

Visualizing Uncertainty with Simulated Chromatic Aberration

Rashidul Islam and Stephen Brooks

Dalhousie University, 6050 University Avenue, Halifax, B3H 1W5, Canada
E-mail: sbrooks@cs.dal.ca

Abstract. *The area of uncertainty visualization attempts to determine the impact of alternative representations and evaluate their effectiveness in decision-making. Uncertainties are often an integral part of data, and model predictions often contain a significant amount of uncertain information. In this study, we explore a novel idea for a visualization to present data uncertainty using simulated chromatic aberration (CA). To produce uncertain data to visualize, we first utilized existing machine learning models to generate predictive results using public health data. We then visualize the data itself and the associated uncertainties with artificially spatially separated color channels, and the user perception of this CA representation is evaluated in a comparative user study. From quantitative analysis, it is observed that users are able to identify targets with the CA method more accurately than the comparator state-of-the-art approach. In addition, the speed of target identification was significantly faster in CA as compared to the alternative, but the subjective preferences of users do not vary significantly between the two.*

Keywords: visualization, chromatic aberration, user study, uncertainty

© 2024 Society for Imaging Science and Technology.
[DOI: 10.2352/J.Percept.Imaging.2024.7.000403]

1. INTRODUCTION

Data is an essential part of daily life, and many datasets have a degree of uncertainty either in known or unknown form. The uncertainties in data can originate from different sources, and it is important to analyze and measure the amount of uncertainty in the data. The most commonly considered aspect of uncertainty is measures of error, which can emerge from both data collection and processing of collected data [9]. However, a third source of error could be referred to as “use error,” which is associated with the application of data [2].

The first opportunity for error is at the data collection phase, which is also often referred to as source error. Source errors can include errors in the data itself or even in the identification of the data. Missteps on the part of the data collector, time pressure constraints, difficult ambient or environmental conditions, or inherent limitations in instruments used to collect the data are just some of the contributing factors to source errors. But error can also be the result of compromises and tradeoffs since the cost of highly

precise data collection may exceed its value. This cost can therefore impact both the accuracy and the completeness of the obtained dataset.

Subsequent modifications to the data collected may include abstraction, scale changes, projections, dimensionality reductions, and analog-to-digital conversions [2] as well as many types of errors resulting from modeling including machine learning. We refer to these as process or modeling errors, and the potential for such errors is ever present especially when data is subjected to a chain of multiple manipulations, each of which can contribute to compounding errors into the data.

In a non-domain specific way, we might define error as the discrepancy between measurement and true value. However, it has been noted that a universally complete definition of data quality may be difficult to define as the particular application area can be a factor [9]. Nevertheless, a variety of common measures have been developed as a metric for reliability or confidence and used across many disciplines [3].

Uncertainty visualization is an ongoing area of research [39, 49] but a topic that many commercial practitioners avoid due to the additional complexity that it introduces. However, Greis et al. [18] explored game-like experimental tasks and compared representations that communicate different amounts of uncertainty information to the user, and results showed that participants did not favor representations with no uncertainty as they valued the additional information. Deitrick et al. [14] also studied whether uncertainty visualization influences, or results in, different decisions and found through a human-subject experiment that it can be a factor.

There are several traditional approaches to handling uncertainty, including error bars [3]; however, these can often be difficult to integrate directly into more general visualizations [20]. In addition, there are various studies conducted for uncertainty representations; for example, textual representation such as captions or tooltips [29], graphical representations such as glyphs [29, 39], custom color palettes such as value-suppressing uncertainty palette (VSUP) [12], bivariate choropleth maps [35], and texture patterns [4]. But to our knowledge, no uncertainty representation has made use of chromatic aberration (CA).

Chromatic aberration is a well-known phenomenon of color distortion or alteration that is sometimes seen around high contrast edges of objects in photographs and can also

Received May 6, 2024; accepted for publication Aug. 9, 2024; published online Sept. 25, 2024. Associate Editor: James A. Ferwerda.

2575-8144/2024/7/000403/16/\$00.00

result from impaired vision. Since different colors of light refract to different angles upon traveling through materials with refractive indices [47], the resulting images may appear to be distorted [24]. Since CA is an image quality problem, most of the research concerning CA is conducted to fix the problem and improve image quality. On the other hand, uncertainty is the problem of data quality, and relevant research is most often conducted to reduce it to improve data certainty. Our goal is neither to improve image quality nor data quality; our aim is to utilize simulated and approximate chromatic aberration as a novel representation of uncertainty visualization.

To evaluate this novel approach, we first collected relevant data from well-known sources and generated uncertainty data from model predictions using machine learning models. Uncertainties were then calculated from the resultant forecasts [16]. We then visualized the uncertainty in the data using CA as well as a recent state-of-the-art competing method, VSUP. We then conducted a controlled human-computer interaction experiment to evaluate whether our new visual representation was more beneficial than the existing approach.

We chose to compare our method with VSUP [12] as it is a recent and novel visualization technique that emphasizes uncertainty over specific data values. By using color gradients, VSUP palettes help users focus on the range and impact of uncertainty rather than on precise numerical details. As with our method, VSUP is intended for situations in which understanding the variability and confidence in the data is essential for making well-informed decisions. We further chose VSUP as a comparator since they also conducted a user study which showed that it offers benefits over traditional bivariate mappings of uncertainty and value. This is due to the non-uniform budgeting of visual channels, which allows VSUP to make more efficient use of the limited visual encoding space. Our own study will test for improvements from our CA approach over VSUP itself.

2. RELATED WORK

From a vision perspective, chromatic aberration leads to various forms of color imperfections in the image. Koh et al. [30] presented a user study to observe the effect of CA on users' judgment with lateral chromatic aberration for chart reading in information visualization on display devices and suggested some guidelines for designers to avoid such issues. Other work [5, 25] proposed image warping techniques to resolve the problem. Real cameras have an aperture through which light falls on an image plane to register the image, but diffraction is an issue in this process. Therefore, [33] presents a novel rendering system for defocus blur and lens effects by approximating optical aberrations. But our purpose is quite different from these prior works as we explore the use of an approximate and simulated CA as a means to represent uncertainty.

Uncertainty is an unavoidable part of data, and due to the complexity it introduces, practitioners often avoid it in their visualizations. The term uncertainty can refer to data

quality, errors in data, or the accuracy of predictions. Given that errors are inherent in many types of data, improper or absent uncertainty representations can mislead decision-making for data analysts. Prediction generation is becoming increasingly important when using machine learning models such as neural networks (multilayer perceptron [MLP], LSTM, and GRU) [40] for performance evaluation, ARIMA or PROPHET [16, 46] for time series analysis, and XGBoost for epidemic predictions [34]. For example, a decision-support tool [38] for medical centers and health-care services has been proposed for influenza prediction and liver disease predictive analysis [47]. All these works have been conducted without specific concern for uncertainty visualization.

Botchen et al. [4] focus on the uncertainty that occurs during data acquisition and utilize texture-based techniques to visualize uncertainty in time-dependent 2D flow fields. In their system, the user can interactively manipulate aspects of the system such as particle density, error influence, or dye injection to affect the visualization of the uncertainty within the flow field. They cite several potential sources of uncertainty but focus on those resulting from data acquisition. Their solution is to use semi-Lagrangian texture advection to show flow direction by streaklines and convey uncertainty by blurring these streaklines. However, unlike our more abstract visualization approach, their solution is tailored to a specific problem, namely the representation of flow fields.

A common task in medical visualization is the partitioning of images or volumetric data into salient regions that correspond to a variety of structures, materials, or pathologies. Medical data collection often includes noise, which is a source of data uncertainty. Quite often, the segmentation task employs sophisticated computational models, which can also introduce a second layer of model uncertainty. Lundstrom et al. [36] propose probabilistic transfer functions in order to assign material probabilities to model cases. This produces a distribution of materials at every 3D location, for which animation is used where each material is shown for a duration that is proportional to its probability. There are interesting ideas introduced in this paper; however, this is also a specialized solution to a particular problem. In addition, we did not pursue an animated solution due to the well-known issues of limited short-term memory and change blindness, which can be a factor to consider for users [10].

Finger et al. [15] describe two studies in which blended icons were used to convey uncertainty regarding the identity of a radar contact as hostile or friendly. A classification study first showed that participants could sort, order, and rank icons from five sets intended to represent different levels of uncertainty. Three icon sets were selected for further study in an experiment in which participants had to identify the status of contacts as either hostile or friendly. Contacts and probabilistic estimates of their identities were depicted on a simulated radar screen in one of three ways: with degraded icons and probabilities, with non-degraded icons and probabilities, and with degraded icons only. Results

showed that participants using displays with only degraded icons performed better on some measures than the other tested conditions.

Kay et al. [29] present a novel mobile interface design and visualization of uncertainty for transit predictions on mobile phones based on discrete outcomes. In a controlled experiment, they found that quantile dot plots reduce the variance of probabilistic estimates compared to density plots and facilitate more confident estimation by end users in the context of real-time transit prediction scenarios. Other researchers [31, 35] investigated how data uncertainty visualized in maps might influence the process and outcomes of spatial decision-making, especially when made under time pressure. According to researchers, the limitations of this research are that they did not consider the effect of stress along with time constraints and it was limited to a cartographic display.

Bubble treemaps [17] combine the representation of treemaps with the quantitative encoding of bubble charts, offering a visualization method for hierarchical datasets with an element of uncertainty. It maintains the hierarchical structure of data by representing nodes as bubbles within hierarchically nested areas as represented and bounded by contoured lines. The size of the bubbles represents the magnitude of a particular measure, while the wiggles of the contours can indicate uncertainty. This method is designed for exploring hierarchical structures and understanding how uncertainty varies across different levels of the hierarchy. They present some interesting use cases but, unlike for example VSUP [12], do not evaluate its effectiveness with a user study. And while interesting and novel, it is worth noting that these bubble treemaps are a custom type of visualization for a particular type of data, namely hierarchical data, whereas our aim is to introduce a new basic visualization element that may have applicability to a variety of visual designs and data types.

