g., food services). Moreover, scrollytelling visualizations have been designed to explain who should get access to medical services first when resources are limited, e.g., sickest first, an equal chance, and maximizing treatment benefits [14]. Such visualizations highlight a socioeconomic challenge in terms of imperfectness of healthcare system and may shed light on how to reform and improve the system.

Examining Social Determinants of Health Factors like chronic illness conditions, accessibility to healthcare are also influenced by the social determinants of health focusing on socioeconomic status (SES) that influence the health of individuals and communities, such as income, education, and employment [24]. Disparities in conditions in turn are associated with health inequities. Visualizations have been created to examine associations of SES and the pandemic (15 out of 81, 19%). Various types of visualizations have been adopted in this category, including scatter plots for displaying the relationship of physical proximity and the exposure to diseases, bar charts to show ranked order of jobs based on forced policies of physical distancing, and index charts that compare high and low income disparities in movement change.

Despite the invaluable effort to unpack and address the vulnerability and equity issues, we also found that visualizations in this category were not equally created across information outlets. In our collection, visualizations created by government agencies only show basic population distribution by age, race, and gender, with a couple of visualizations (n=2) displaying underlying condition. However, none examines other SES-related factors. Instead, work examining aspects regarding SES, accessibility to resources, and living conditions was created by the news outlets independent media, companies, and NGOs, indicating the effort being put in revealing vulnerability and equity challenge. In addition, when visualizing risk, vulnerability, and equity, one caveat is to avoid stigmatization that the pandemic is not targeting one specific group.

4 Discussion

So far, we have presented how visualizations have been created to convey various massages during the pandemic, and challenges and opportunities of each type of message. We further reflect upon our observations and analysis of existing crisis visualization.

Dealing with uncertainty: Crises take a toll on human life and are inherently characterized by change, uncertainty, and complexity [35]. Uncertainty is one the greatest concerns for most in a crisis [30]. Uncertainty can also be caused by currently unavailable, missing, incomplete, and inconsistent information. Unlike direct uncertainty that emphasizes quantifiable statistical uncertainty (e.g., Section 3.2), indirect uncertainty is about the meta-level evidence of underlying information that forms claims of the fact, number, or hypotheses [38]. For example, research has investigated the COVID-19 data quality and suggested that countries with lower Healthcare Access and Quality Index may have underreported cases or are not capable to detect them adequately [21]. Therefore, care must be taken when interpreting and visualizing crisis data.

Our analysis identifies some attempts to communicating indirect uncertainty, including disclaimers and explanations of data collection and visualization methodology (e.g., providing guidance on how to interpret data and describing potentially missing data or inconsistency in reporting), assessment of overall quality of evidence (e.g., assigning data-quality grade based on assessment of the completeness of the reporting by the COVID tracking project [37], and certain conceptual simulation visualization that allow users to set parameters.

Communicating uncertainty is not only about examining quantifiable direct uncertainty (e.g., showing probabilities with graphical annotations), but also about self-efficacy, value judgments, and assessments of intention [30]. Prior work has suggested that uncertainty increases engagement and enhances active learning [26]. Attempts to communicate uncertainty might bring opportunities in educating the general public about how to understand, interpret, and handle uncertainty in times of crisis. More research is needed to further examine effective approaches to communicate uncertainty, both direct and indirect uncertainty, with the public.

Scaffolding crisis visualization literacy: We believe there is a need to scaffold *crisis visualization literacy*—the ability to access, understand, analyze, evaluate, and present crisis information in visual formats. Crisis visualization has its own unique traits.

First, as we mentioned, visualizing crisis information involves high-level uncertainty. To effectively communicate with the public during a crisis, communication should be open, transparent, and culturally relevant about what is known and unknown [39]. On the other hand, flexibility should also be given [30], but be coupled with skepticism and critical thinking about reported visualization which may be misleading, confusing, or failing to make a point due to incomplete and low-quality of data, inappropriate mapping between data and charts, and intentional visual disinformation.

Second, the visualization of crisis information evolves as the crisis develops. During the past months, many visualizations on the severity of the pandemic have progressively included new features, such as moving-average lines, functions to toggle between linear and log scale, changing color scheme in choropleth maps, and applying normalization for choropleth maps to address potential bias.

Scaffolding crisis visualization literacy is critical, especially for groups vulnerable to visual misinformation (low-SES and older adults [45]). Building such literacy is crucial for creating an equitable society in which people of all backgrounds are able to utilize, benefit from, question, and challenge information that is disseminated about the humanitarian challenges that impact our world.

