

PERCEPTION IS REALITY: MEASURING THE EFFICACY
OF HUMAN-CENTERED DATA VISUALIZATION
METHODS WITH COVID-19 CASE DATA

by

STEPHANIE SCHOFIELD

A THESIS

Presented to the Department of Computer Information Science
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

June 2021

An Abstract of the Thesis of

Stephanie Schofield for the degree of Bachelor of Arts
in the Department of Arts and Sciences to be taken June 2021

Title: Perception is Reality: Measuring the Efficacy of Human-Centered Data
Visualization Methods with COVID-19 Case Data

Approved: Professor Joseph Sventek, Ph.D.
Primary Thesis Advisor

Data visualization is a tool used to represent the vast amount of information created each day. Although current literature about data visualization illustrates many ways to visualize data, only a few methods are used on a recurring basis. Common visualizations like bar charts, pie charts, and scatter plots are seen across the internet in news articles, scholarly works, and scientific papers. These methods, although prevalent and easy to produce, are not always created with the viewer in mind. They can be too complex, over-simplified, or misrepresent the data. This research aims to measure whether a human-centered approach to designing data visualizations could portray the information more effectively, and allow the user to understand the argument the data presents. More specifically, this study will measure whether a human-centered approach to visualizing COVID-19 case data will more effectively change behavior intentions with regards to social-distancing. Results of this research reflect that the human-centered approach did not produce higher rates of comprehension, visual appeal, or willingness to change behavior, and suggests that a more distinct and interactive method is necessary for meaningful differences in understanding of the information.

Acknowledgements

I would like to express my sincere gratitude to my thesis committee who have helped me in every step of this process--not only for their time and extreme patience, but for their intellectual contributions to my development as a computer scientist. Professor Joe Sventek's guidance, wisdom, and support, from our initial discussions about data visualization to now, have inspired me throughout this project. To Professor Nicole Dudukovic, I thank you for bringing your unique, balancing perspective as someone with rigorous experience in psychology, memory, and attention. To Professor Juan Flores, thank you for providing me invaluable guidance throughout the writing process.

I would like to thank my family for being the best support system I could ask for. Dad, thank you for inspiring my love of math and understanding how complex systems work, and sending me the article that influenced this project. Mom, thank you for taking time away from your own dissertation to advise me on mine. Sean, thank you for frequently interrupting my workflow by showing me funny Youtube videos. I would also like to thank all of my friends who offered words of encouragement, a hug, or a latte to support me while I wrote this thesis.

Completing a thesis as an undergraduate is a difficult endeavor, but I am proud to say my experience working on this project was intellectually stimulating, fun, and rewarding, thanks to the support of each of you. Thank you for having confidence in me and offering your guidance through this process.

Table of Contents

Chapter 1: Introduction	1
1.1 Motivation for this Research	1
1.2 Research Questions	3
1.3 Hypothesis	4
1.4 Outline of Thesis	4
Chapter 2: Literature Review	5
2.1 What is Data Visualization, and Why Do It?	5
2.2 Common Mistakes in the Default Design Process to Visualize Data	9
2.3 An Argument for a Human-Centered Approach	10
2.4 Current Methods in Human-Centered Design	11
Chapter 3: Methodology	13
3.1 Data Collection	13
3.2 Creating Visualizations	17
3.3 Designing the Survey	19
3.3.1 Structure and Question Design	22
3.3.2 Social Distancing and Mask Behavior Questions	23
3.3.3 Measuring Comprehension Questions	25
3.3.4 Visual Appeal Questions	28
3.3.5 Demographic Questions	28
3.4 Participant Criteria	29
3.5 Data Analysis	30
Chapter 4: Results	33
4.1 Participant Demographic	34
4.1 Overview of Entire Sample	34
4.2 Age	37
4.2.1 Ages 18-24	37
4.2.2 Ages 35+	38
4.3 Gender	39
4.3.1 Female Participants	39
4.3.2 Male Participants	40
4.4 Colorblindness	41

4.4.1 Colorblind Participants	42
4.4.2 Non-Colorblind Participants	43
Chapter 5: Discussion and Recommendations	45
5.1 Producing Effective Visuals	45
5.2 Changing Public Behavior	46
5.3 Future Research	47
Conclusion	50
Appendix	51
Glossary	51
Survey Materials	53
Question Breakdown	53
Recruitment Statement	66
Survey Information Statement	67
Charts	69
Debriefing Statement	73
Notice of Amendment Review and Exempt Determination	74
References	76