We can also consider uncertainty handling in widely used software packages such as `ggdist` [26], which is an R package that enhances data distribution visualization by extending `ggplot2` with specialized tools. This also extends to `tidybayes` [21], which is an R package that integrates Bayesian statistical methods and provides a set of functions that make it easier to interpret Bayesian inferences. The package `tidybayes` builds on top of (and re-exports) several functions for visualizing uncertainty from its sister package `ggdist` [26]. One of these functions is to create curves to represent time series data, incorporating uncertainty bands to show confidence intervals. This type of plot aims to help users understand the variability and uncertainty for this type of data, providing a view of the range of possible values. However, as with our previous note regarding bubble treemaps, with the introduction of our approach, our initial aim is first to assess CA as a new basic visualization element that may have applicability to a variety of visual designs and data types rather than focusing on any particular type of data such as time series or hierarchical data. Once we have established CA's potential as a basic visual element, further

work may explore how it might be effectively incorporated into particular visualization designs, a topic that we will revisit in future work.

One approach of uncertainty visualization is to encode data values and uncertainty values independently, using two visual attributes in a bivariate map. But these resulting bivariate maps can be difficult to interpret, and the discriminability of marks can be reduced due to the interference between visual channels. To address this issue, Correll et al. [12] introduce VSUPs, and we highlight this approach as it is the comparator approach of our user study. The VSUP allocates smaller color ranges of the visual channel to data when uncertainty is high and larger ranges when uncertainty is low. This allocation of visual variables promotes patterns of decision-making that make use of uncertainty information, discouraging comparison of values in unreliable regions of the data and promoting comparison in regions of high certainty. In traditional approaches, the outputs for each combination of value and uncertainty might be represented as a 2D grid of colors whereas the VSUP approach produces a grid of pie-shaped arcs, mapping data to a smaller set of colors for higher levels of uncertainty.

However, the main limitation of VSUPs [12] results from their key design decision to filter out higher uncertainty values by grouping them altogether, which suppresses the values for decision-making when uncertainties are high. Due to this elimination of uncertainty information, designers need to carefully consider whether this representation is suitable for specific systems. Another limitation is that both uncertainty and value are represented by color, and the perceptual challenges associated with color channels are well-known. The limited ability of users to distinguish fine differences of hue means that users may struggle to match an array of hues that are simultaneously mapped to both value and uncertainty. This requires the concept of a limited "budget" of distinguishable colors. Given the limited budget of discernible hues, it necessitates that the data be quantized; due to this quantization, uncertainty visualization for continuous values is not possible.

3. UNCERTAINTY DATA GENERATION

We require data with uncertainty to proceed with our study. Although the novelty of this work does not lie in the chosen method of data prediction, for clarity, we now discuss the collected dataset, data manipulation and pre-processing, a brief description of the predictive model, and the generation of uncertainty from predicted data.

3.1 Data Collection

We chose the Covid-19 dataset from the WHO-authorized data repository, and we consolidated and validated the raw data before feeding it into several machine learning models [16, 47]. In particular, the date, location, `new_cases`, and `total_cases` are some of the useful attributes for our predictions. The full dataset includes hundreds of thousands of records for Covid data from more than 237 countries and territories.

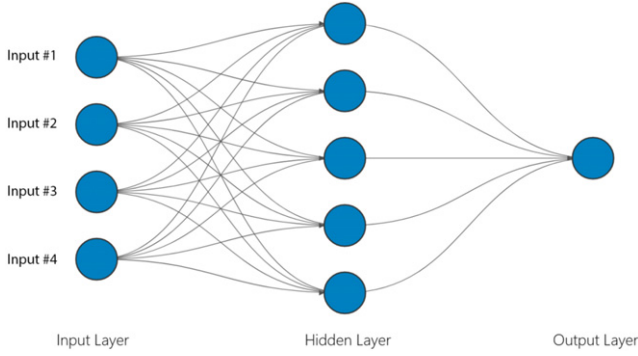


Figure 1. Essential architecture of an MLP network.

The dataset is collected as an Excel file, which includes daily occurrences and/or counts of all properties. The total_* fields such as total_cases and total_deaths are cumulative, and so every day, that is updated from the previous day's counts. If there is no value in a cell for a certain date for a country, that cell would be empty; so that needed to be handled during pre-processing. In practice, we used the average of the prior and subsequent values to fill in missing data points.

3.2 Predictive Models with Uncertainty

A time series forecasting model comprises a sequence of data points captured, using time as an input parameter. It uses historical data to predict values for the next duration, for instance, data for the next few weeks. Although we explored prediction experiments with several machine learning and statistical algorithms (MLP, CNN, LSTM, and ARIMA), for this visualization study, we used the results produced from an MLP, which is a class of feedforward artificial neural networks [8].

An MLP is a neural network connecting multiple layers in a directed graph, which means that the signal passes through the nodes only in one direction. Figure 1 shows the essential architecture of an MLP network. It can be used for time series forecasting by taking multiple observations at prior time steps, called lag observations, and using them as input features and predicting one or more time steps from those observations. The training dataset is therefore a list of samples, where each sample has some number of observations from days prior to the time being forecasted, and the forecast is the day in the sequence.

The specific process we used for setting up and running the network is as follows. We first initialize a "Sequential" model from the Keras deep learning library [11], and a dense layer is added to the model with 24 inputs and 500 nodes using the rectified linear activation function (relu). We then add another dense layer with one output. We then compile the model with the mean square error loss function and the Adam optimizer, fitting the model with the training dataset, and then obtain the prediction output for each time step t (per day). Finally, we calculate the ranges (lower bound L_t , mean M_t , and upper bound U_t) of each prediction at time t .

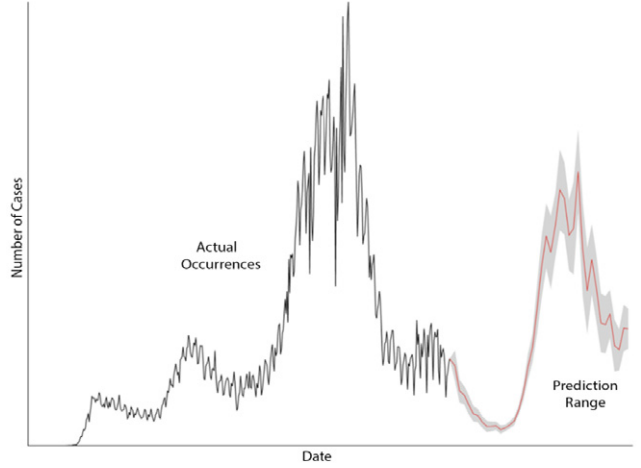


Figure 2. Example of daily Covid forecasting for 200 days.

To train the model, we chose the top 100 most impacted countries (based on the number of new cases). Uncertainties were then calculated from the ranges of predicted values for every time step (daily) during the specified 200 days of the forecasting period. Figure 2 shows the daily forecasting of the number of new cases for the United States based on previous statistics. The black line on the left shows the actual occurrences, the red line towards the right shows the predicted number of cases, and the grayed background surrounding the predicted line represents the ranges of model prediction. This means the model can predict a value between the lower and upper values for a particular day, and that gray area represents the area of uncertainty.

To map values and their variations onto display devices, we first normalize the uncertainty. We scale the range of variation in the data based on the maximum uncertainty in the data found in the country that exhibits the most uncertainty in its predictions. For this, we generate uncertainty data for each country, U_t , from the predictive results (the range between lower bound L_t and higher bound H_t) for each of 200 days, t :

$$U_t = L_t - H_t,$$

and we then sum the average uncertainty A_C of each country C dividing by the number of days t :

$$A_c = \sum \frac{U_t}{t}.$$

We then find the maximum average uncertainty from all countries, M , and then divide each country's average uncertainty by that maximum to produce a normalized average uncertainty for each country:

$$\hat{A}_c = \frac{A_C}{M}.$$

Once we have the normalized form of uncertainty for every country, \hat{A}_c , we then map the normalized uncertainty to a radial displacement r_c for each country C in our application with a scaling factor of 9 pixels, yielding radial displacements

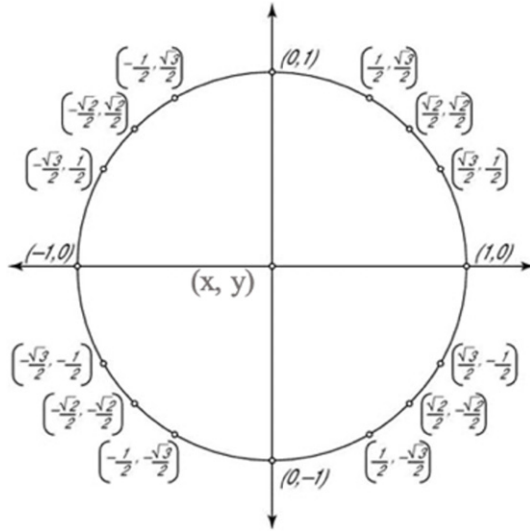


Figure 3. Underlying geometry of CA.

of 0 to 9 pixels. For example, countries that have higher uncertainties might have normalized uncertainty values \hat{A}_c such as 1, 0.9, and 0.8, which map to normalized radial displacements r_c of 9, 8.1, and 6.4, respectively.

Our guiding principle for mapping the values is to be as fair as possible to the competing method in our study (VSUP), and so we mapped the data values themselves to the eight color levels used in the pie-shaped VSUP color map along the top. So although this constraint of VSUP is somewhat arbitrary for our approach, we again normalized the data values and then scaled up to a range suitable for eight color levels. In the user study itself, we felt it more important to be as fair and consistent as possible between the two methods rather than express the absolute numerical values.

4. DESIGN OF CA VISUALIZATION

Chromatic aberration occurs when light of different wavelengths does not focus to the same convergent point because light with shorter wavelengths refracts more than light with longer wavelengths. Inspired by this phenomenon, we can consider a circle that represents the predicted number of new cases for a country on a specific day. But since there is an associated uncertainty with each prediction, a single circle will not be sufficient to represent the bivariate (number of cases and uncertainty) distribution. Instead of a single circle, we use three circles with separated RGB color channels, applying lateral shifts from the center of the circle by the amount of uncertainty, and blend them together in the center. Figure 3 shows this geometric arrangement on a unit radius circle.