5 Limitations and Conclusion

We curated crisis visualizations amidst COVID-19. It is inevitable that our collection fails to capture all visualizations. We welcome more visualizations to be added since we have been continuously maintaining the collection. Yet, our goal is to present the diversity and versatility of these crisis visualizations, as well as issues and considerations when designing for each type of message. We have been working on creating an online digital archive system to allow people to explore and analyze COVID-19 crisis visualizations. Not a single visualization or message can tell the whole story. Instead, multiple views of the pandemic across different fields, if designed carefully and responsibly, may help reveal the multi-faceted impacts of the crisis and fight against the pandemic.

Acknowledgments

We thank Paul Kahn for co-leading, developing, and maintaining the database, Hugh Dubberly and Dubberly Design Office for support, and people who have facilitated with data collection and coding. We also thank Ben Shneiderman, Catherine Plaisant for early discussions and advice, Anamaria Crisan, Lace Padilla, Jennifer Howell, Rumi Chunara, Racquel Fygenson, Helia Hosseinpour, Wellness lab and VIS lab at Northeastern for feedback, and Jane Kokernak for editing.

References

[1] N. Akpan and K. Elliott. How coronavirus compares to flu, ebola, and other major outbreaks. Accessed: 2020-06-28.

[2] A. Arcia, N. Suero-Tejeda, N. Spiegel-Gotsch, J. A. Luchsinger, M. Mittelman, and S. Bakken. Helping hispanic family caregivers of persons with dementia "get the picture" about health status through tailored infographics. *The Gerontologist*, 59(5):e479–e489, 2019.

[3] Bhatia, Aatish and Reich, Henry. Trajectory of world covid-19 confirmed cases. Accessed: 2020-07-02.

[4] K. Bhattacharjee, M. Chen, and A. Dasgupta. Privacy-preserving data visualization: Reflections on the state of the art and research opportunities. *STAR*, 39(3), 2020.

[5] S. E. Brownell, J. V. Price, and L. Steinman. Science communication to the general public: why we need to teach undergraduate and graduate students this skill as part of their formal scientific training. *Journal of undergraduate neuroscience education*, 12(1):E6, 2013.

[6] CDC. Assessing risk factors for severe covid-19 illness. Accessed: 2020-06-25.

[7] J. Corum and C. Zimmer. How coronavirus mutates and spreads. Accessed: 2020-06-28.

[8] B. Dent, J. Torguson, and T. Hodler. *Cartography: Thematic Map Design*. McGraw-Hill New York, New York, NY, 2008.

[9] J. Dirksen. *Design for how people learn*. New Riders, 2015.

[10] B. I. U. Dur. Data visualization and infographics in visual communication design education at the age of information. *Journal of Arts and Humanities*, 3(5):39–50, 2014.

[11] K. T. Eames and M. J. Keeling. Contact tracing and disease control. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1533):2565–2571, 2003.

[12] A. Elassar. A florida county is reminding people to maintain a distance of at least one alligator between each other. Accessed: 2020-07-07.

[13] M. Fairs. Iconic covid-19 images designed to create "a feeling of alarm" says cdc medical illustrator dan higgins. Accessed: 2020-06-28.

[14] C. Feilding. The toughest triage. Accessed: 2020-06-28.

[15] M. Foster. Statistical mapping.

[16] N. Gale and W. C. Halperin. A case for better graphics: The unclassed choropleth map. *The American Statistician*, 36(4):330–336, 1982.

[17] L. Gamio and K. Yourish. See how the coronavirus death toll grew across the u.s. Accessed: 2020-07-04.

[18] J. Glanz, L. Leatherby, M. Bloch, M. Smith, L. Buchanan, J. Wu, and N. Bogel-Burroughs. Coronavirus could overwhelm u.s. without urgent action, estimates say. Accessed: 2020-06-28.

[19] M. G. Katz, S. Kripalani, and B. D. Weiss. Use of pictorial aids in medication instructions: a review of the literature. *American journal of health-system pharmacy*, 63(23):2391–2397, 2006.

[20] R. Kosara. The visual evolution of the "flattening the curve" information graphic. Accessed: 2020-06-28.

[21] H. Lau, V. Khosrawipour, P. Kocbach, A. Mikolajczyk, H. Ichii, J. Schubert, J. Bania, and T. Khosrawipour. Internationally lost covid-19 cases. *Journal of Microbiology, Immunology and Infection*, 2020.