List of Figures

Figure 1.10	2
Figure 2.10	6
Figure 2.11	7
Figure 3.10	14
Figure 3.11	15
Figure 3.12	16
Figure 3.3.1: Human-Centered Daily Cases Visualization	20
Figure 3.3.2: Default Daily Cases Visualization	20
Figure 3.3.3: Human-Centered Cumulative Cases Visualization	21
Figure 3.3.4: Default Cumulative Cases Visualization	21
Figure 3.30	24
Figure 3.31	24
Figure 3.32	27
Figure 3.33	27
Figure 4.0: Question Legend	33
Figure 4.11	34
Figure SM1: Q62	53
Figure SM2: Q63	54
Figure SM3: Q97	54
Figure SM4: Q130	55
Figure SM6: Q98	57
Figure SM7: Q131	57
Figure SM8: Q50 / Q165 / Q122 / Q166	58
Figure SM9: Q135 / Q136	58
Figure SM10: Q134 / Q137	58
Figure SM11: Q66 / Q170 / Q124 / Q173	59
Figure SM12: Q145 / Q171 / Q146 / Q174	59
Figure SM13: Q139 / Q172 / Q138 / Q175	60
Figure SM16: Q71	61
Figure SM17: Q100	62
Figure SM18: Q142	62

Figure SM19: Q52	62
Figure SM21: Q54	63
Figure SM23: Q56	64
Figure SM24: Q57	65
Figure SM25: Q58	66
Figure SM26: Random ID	66
Figure C1: Human-Centered Daily Cases Visualization	71
Figure C2: Default Daily Cases Visualization	71
Figure C3: Human-Centered Cumulative Cases Visualization	72
Figure C4: Default Cumulative Cases Visualization	72
Figure N1: IRB Exempt Determination	74
Figure N2: IRB Amendment Review	75

Chapter 1: Introduction

1.1 Motivation for this Research

Data is the most abundant resource on earth. Everything we touch, hear, see, and breathe is information. In fact, we create about 2.5 exabytes of data online every day: 500 million tweets, 5 billion internet searches, and 294 billion emails are sent daily (Desjardins, 2019). With so much data accessible on the internet, it is critical that we analyze how this information is presented, and how it is perceived. Now in the face of a global pandemic, articles, images, and videos of the Coronavirus's development have spread across the world. Anxious people are seeking answers, but some of the visualizations in these publications misrepresent the data. Or worse, the data lack sufficient scientific support to be correct. In order for us to stop the spread of this virus, the public needs to understand the urgency and reasoning behind wearing a mask and practicing social-distancing. Understandable, unambiguous graphics are critical in order for the public to understand these ideas.

In April 2020, I left my job as a Resident Assistant in Eugene to fly home to Honolulu, Hawaii due to a pandemic-induced domestic travel ban placed on military dependents. At this time, very few people were wearing masks, socially-distancing, or practicing precautions that would signal an understanding of the seriousness of the virus. I wore two face masks and a shield to protect myself on the plane, but received many negative stares as I walked through the airport. Just before taking off, my dad sent me an article from the New York Times titled "This 3-D Simulation Shows Why Social

Distancing Is So Important” (Parshina-Kottas et al., 2020). This article features a visualization of how water droplets spread around an open room after a person sneezes.