To draw a circle representing uncertainty, we can draw three chromatic circles. We first set the center of the target circle at (x, y) and let the radial offset r represent the uncertainty. This can be implemented using many available visualization toolkits. For example, in the widely used D3 (d3js.org) library, this can be easily achieved with a

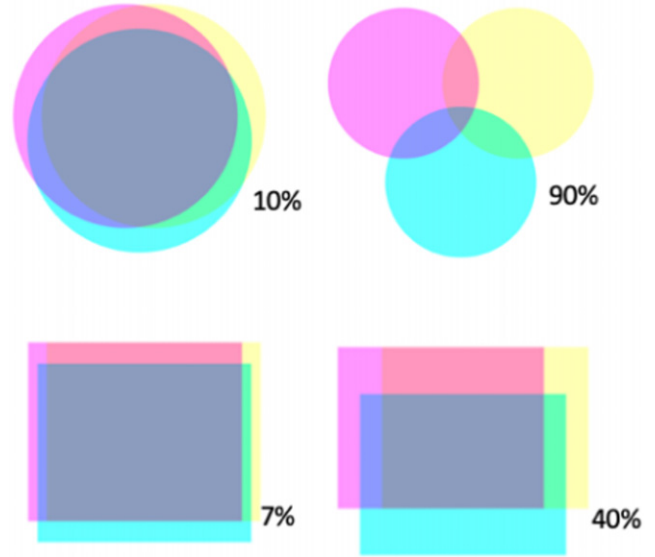


Figure 4. CA representation on circles (top) and rectangles (bottom).

commonly available blending mode (such as the CSS mix-blend-mode “darken”) to blend all three circles. Using the normalized radial displacements r_c calculated in Section 3.2, we then draw the first chromatic circle C_1 with color (R, 255, 255) at a shifted location of

$$(x, y + r_c),$$

the second chromatic circle C_2 with color (255, G, 255) at a shifted location of

$$\left(x + r_c \times \frac{+\sqrt{3}}{2}, \quad y + r_c \times \frac{-1}{2} \right),$$

and the third chromatic circle C_3 with color (255, 255, B) at a shifted location of

$$\left(x + r_c \times \frac{-\sqrt{3}}{2}, \quad y + r_c \times \frac{+1}{2} \right).$$

Using this approach, the resultant aberration is presented with the uncertainty for a country in Figure 4. The center area’s color represents the predicted number of new cases, and the color separated edges represent the amount of uncertainty in that prediction. In this way, each of the items in the figure represents both the predicted value and the amount of uncertainties.

5. USER STUDY DESIGN

Uncertainty visualization presents a complex challenge that requires careful design of user studies. Various study types include experimental, descriptive, observational, and within-/between-subject studies. Our research involved a within-subject comparative study evaluating uncertainty visualizations using metrics such as task time, error rate, and subjective assessments (NASA-TLX, System Usability

Scale [SUS]). We aim to understand how well chromatic aberration performs compared to the state-of-the-art VSUP visualization in terms of accuracy, efficiency (user response times), and user preference. This helps in assessing how well the visualization supports the intended tasks and identifying any additional needs or requirements.

5.1 Primary Study Components

Beyond the initial testing, Component Questions and Post-Session Questionnaires (PSQs) are the two main categories of question for the participants. Component Questions requires the user to select visualization elements defined by both the data value and the degree of uncertainty. As already noted, VSUP is a recent and prominent technique for uncertainty visualization. In that paper, they used a grid-chart representation, but in our study we broadened the test cases somewhat using both a grid chart and a circle chart, which form the following core components of our study:

- CA + Circle: Chromatic aberration is applied on circles in a circle chart.
- CA + Grid: Chromatic aberration is applied on squares in a grid chart.
- VSUP + Circle: Uncertainties are presented using color with circular shapes.
- VSUP + Grid: Uncertainties are presented using color with square shapes.

The two chart representations and two uncertainty visualizations are implemented in four different components.

The comparator method, VSUP, makes use of a grid chart with its custom color set. To make the comparison fair, we have grouped our uncertainties into four levels since VSUP also uses four levels of uncertainties. In our case, we have quantized our CA data and made four equidistant values of uncertainties such as 33, 52, 71, and 90 to represent the chromatic aberration in both circles and rectangles. In addition, to fill the circles and rectangles of CA, we have used the eight standard VSUP colors to make the evaluation consistent.

5.2 Recruitment

We recruited 32 participants to ensure balanced representation for all components. Most participants (97%) were undergraduate or graduate students with a background in computer science or information and communication technology. To ensure accurate results, we conducted color blindness tests and required participants to have a reliable Internet connection and a computer as the study was conducted entirely online.

5.3 Counterbalancing

Many empirical evaluations of input devices or interaction techniques involve comparing a new device or technique against alternatives. Our study used a within-subject design, allowing each participant to test every component while addressing potential learning effects. Each component includes eight questions, presented in random order to each



Figure 5. Possible flow of questionnaires for a participant.

participant. We employed the balanced Latin squares method to counterbalance the presentation of components, ensuring that no two consecutive participants received the same order and that each component was presented first to an equal number of participants (8 out of 32). This approach helps mitigate learning effects and ensures balanced emphasis across components.

5.4 Study Procedure

The study session contained several stages that included a color blindness test, introductions before starting a session, sessions for each of the core components, and Post-Session Questionnaires. In the following, we explain each of the stages. Over the course of the interactive sessions, the following data was collected:

- Answers to the questionnaire questions
- Videos (screen only)
- Audio recordings of the session.

Timing information was also recorded to facilitate a comparison of the time requirements of each competing visualization approach.

Since the study was conducted online during Covid, an event was created through the online meeting platform MS Teams. When the participant logged in to the system, the researcher greeted and welcomed the participant and exchanged formal greetings. If the participant faced any technical difficulty, then the researcher tried to help by any possible means. The researcher then briefed the participant about the steps he or she had to go through and explained how he was going to conduct the session. Participants were also asked if his/her system had a Firefox/Edge browser installed, which is mandatory for the study. If not, he/she would be requested to install it and the researcher might instruct further if they needed any help regarding the installation of the browser. After this, the participant was requested to open it and informed that the researcher would give two URLs for the session: (i) for the color blindness test and (ii) for the questionnaire about the study.

To maintain similarity with the work of Correll et al. [12], a color blindness/vision test was first administered to ensure participants were capable of discerning color accurately. Specifically, we presented a set of Ishihara plates [22] as a color blindness test, and participants were asked to detect the embedded numbers. Participants who failed to pass the color vision test were politely asked to withdraw from the study.

After the color blindness test was passed successfully, the researcher asked participants some basic questions, which we thought might be relevant to their performance. For instance, the following information was noted by the main researcher:

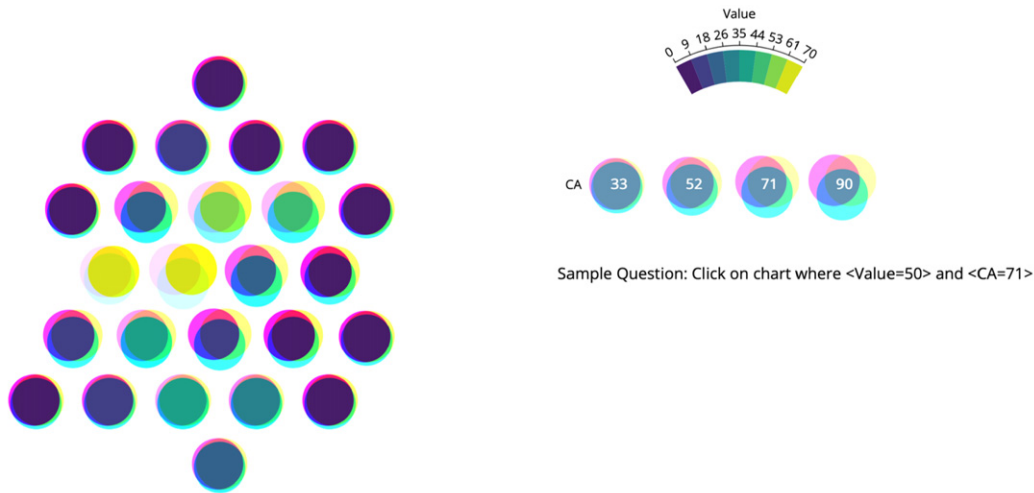


Figure 6. Sample Question. The number inside CA bubbles denotes the amount of uncertainty represented by those chromatic circles; e.g., CA = 71 refers to 71% uncertainty.

- Educational (science, arts, etc.) background
- Professional background (IT, accountant, etc.)
- Computer skills (basic, intermediate, expert)
- Mathematical and geometric knowledge
- Visualization and computer graphics knowledge
- Computer gaming skill
- Measurement knowledge (inch, feet, pixel, etc.).

After the pre-session discussion, there were two types of questions posed in our user study design: Component Questions and PSQs. By Component Questions, we refer to the questions relevant to those four core components. On the other hand, PSQs refer to the questions to obtain user feedback from the experience of using the four core components of the system. The PSQs include SUS questions and NASA-TLX Workload Scale questions.

If we consider *A*, *B*, *C*, and *D* as four components of the study, then Figure 5 shows a possible flow of the components that come one after another randomly during the session of the participants as discussed in Section 5.3. It also shows that PSQs appear at the completion of four modules.

At the beginning of every section, the bottom-right part of the user interface (UI) showed the session description. The researcher described the features (chart, legend, how questions will be asked, etc.). After completion of the explanation, the participant was asked to click the “Start” button when ready. Once he or she pressed the “Start” button, the questionnaire began immediately and questions were presented one at a time.

When presented with a question, the user needed to select a cell (bubble or rectangle) from the chart based on the provided Value and Uncertainty/CA combination. An example question is shown in Figure 6. After a cell was selected by the user, the next question appeared at the same place until the eighth question of the section was reached. We presented one example question prior to the questionnaire of each section. In the actual sessions, it was also described

verbally to the participant along with the opportunity for the participant to ask any questions they may have had.

The order of the questionnaire was changed by counterbalancing for individual participants. And so, this could be one possible sequence for a particular user:

1. Example of CA + Bubble
2. Questionnaire on CA + Bubble
3. Example of VSUP + Bubble
4. Questionnaire on VSUP + Bubble
5. Example of CA + Grid
6. Questionnaire on CA + Grid
7. Example of VSUP + Grid
8. Questionnaire on VSUP + Grid.