[22] N. LePan. History of pandemics. Accessed: 2020-06-28.

[23] L. Liu, A. P. Boone, I. T. Ruginski, L. Padilla, M. Hegarty, S. H. Creem-Regehr, W. B. Thompson, C. Yuksel, and D. H. House. Uncertainty visualization by representative sampling from prediction ensembles. *IEEE transactions on visualization and computer graphics*, 23(9):2165–2178, 2016.

[24] E. C. Lowcock, L. C. Rosella, J. Foisy, A. McGeer, and N. Crowcroft. The social determinants of health and pandemic h1n1 2009 influenza severity. *American journal of public health*, 102(8):e51–e58, 2012.

[25] Y. Ng, Z. Li, Y. X. Chua, W. L. Chaw, Z. Zhao, B. Er, R. Pung, C. J. Chiew, D. C. Lye, D. Heng, et al. Evaluation of the effectiveness of surveillance and containment measures for the first 100 patients with covid-19 in singapore–january 2–february 29, 2020. 2020.

[26] E. Ozcelik, N. E. Cagiltay, and N. S. Ozcelik. The effect of uncertainty on learning in game-like environments. *Computers & Education*, 67:12–20, 2013.

[27] Pandaid. Social distancing. Accessed: 2020-06-10.

[28] Y. Parshina-Kottas, B. Saget, K. Patanjali, O. Fleisher, and G. Gianordoli. This 3-d simulation shows why social distancing is so important. Accessed: 2020-07-01.

[29] M. Q. Patton. *Qualitative evaluation and research methods*. SAGE Publications, inc, 1990.

[30] B. Reynolds, S. Deitch, and R. A. Schieber. Crisis and emergency risk communication pandemic influenza. 2007.

[31] A. Romano, C. Sotis, G. Dominioni, and S. Guidi. Covid-19 data: The logarithmic scale misinforms the public and affects policy preferences. *Available at SSRN 3588511*, 2020.

[32] R. Sakai. What i learned from covid-19 data visualization. Accessed: 2020-06-28.

[33] K. B. Schloss, C. C. Gramazio, A. T. Silverman, M. L. Parker, and A. S. Wang. Mapping color to meaning in colormap data visualizations. *IEEE transactions on visualization and computer graphics*, 25(1):810–819, 2018.

[34] J. Sears, J. M. Villas-Boas, S. B. Villas-Boas, and V. Villas-Boas. Are we stayinghome to flatten the curve? *Department of Agricultural and Resource Economics, first version April*, 5, 2020.

[35] M. W. Seeger, T. L. Sellnow, and R. R. Ulmer. *Communication and organizational crisis*. Greenwood Publishing Group, 2003.

[36] South China Morning Post Graphics. Coronavirus: the disease covid-19 explained. Accessed: 2020-06-28.

[37] The COVID Tracking Project. State grades. Accessed: 2020-06-28.

[38] A. M. Van der Bles, S. van der Linden, A. L. Freeman, J. Mitchell, A. B. Galvao, L. Zaval, and D. J. Spiegelhalter. Communicating uncertainty about facts, numbers and science. *Royal Society open science*, 6(5):181870, 2019.

[39] E. Vaughan and T. Tinker. Effective health risk communication about pandemic influenza for vulnerable populations. *American Journal of Public Health*, 99(S2):S324–S332, 2009.

[40] B. Welsh. Pez charts. Accessed: 2020-06-28.

[41] A. Williams and A. Blanco. How the coronavirus exposed health disparities in communities of color. Accessed: 2020-07-01.

[42] M. Wilson. The story behind 'flatten the curve,' the defining chart of the coronavirus. Accessed: 2020-06-29.

[43] Y. Zhang, S. D. Bartolomeo, F. Sheng, H. Jimison, and C. Dunne. Evaluating alignment approaches in superimposed time-series and temporal event-sequence visualizations. In *2019 IEEE Visualization Conference (VIS)*, pp. 1–5, 2019.

[44] Y. Zhang, K. Chanana, and C. Dunne. Idmvis: Temporal event sequence visualization for type 1 diabetes treatment decision support. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):512–522, 2019.

[45] Y. Zhang, N. Suhaimi, R. Azghandi, M. A. Joseph, M. Kim, J. Griffin, and A. G. Parker. Understanding the use of crisis informatics technology among older adults. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020.