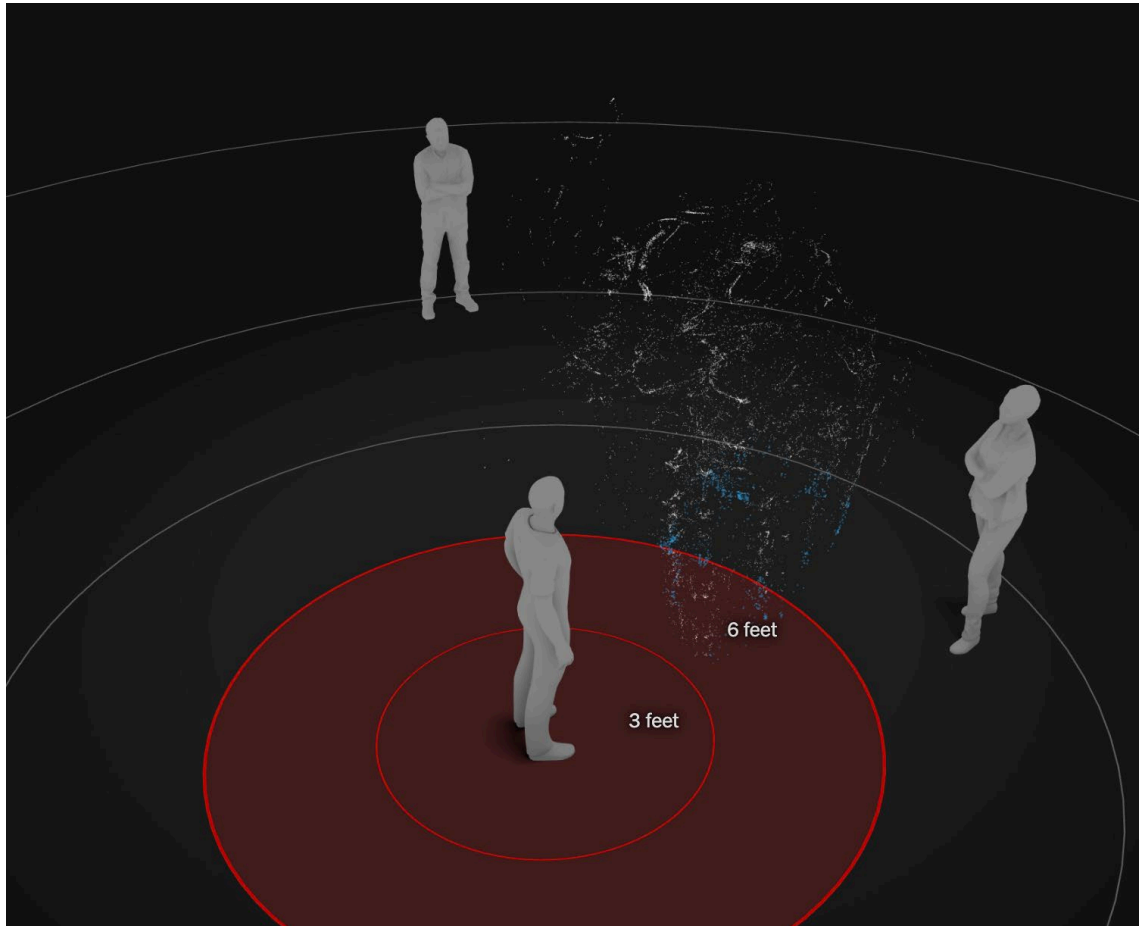


Figure 1.10

Image showing the progression of a person’s sneeze from three feet, six feet, and farther (Parshina-Kottas et al., 2020).

After looking at this graphic, I understood how the virus spreads, and was reminded how important it is to wear a mask to protect others. Although I had read articles about which protocols I should follow, this visualization made it clear why I need to be social distancing and masking. I found it easy to connect with this type of interactive visual, and see the science and data in a more comprehensible format. This is what inspired me to think about the way these visualizations are produced, and how

effective charts could motivate people to behave a certain way. I wondered: what if a chart representing COVID-19 data could compel someone to socially-distance? This is where I began my research and the development of this project.

1.2 Research Questions

In order to measure if data visualizations are truly compelling, I had to ask questions surrounding the approach to creating those visuals. I wanted to know how visualizations are created, what current literature proposes about these methods, and how to improve upon that methodology to make more effective data visualizations. The first step in the research process was to conduct a literature review. The review was motivated by a few important questions: Why use a human-centered design approach to creating visualizations? Can visualizations be clear, and convincing enough to change people's behavior after looking at them? What is the best way to present visualizations of data for maximum human-understanding? Addressing these questions in my literature review narrowed my focus for what I wanted answered in my study.

I determined that there were questions to be answered about the effectiveness of data visualization methods, and areas of public health and data science where I could contribute this research. My study addressed: Which method of visualizing COVID-19 case data will result in the highest number of responses showing a willingness to continue social-distancing and mask-wearing behavior, or change their behavior if not doing so? Once I completed secondary research and formulated my research questions, I had enough context to make a prediction about the results of this study.

1.3 Hypothesis

My hypothesis is that participants who view COVID-19 case data visualizations designed with a human-centered approach will show the highest rate of changing their current social distancing and mask-wearing behavior. Additionally, I predict that human-centered visualizations will be preferred among participants, and be understood more clearly than my default charts.

1.4 Outline of Thesis

I start by giving context to current methods in data visualization, human-centered design, and how data visualization affects public health. My literature review contains the default practices for designing visualizations, and an argument for why human-centered design is critical to creating effective visualizations. I will then provide my research methodology in finding the data sets, generating the visualizations, designing the survey, recruiting participants, and analyzing the data. Then I will share the findings of this study based on several key participant demographics, followed by a discussion of the results and future work in this research.

Chapter 2: Literature Review

I start this chapter by introducing significant applications of data visualization, the motivation behind portraying data as comprehensible graphics, and then I explore current methods for producing visualizations. I will make an argument for why data visualizations need to be created with human-comprehension at the forefront of design. Furthermore, I will explain why compelling visualizations are so critical to public health.

2.1 What is Data Visualization, and Why Do It?

Tamara Munzner, an expert in information visualization and author of *Visualization Analysis & Design*, defines computer-based visualizations as models that provide visual representations of datasets designed to help people carry out tasks more effectively (Munzner, 2015). To be more concise, data visualization is modeling large amounts of data in a visual representation. Munzner argues that visualization is most necessary when one desires to enhance human capabilities rather than replace them with computation (Munzner, 2015). This means that visualizations should help people understand complex ideas without having to understand the computational analysis behind it.