Then we asked them to answer the following two types of additional questionnaires:

9. Questions on SUS
10. Questions on NASA-TLX.

5.5 Component Questions

We now present a sampling of questions that were presented to the user, with additional explanatory information placed within them. Figs. 3 and 4 show example component questions for the CA + Circle and VSUP + Circle modules with the addition of explanatory markups for better understanding. We note that in the study session, the markups were not shown since the primary researcher clarified the underlying mechanism to the participants and/or answered any question the participants had before and during the session. There are significant commonalities in the two figures. Both present the following:

- The clickable chart in the left area
- A legend for both value and uncertainty in the top right
- The question asked of the participant (select a value & uncertainty level) in the bottom right.

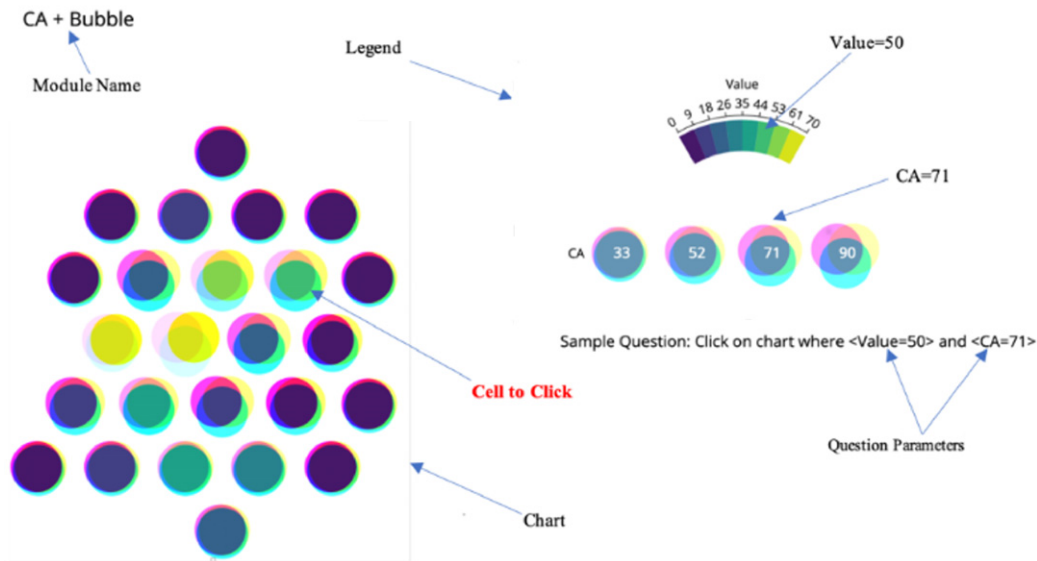


Figure 7. CA+Circle questionnaire interface.

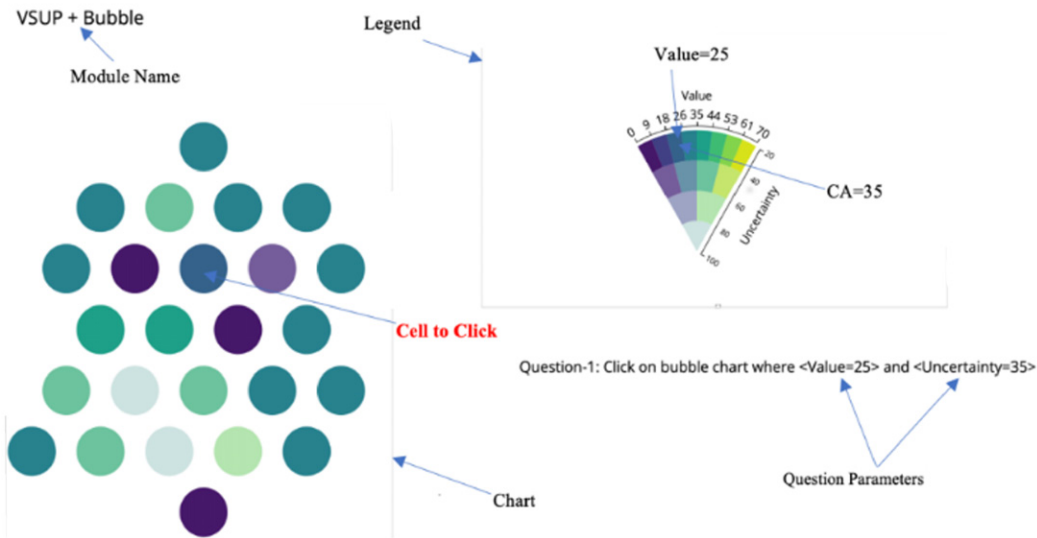


Figure 8. VSUP+Circle questionnaire interface.

Apart from the uncertainty method itself, the main point of difference between Figures 7 and 8 is found in the legend. The CA requires a composite legend, with colors shown for value and levels of CA for uncertainty. The VSUP's legend is more integrated since both value and uncertainty are mapped to a pie-shaped color map. Note, however, that the top row of colors used for value in both CA and VSUP was intentionally made the same to facilitate comparison. The VSUP legend is pie-shaped because data points with higher levels of uncertainty are mapped to fewer colors under the assumption that higher uncertainty yields a reduced discrimination of values.

Figures 9 and 10 show example component questions for the CA+Grid and VSUP+Grid modules including the added explanatory markups for better legibility. We note that apart

from the circles of the circle chart rather than the squares of the grid chart, the figures are very similar to Figs. 7 and 8.

Figures 11 and 12 show the PSQ interfaces for the NASA-TLX Workload related questions and the SUS questions. We separated both UIs in the figure where CA and VSUP occupy the top and bottom, respectively. Since the mechanism is the same for both CA+Circle and CA+Grid, they are grouped together and placed at the top of the UI in the CA section. Similarly, VSUP+Circle and VSUP+Grid are grouped together for the same reason and placed at the bottom in the VSUP section of the UI. For both CA and VSUP, we have shown the same question. For SUS, the questions are in the scale range of 1 to 5, where 1 means "Strongly Disagree" and 5 means "Strongly Agree" and the rest of the scales 2, 3, and 4 carry in between weights based

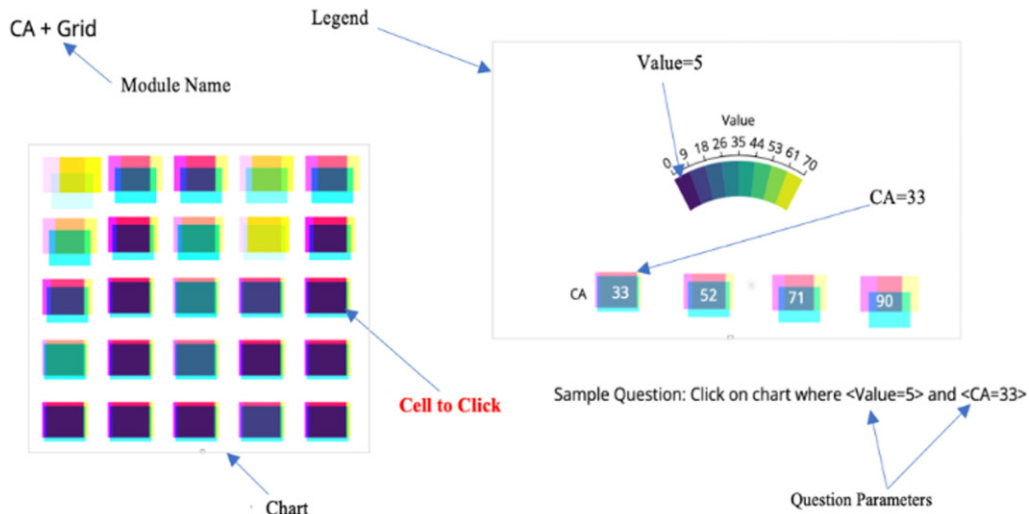


Figure 9. CA+Grid questionnaire interface.

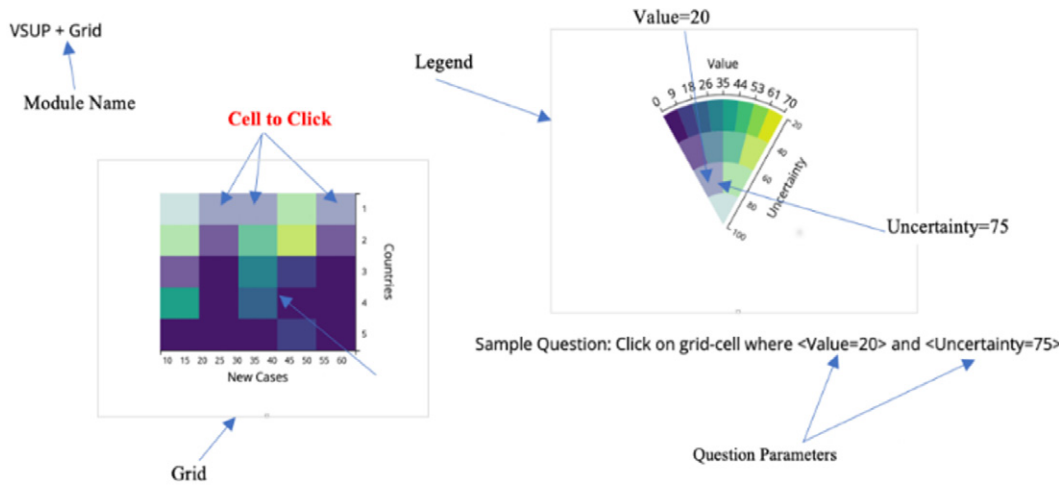


Figure 10. VSUP+Grid questionnaire interface.

on their values whereas for NASA-TLX, it is on a scale of 1 (Very Low) to 21 (Very High).

6. RESULTS AND NUMERICAL ANALYSIS

We recorded several sources of data from the user study, which include (i) quantitative questionnaire results; (ii) time utilization data for each component; (iii) SUS data for CA and VSUP; and (iv) NASA-TLX for CA and VSUP. We analyzed these data and present the results in the following sections.

6.1 Quantitative Questionnaire Results

As we have four components (CA + Circle, CA + Grid, VSUP + Circle, VSUP + Grid), we collected the performance data for each component separately. As previously stated, there were eight questions for each component and every question carried 1 point. For answering correctly, the participant gained 1 point and did not lose any points for incorrect answers. Our 32 participants could therefore gain a

maximum of 8 points for a component. Figure 13 graphically shows the correct response scores for both the circle chart and the grid chart for all 32 participants, where higher scores are better. As we will see in the subsequent analysis, when using chromatic aberration, users scored approximately 10% better on average. We note that this is not uniformly so, as there are a few instances of users performing better with VSUP.

We analyzed user performance (accuracy) with ANOVA for the four components and we subsequently used the t -test for two grouped (CA and VSUP) components. We first define the null (H_0) and alternative hypotheses (H_a) as follows:

H_0 : $\mu_1 = \mu_2 = \mu_3 = \mu_4$ (user accuracy was equal for all components)

H_a : Not all means are equal (user accuracy was not equal for all components).

Specifically, these hypotheses are tested using an F -ratio for one-way ANOVA. We obtained the test results shown

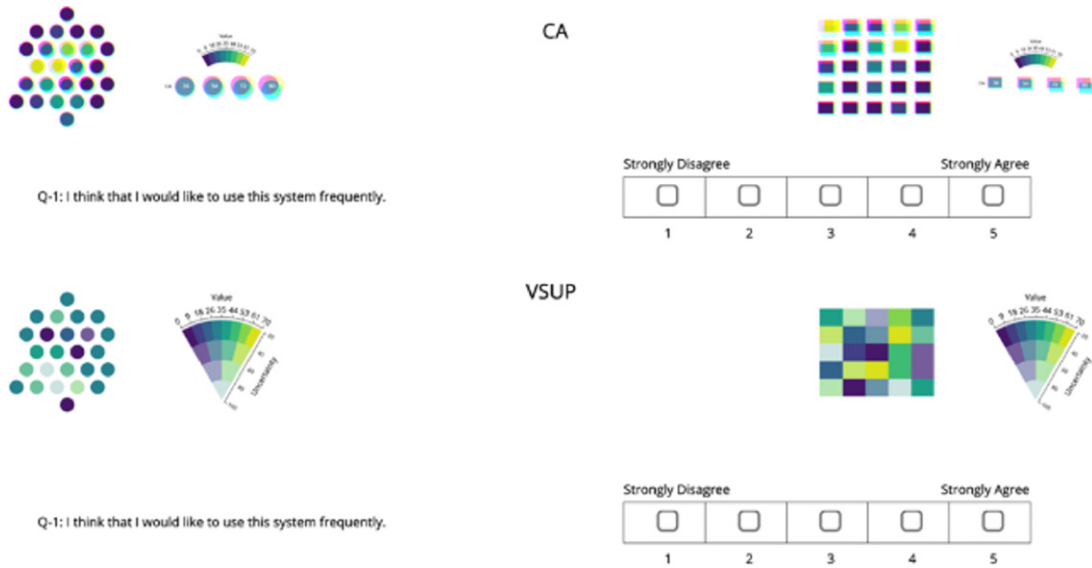


Figure 11. PSQ interface for SUS.

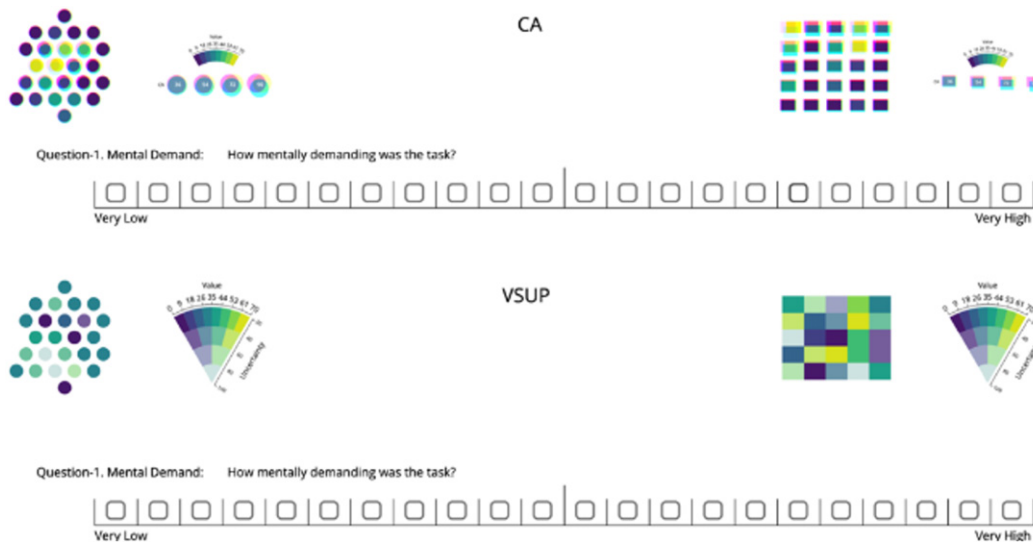


Figure 12. PSQ interface for NASA-TLX.

Table I. Summary of ANOVA test results.

Source	DoF	Sum of squares	Mean square	F-stat	p-value
Between groups	3	19.588	6.529	3.850	0.0113
Within groups	124	210.285	1.696		
Total	127	229.873			

in Table I for the significance level $\alpha = 0.05$, and the degrees of freedom are $df1 = 3$ and $df2 = 124$. Therefore, the rejection region for this F -test is $R = \{F: F > 2.678\}$ and the computed test statistic F equals 3.8499, which is not in the 95% region of acceptance: $[-\infty; 2.678]$. Given that $F = 3.85 > F_c = 2.678$, it is concluded that the null

hypothesis is rejected at the $\alpha = 0.05$ significance level, and the user accuracy was not equal for all components.

We then compare the combined CA and VSUP data from the four components' performance data by grouping the two pairs $CA = (CA+Circle \text{ and } CA+Grid)$ and $VSUP = (VSUP+Circle \text{ and } VSUP+Grid)$. The statistical summary of user accuracy performance is CA (mean = 5.938, SD = 1.105, SEM = 0.195, $N = 32$) and VSUP (mean = 5.422, SD = 1.078, SEM = 0.191, $N = 32$). Here we define the null (H_0) and alternative hypotheses (H_a) as follows:

$$H_0: \mu D = (\mu_1 - \mu_2) \geq 0 \text{ (performance of CA is higher or equal to the performance of VSUP)}$$

$$H_a: \mu D = (\mu_1 - \mu_2) < 0 \text{ (performance of CA is less than the performance of VSUP).}$$

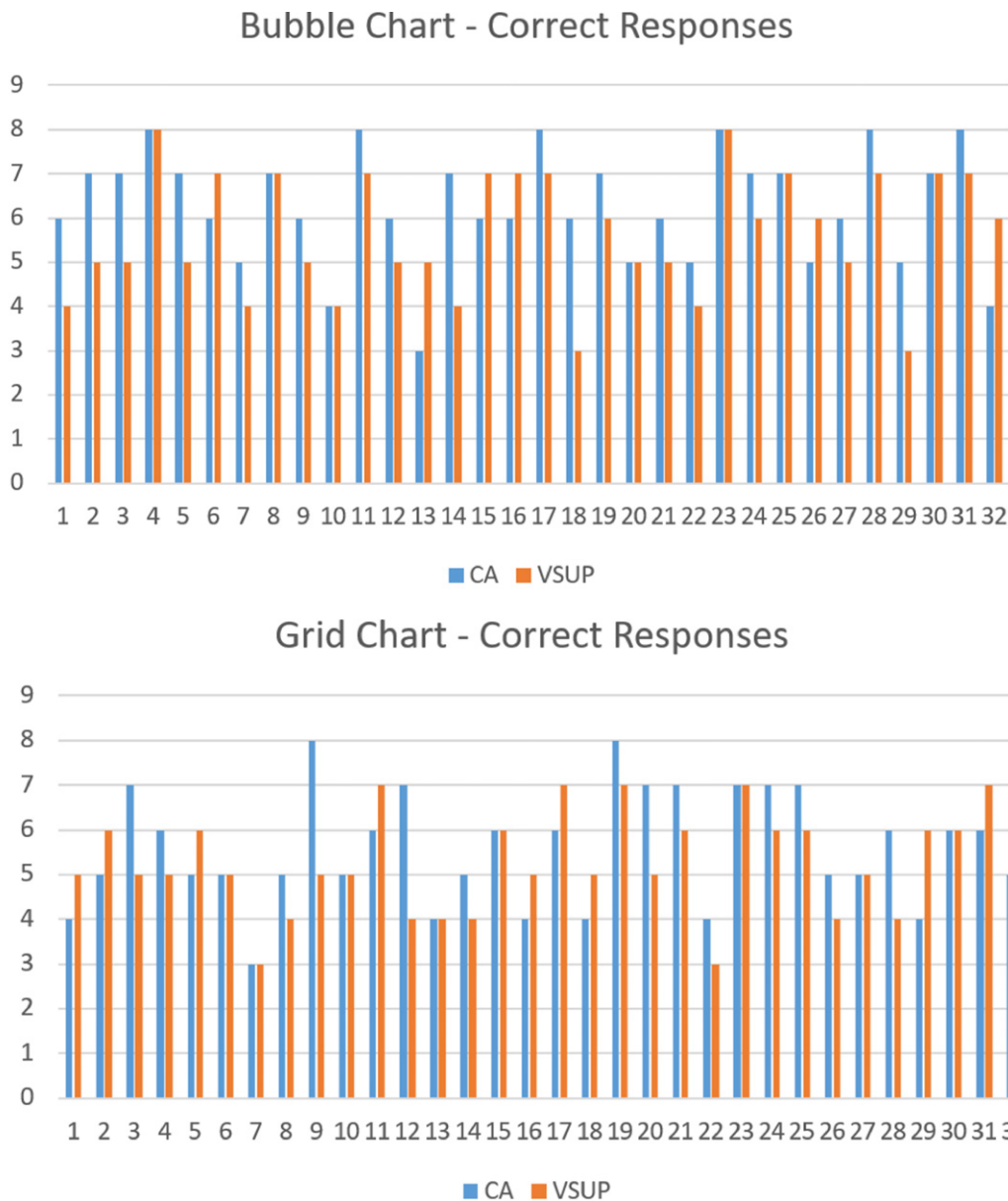


Figure 13. User accuracy scores for the circle chart (top) and the grid chart (bottom) when using chromatic aberration (blue) and value-suppressing uncertainty palettes (orange). Higher scores are better.

This corresponds to a left-tailed test, for which a t -test for paired samples is used. Using a significance level of $\alpha = 0.05$, the critical value for a left-tailed test is $t_c = -1.696$ and the rejection region for this left-tailed test is $R = \{t : t < -1.696\}$. The computed test statistic is 3.61 and since it is observed that $t = 3.61 \geq t_c = -1.696$, it is concluded that *the null hypothesis is not rejected*. We can say that the accuracy performance of CA quantitatively surpassed the performance of VSUP.

Our system also tracked the user response times for every component. Figure 14 graphically shows the user response times for tasks completed for both the circle chart and the grid chart for all 32 participants, where lower response times are better. As we will see in the subsequent

analysis, when using chromatic aberration, users were able to complete the tasks approximately 11% faster on average. We also note anecdotally that some individuals have very different response times with CA versus VSUP. In particular, the chart exhibits some markedly longer VSUP response times for some users. However, the results are not uniform as there are some instances of users performing faster with VSUP.

The statistical summary of the timing data is CA (mean = 8.675, SD = 2.320, SEM = 0.410, $N = 32$) and VSUP (mean = 9.647, SD = 3.123, SEM = 0.552, $N = 32$), where a shorter response duration is preferred. For user response times, we define the null (H_0) and alternative hypotheses (H_a) as follows:

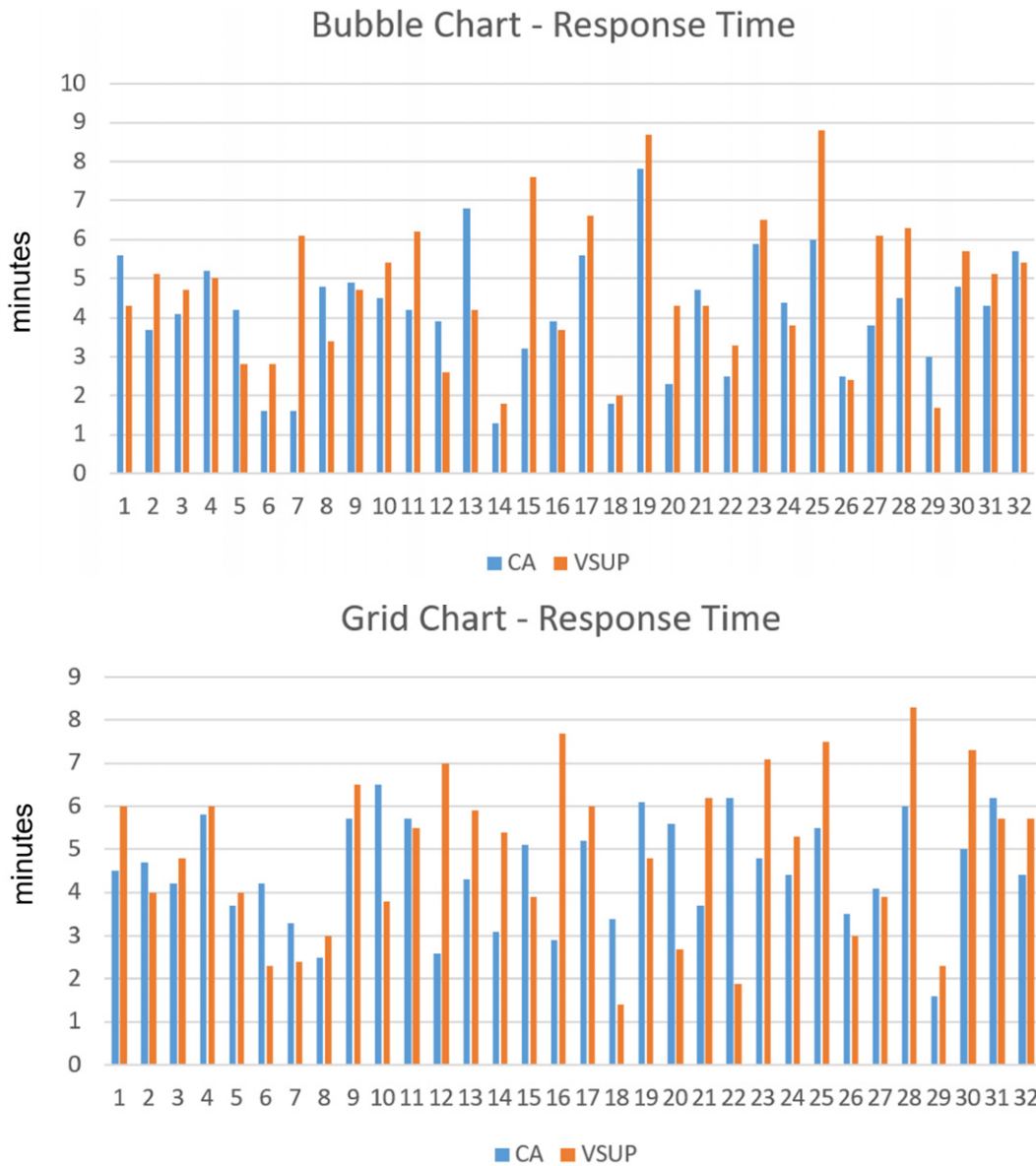


Figure 14. User response times for the circle chart (top) and the grid chart (bottom) when using chromatic aberration (blue) and value-suppressing uncertainty palettes (orange). Lower response times are better.

$H_0: \mu D = (\mu_1 - \mu_2) \leq 0$ (CA response was equal to or faster than VSUP response)

$H_a: \mu D = (\mu_1 - \mu_2) > 0$ (CA response was slower than VSUP response).

This corresponds to a right-tailed test, for which a t -test for paired samples is used. Again, using a significance level of $\alpha = 0.05$, the critical value for a right-tailed test is $t_c = 1.696$ and the rejection region for this right-tailed test is $R = \{t : t > 1.696\}$. The computed test statistic is equal to -2.656 . Since it is observed that $t = -2.656 \leq t_c = 1.696$, it is then concluded that *the null hypothesis is not rejected*. We can say that the response time performance of the CA method was faster than the VSUP method.

6.2 Qualitative Results

The SUS test provides a useful tool for measuring the usability of systems based on subjective user experience [6]. It consists of a ten-item questionnaire with five scale responses from participants from Strongly Agree to Strongly Disagree, which classifies the ease of use of the system being tested. Figure 15 (top) graphically shows the average SUS scores for all ten questions:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.

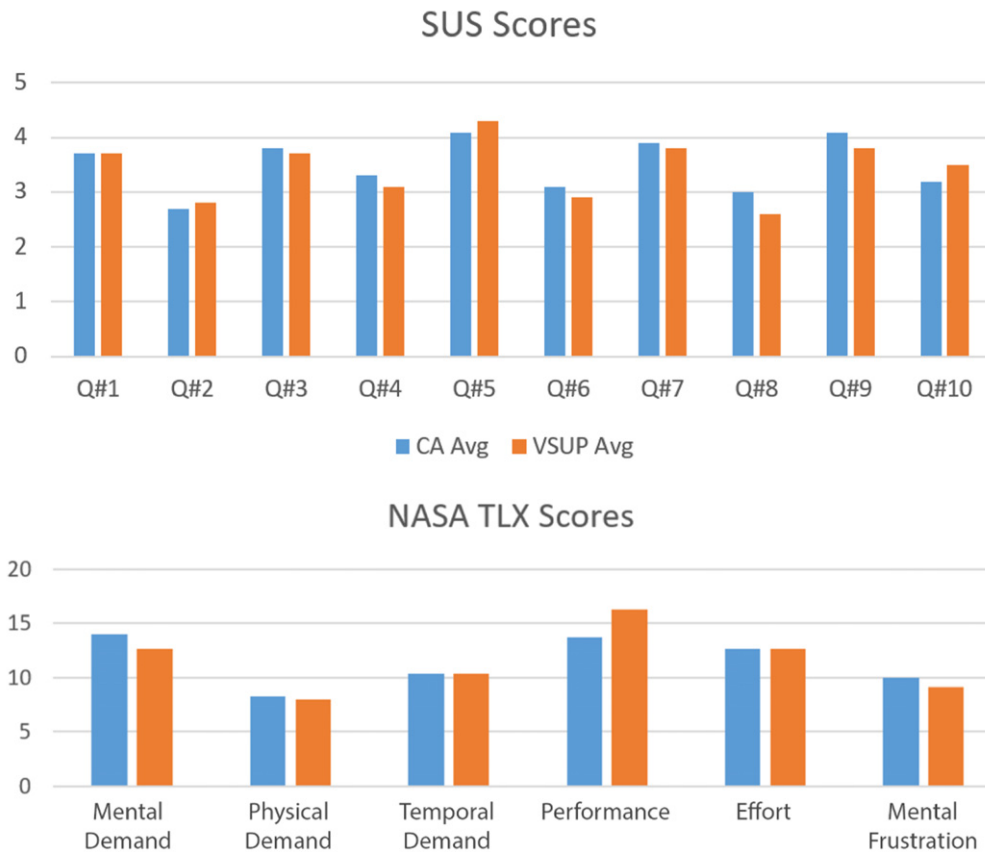


Figure 15. System Usability Scale (top) and NASA-TLX (bottom) scores recorded for chromatic aberration (blue) and value-suppressing uncertainty palettes (orange).

6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

As can be seen from the chart, there is a general pattern of consistency between the scores for the two methods of visualization with only minor variations, which is reflected in the following analysis, which shows a lack of statistically significant difference.

We interpret the results by normalizing the scores to produce a percentile ranking. By convention of SUS scoring, based on Sauro [44], we converted SUS results to SUS scores by the following rules:

1. For odd-numbered items: subtract 1 from the user response.
2. For even-numbered items: subtract the user responses from 5. This scales all values from 0 to 4 (with 4 being the most positive response).
3. Add the converted responses for each user and multiply that total by 2.5. This converts the range of possible values to a range from 0 to 100.

The statistical overview of the scores is CA (mean = 60.078, SD = 16.307, SEM = 2.883, $N = 32$) and VSUP (mean = 61.094, SD = 14.227, SEM = 2.515, $N = 32$). The Shapiro–Wilk tests on both distributions showed that they do not meet the normality test for CA ($W(32) = 0.913$, $p = 0.013$) and VSUP ($W(32) = 0.889$, $p = 0.003$). We therefore used the Kruskal–Wallis test on the data, which is a non-parametric alternative to the paired t -test since the distributions are not normal. The purpose of the test is to assess whether the samples come from populations with the same population median. We define the null (H_0) and alternative hypotheses (H_a) as follows:

H_0 : The samples come from populations with equal medians.

H_a : The samples come from populations with medians that are not all equal.

Using a significance level of $\alpha = 0.05$, and with the number of degrees of freedom equal to $df = 2 - 1 = 1$, the rejection region for this chi-square test is $R = \{\chi^2 : \chi^2 > 3.841\}$. The computed test (H) statistic is 0.146 and since $\chi^2 = 0.146 \leq 3.841$, it is concluded that the null hypothesis is not rejected. This implies that although the subjective SUS scores of the two methods varied slightly in our experiment, the differences were not statistically significant as per the Kruskal–Wallis test at $\alpha = 0.05$.

Table II. Kruskal–Wallis test results of NASA-TLX.

NASA-TLX	p	H	Result
Mental demand	0.6626	0.190	Not rejected
Within groups	0.8038	0.062	Not rejected
Temporal demand	0.8932	0.018	Not rejected
Performance	0.0574	3.610	Not rejected
Effort	0.8038	0.062	Not rejected
Mental frustration	0.6772	0.173	Not rejected

TLX stands for Task Load Index and is a measure of perceived workload [37]. As with the SUS test, we have also collected the NASA-TLX data from our system. TLX uses increments of high, medium, and low for each point resulting in 21 gradations on the scales. To compute the final scores, we subtract 1 from the given rating in the range of 1–21 and multiply by 5. Fig. 15 (bottom) graphically shows the average TLX scores for all six questions. As can be seen from the chart, there is once again a general pattern of consistency between the scores for the two methods of visualization with only minor variations.

To begin the analysis of this data, we applied the Shapiro–Wilk normality test at $\alpha = 0.05$. We determine the test results as follows:

CA: mental demand ($W = 0.906$, p -value = 0.009), physical demand ($W = 0.914$, p -value = 0.014), temporal demand ($W = 0.948$, p -value = 0.128), performance ($W = 0.948$, p -value = 0.014), effort ($W = 0.942$, p -value = 0.085), mental frustration ($W = 0.916$, p -value = 0.017).

VSUP: mental demand ($W = 0.863$, p -value = 0.001), physical demand ($W = 0.903$, p -value = 0.007), temporal demand ($W = 0.938$, p -value = 0.067), performance ($W = 0.887$, p -value = 0.003), effort ($W = 0.901$, p -value = 0.006), mental frustration ($W = 0.877$, p -value = 0.002).

Other than temporal demand, none of the perceived workload groups were found to be normal in distribution. Hence, we again use the Kruskal–Wallis non-parametric test to evaluate the differences between the two methods of uncertainty representation (CA and VSUP) for the NASA-TLX ratings provided by the participants. We define the null (H_0) and alternative hypotheses (H_a) as follows:

H_0 : The samples come from populations with equal medians.

H_a : The samples come from populations with medians that are not all equal.

Table II shows the summary of the test results at the $\alpha = 0.05$ significance level, with $df = 1$ and $\chi^2 = 3.841$.

As shown in Table II, since none of the perceived workload H values exceeded the χ^2 values, no statistically

significant differences were found at $\alpha = 0.05$ in the subjective NASA-TLX test.

Although participants did not offer many verbal comments, we note those that were made during the experiment. Participants (4, 21) commented that the “CA representation is deterministically difficult,” but we also noted that in these cases, the comment was the opposite of their performance given that they performed better in CA than VSUP. It is noteworthy nonetheless. Some other participants (19, 24) made a more nuanced comment, stating that “CA representation is complex but gives more confidence to find target.” Another comment that was commonly expressed by participants (14, 25, 31) is that “Colors are very close in VSUP which made them puzzled to select target.”

7. LIMITATIONS

There are several limitations of this work that we wish to highlight. Although we did not have a prerequisite for participants to be university students, based on the responses we received, most were from universities (undergraduate and graduate students) with only a subset of the population being working professionals. Based on this, one must be cautious not to generalize to significantly younger or older demographics when considering issues such as fatigue. We believe the prior VSUP study [12] also had participants with similar backgrounds.

Related to the sample demographics, in order to maintain coherence with the prior work of Correll et al. [12], and more generally, based on standard practice in such experiments, users with color blindness did not participate in the experiment. However, with regard to color vision impairment, a point of note is that although we did not design or test the visualization for color blind support, when compared with an alternative such as VSUP, which is purely based on very subtle differences of color, our approach at least offers a spatial encoding of the uncertainty information, namely when the R, G, and B elements separate. The limited discernible color palette of someone with limited color vision might be solely applied to the value, leaving uncertainty to be encoded with spatial separation. We speculate that because we are not trying to encode both uncertainty and the data values into color, our method may be more adaptable for color impaired applications than VSUP, but this would need to be tested with further studies.

However, if one considers other visual impairments such as myopia, presbyopia, or astigmatism, which affect the sharpness of the image, then other issues may come into play such as fatigue, eye strain, or general effectiveness. This also warrants further study. A general countermeasure worth mentioning is providing some choice for users with visual impairments, whether that be VSUP, CA, or CA with a limited color palette for value.

In addition, in the CA representation, one needs to be careful so that chromatic objects and adjacent objects do not overlap. An additional level of care must be applied in the case of implementing zooming based on the zoom scale of the visualization in order to keep them consistent. This relates

to the potential use of CA in geographic information system (GIS) applications, which we will return to in the following section. Another potential limitation is how CA may interact with other glyph shapes.

Lastly, although our approximation of CA is computationally inexpensive and very accessible for use in widely used web-based visualizations, if one were to implement a more complex CA rendering method, then further study with participants may be required. In particular, we note that in real-world chromatic aberration, chromatic blurring appears continuously from the inner edge to the outer edge.

8. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a novel approach for uncertainty visualization, namely chromatic aberration. We conducted a within-subject comparative user study between the state-of-the-art alternative, VSUP, and our system to assess user performance accuracy/error rate, task completion time, and subjective assessment (with NASA-TLX and SUS). From the numerical analysis of the results, we see that the user performance of CA is both statistically more accurate (10%) and faster (11%) when compared to VSUP, whereas in the subjective assessment, the two methods do not vary significantly. Future studies could look at further task-specific uses of CA, which might shed more light on the lack of statistical differences with regard to subjective preference.

We noted that in real chromatic aberration, blurring appears continuously from the inner edge to the outer edge. However, our simplified implementation allows us to reduce the aberration to both double and/or single parameter(s), which facilitates the representation of uncertainty. It also allows one to implement the approach relatively easily using standard and widely available graphical operations. However, additional research could be conducted that examines more sophisticated chromatic aberration effects. In addition, further research could be carried out with more levels of uncertainties than were tested by Correll et al. [12].

Related to this, one might look more closely at the mapping of uncertainty to the scale of CA displacements. Additional research would be required to determine the optimal or minimal displacement required to convey uncertainty. More broadly speaking, as this is the first attempt at using chromatic aberration for this novel purpose, there likely remains significant scope to optimize and adapt its use with further refinements and ideas.

Other possibilities include studying the use of CA in animated and interactive visualizations, using CA for contextual highlighting, and testing the effectiveness of CA for use with more complex glyphs, as well as the potential intersection between CA and known pre-attentive visual effects. Another aspect that could be explored is using CA with alternative data types such as categorical, hierarchical, and time series data, which might then relate more specifically to other existing work such as bubble treemaps [17], ggdist [26], and tidybays [21].

It would also be interesting to explore the limits of CA in applications requiring choropleth maps and other GIS-related visualizations. Despite the spatial requirement that the CA separation requires, a choropleth map might be possible using a buffer between each element to allow for some degree of chromatic aberration at the boundaries. Alternatively, the R, G, and B chromatic displacements for each element in the map might instead maintain the outer boundary shape as a maximum area. Then each R, G, and B displacement would need to shrink somewhat and move away from the center internally, within that original boundary of each map element. This may offer a stronger overall Gestalt of the uncertainty in the map, but the tradeoff might be a tighter limit on how small individual elements could become.

Beyond GIS maps, one might speculate on the breadth of applicability to other types of data, such as temporal data, and other forms of visualization. With some graphs that are defined by lines and curves, the CA would need to be adapted and applied locally along the line or curve, rather than uniformly displacing the entire curve into three renderings for R, G and B. Other graphs are primarily defined with their silhouettes, such as violin charts, density charts, and ridgeline charts. And as with curve and line-based charts, for these charts that provide most of their information along silhouettes, the CA would likely need to be applied locally along those boundary curves. Some other types of graphs, such as stream graphs, may often contain large flat areas of color and these might benefit from the overlay of an artificial texture prior to the local application of simulated CA. This might allow the CA to differentiate that texture locally based on uncertainty. But given that our method has purposely been designed for implementation in widely used online visualization libraries, such as D3, it is our hope that other practitioners begin to experiment with possibilities that we are as yet unable to anticipate.

REFERENCES

1. B. Auffarth, *Machine Learning for Time-Series with Python* (Packt Publishing, United Kingdom, 2021).
2. M. K. Beard, "Use error: The neglected error component," *Proc. AUTO-CARTO 9* (Baltimore, MD, 1989), pp. 808–817.
3. G. Bonneau, H. Hege, C. Johnson, M. M. Oliveira, K. Potter, P. Rheingans, and T. Schultz, "Overview and State-of-the-Art of Uncertainty Visualization," in *Scientific Visualization. Mathematics and Visualization*, edited by C. Hansen, M. Chen, C. Johnson, A. Kaufman, and H. Hagen (Springer, London, 2014).
4. R. P. Botchen, D. Weiskopf, and T. Ertl, "Texture-based visualization of uncertainty in flow fields," *Proc. IEEE Visualization* (IEEE, Piscataway, NJ, 2005), pp. 647–654.
5. T. E. Boulton and G. Wolberg, "Correcting chromatic aberrations using image warping," *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 1992), pp. 684–687.
6. K. Brodli, R. A. Osorio, and A. Lopes, "A review of uncertainty in data visualization," *Expanding the Frontiers of Visual Analytics and Visualization* (Springer, Berlin, 2012), pp. 81–109.
7. J. Brooke, "SUS: A quick and dirty usability scale," in *Usability Evaluation in Industry*, edited by P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland (Taylor and Francis, London, 2014).
8. J. Brownlee, "Deep learning models for time series forecasting," *Machine Learning Mastery* (2018).

- ⁹ B. Buttenfield and M. K. Beard, "Visualizing the quality of spatial information," *Proc. AUTO-CARTO 10* (Baltimore, MD, 1991), Vol. 6, pp. 423–427.
- ¹⁰ E. Cakmak, H. Schäfer, J. Buchmüller, J. Fuchs, T. Schreck, A. Jordan, and D. Keim, "MotionGlyphs: Visual abstraction of spatio-temporal networks in collective animal behavior," *Comput. Graph. Forum* **39**, 63–75 (2020).
- ¹¹ F. Chollet, "Keras: Deep Learning for Humans". Available at <https://github.com/fchollet/keras> (2015, accessed 2023).
- ¹² M. Correll, D. Moritz, and J. Heer, "Value-suppressing uncertainty palettes," *Proc. CHI Conf. on Human Factors in Computing Systems* (ACM, Montreal, Canada, 2018), pp. 1–11.
- ¹³ M. Correll and M. Gleicher, "Error bars considered harmful: Exploring alternate encodings for mean and error," *IEEE Trans. Vis. Comput. Graphics* **20**, 2142–2151 (2014).
- ¹⁴ S. Deitrick and R. Edsall, "The influence of uncertainty visualization on decision making: An empirical evaluation," *Progress in Spatial Data Handling* (Springer, Berlin, Heidelberg, 2006), pp. 719–738.
- ¹⁵ R. Finger and A. M. Bisantz, "Utilizing graphical formats to convey uncertainty in a decision-making task," *Theor. Issues Ergon. Sci.* **3**, 1–25 (2002).
- ¹⁶ E. Gecili, A. Ziady, and R. D. Szczeniak, "Forecasting COVID-19 confirmed cases, deaths and recoveries: Revisiting established time series modeling through novel applications for the USA and Italy," *PLoS One* **16** (2021).
- ¹⁷ J. Görtler, C. Schulz, D. Weiskopf, and O. Deussen, "Bubble treemaps for uncertainty visualization," *IEEE Trans. Vis. Comput. Graphics* **24**, 719–728 (2018).
- ¹⁸ M. Greis, P. Agroudy, H. Schuff, T. Machulla, and A. Schmidt, "Decision-making under uncertainty: How the amount of presented uncertainty influences user behavior," *Proc. Nordic Conf. on Human-Computer Interaction* (ACM, New York, NY, USA, 2016), pp. 1–4.
- ¹⁹ S. Greenberg and B. Buxton, "Usability evaluation considered harmful (some of the time)," *Proc. Conf. Human Factors in Computing Systems (CHI)* (ACM, Florence, Italy, 2008), pp. 111–120.
- ²⁰ H. Griethe and H. Schumann, "The visualization of uncertain data: Methods and problems," *Proc. SimVis* (Institute for Operations Research and the Management Sciences, Magdeburg, Germany, 2006), pp. 143–156.
- ²¹ W. Hadley, "Tidy data," *J. Stat. Softw.* **59**, 1–23 (2014).
- ²² L. H. Hardy, G. Rand, and M. C. Rittler, "Tests for detection and analysis of color blindness; An evaluation of the Ishihara test," *Arch Ophthal* **4**, 268–275 (1945).
- ²³ B. Hanington and B. Martin, *Universal Methods of Design* (Rockport Publishers, Beverly, MA, 2019).
- ²⁴ J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay, "In pursuit of error: A survey of uncertainty visualization evaluation," *IEEE Trans. Vis. Comput. Graphics* **25**, 903–913 (2019).
- ²⁵ M. K. Johnson and H. Farid, "Exposing digital forgeries through chromatic aberration," *Proc. Multimedia and Security* (ACM, New York, NY, USA, 2006), pp. 48–55.
- ²⁶ M. Kay, "ggdist: Visualizations of distributions and uncertainty in the grammar of graphics," *IEEE Trans. Vis. Comput. Graphics* **30**, 414–424 (2024).
- ²⁷ G. Keppel, *Design and Analysis: A Researcher's Handbook*, 4th ed. (Upper Saddle River, NJ, 2004).
- ²⁸ A. Kamal, P. Dhakal, A. Y. Javaid, V. K. Devabhaktuni, D. Kaur, J. Zaiantz, and R. Marinier, "Recent advances and challenges in uncertainty visualization: A survey," *J. Vis.* **24**, 1–30 (2021).
- ²⁹ M. Kay, T. Kola, J. R. Hullman, and S. A. Munson, "When (ish) is my bus?: User-centered visualizations of uncertainty in everyday, mobile predictive systems," *Proc. CHI Conf. on Human Factors in Computing Systems* (ACM, New York, NY, 2016), pp. 5092–5103.
- ³⁰ K. Koh, B. Kim, and J. Seo, "Effect of lateral chromatic aberration for chart reading in information visualization on display devices," *Advanced Visual Interfaces* (ACM, New York, NY, 2014), pp. 289–292.
- ³¹ M. Korporaal, I. T. Ruginski, and S. I. Fabrikant, "Effects of uncertainty visualization on map-based decision making under time pressure," *Human-Media Interaction* (Frontiers Media SA, Lausanne, Switzerland, 2020), Vol. 2.
- ³² H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale, "Empirical Studies in Information Visualization: Seven Scenarios," *IEEE Trans. Vis. Comput. Graphics* **18**, 1520–1536 (2012).
- ³³ S. Lee, E. Eisemann, and H.-P. Seidel, "Real-time lens blur effects and focus control," *ACM Trans. Graphics* **29**, 1–7 (2010).
- ³⁴ J. Leo, E. Luhanga, and K. Michael, "Machine learning model for imbalanced cholera dataset in Tanzania," *Sci. World J.* 1–12 (Wiley, Hoboken, NJ, 2019).
- ³⁵ L. R. Lucchesi and C. K. Wikle, "Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation," *Statistics* **6**, 292–302 (2017).
- ³⁶ C. Lundstrom, P. Ljung, A. Persson, and A. Ynnerman, "Uncertainty visualization in medical volume rendering using probabilistic animation," *IEEE Trans. Vis. Comput. Graphics* **13**, 1648–1655 (2007).
- ³⁷ S. Mach, J. P. Gründling, F. Schmalfuß, and J. F. Krems, "How to assess mental workload quick and easy at work: A method comparison," *Advances in Intelligent Systems and Computing* (Springer, United States, 2019), pp. 978–984.
- ³⁸ G. Miranda, J. Baetens, N. Bossuyt, O. Bruno, and B. Baets, "Real-time prediction of influenza outbreaks in Belgium," *Epidemics* (Elsevier, United Kingdom, 2019), 28.
- ³⁹ A. Pang, C. Wittenbrink, and S. Lodha, "Approaches to uncertainty visualization," *Vis. Comput.* **13**, 370–390 (1997).
- ⁴⁰ M. Panja, T. Chakraborty, S. Nadim, I. Ghosh, U. Kumar, and N. Liu, "An ensemble neural network approach to forecast Dengue outbreak based on climatic condition," *Chaos, Solitons & Fractals* **167** (2023).
- ⁴¹ M. Riveiro, "Evaluation of uncertainty visualization techniques for information fusion," *Int'l. Conf. Information Fusion* (IEEE, Piscataway, NJ, 2007), pp. 1–8.
- ⁴² C. Satrio, W. Darmawan, B. Unrica Nadia, and N. Hanafiah, "Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET," *Procedia Comput. Sci.* **179**, 524–532 (2021).
- ⁴³ J. Sauro and J. R. Lewis, "Quantifying the user experience: Practical statistics for user research," 2nd ed. (Elsevier, Amsterdam, Netherlands, 2016).
- ⁴⁴ J. Sauro and J. Lewis, *Quantifying the User Experience: Practical Statistics for User Research* (Morgan Kaufmann, Waltham, Massachusetts, 2012).
- ⁴⁵ M. Schneider, M. McDowell, P. Guttorp, E. A. Steel, and N. Fleischhut, "Effective uncertainty visualization for aftershock forecast maps," *Natural Hazards and Earth System Sciences* (European Geosciences Union, Munich, Germany, 2022), Vol. 22, pp. 1499–1518.
- ⁴⁶ X. Song, J. Xiao, J. Deng, Q. Kang, Y. Zhang, and J. Xu, "Time series analysis of influenza incidence in Chinese provinces from 2004 to 2011," *Medicine* (Baltimore) **26**, e3929 (2016).
- ⁴⁷ M. Srivenkatesh, "Performance evolution of different machine learning algorithms for prediction of liver disease," *Int. J. Innovative Technology and Exploring Engineering* **9**, 2278–3075 (2019).
- ⁴⁸ S. Weidman, "Deep learning from scratch: Building with python from first principles," O'Reilly Media, 1st edition, Oct. 15, 2019.
- ⁴⁹ C. M. Wittenbrink, A. T. Pang, and S. K. Lodha, "Glyphs for visualizing uncertainty in vector fields," *IEEE Trans. Vis. Comput. Graphics* **2**, 266–279 (1996).
- ⁵⁰ H. S. Yoo, "Color illusions on liquid crystal displays and design guidelines for information visualization," *Master of Science, Virginia Tech*, <http://hdl.handle.net/10919/36372>, December, 7, 2007.