There are many areas of research that benefit from human analysis of data visualizations. Economists, physicians, policy makers, geologists, mathematicians, historians, musicians, and many other industry professionals and creatives alike utilize data modeling. The website visualcapitalist.com features many applications of data visualization across technology, money, energy, healthcare, and more. One recent post shows a unique visualization of the richest families in America, as seen in figure 2.10.

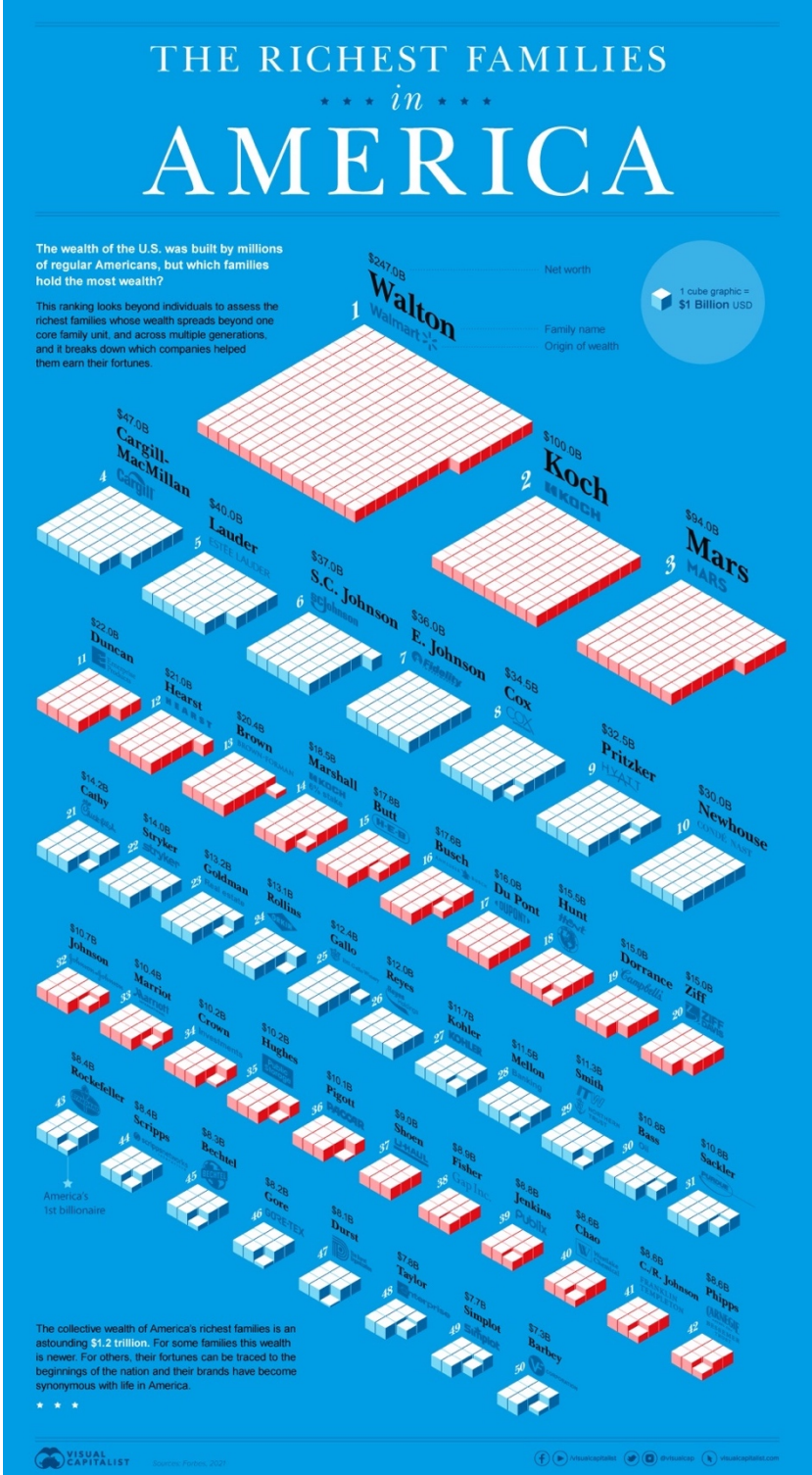


Figure 2.10

“Visualized: The Richest Families in America “, represented with blocks, by Avery Koop (2021).

Another notable visual, shown in figure 2.11, portrays the most popular subscription-based streaming services with slices of a pie chart. These visualizations are noteworthy

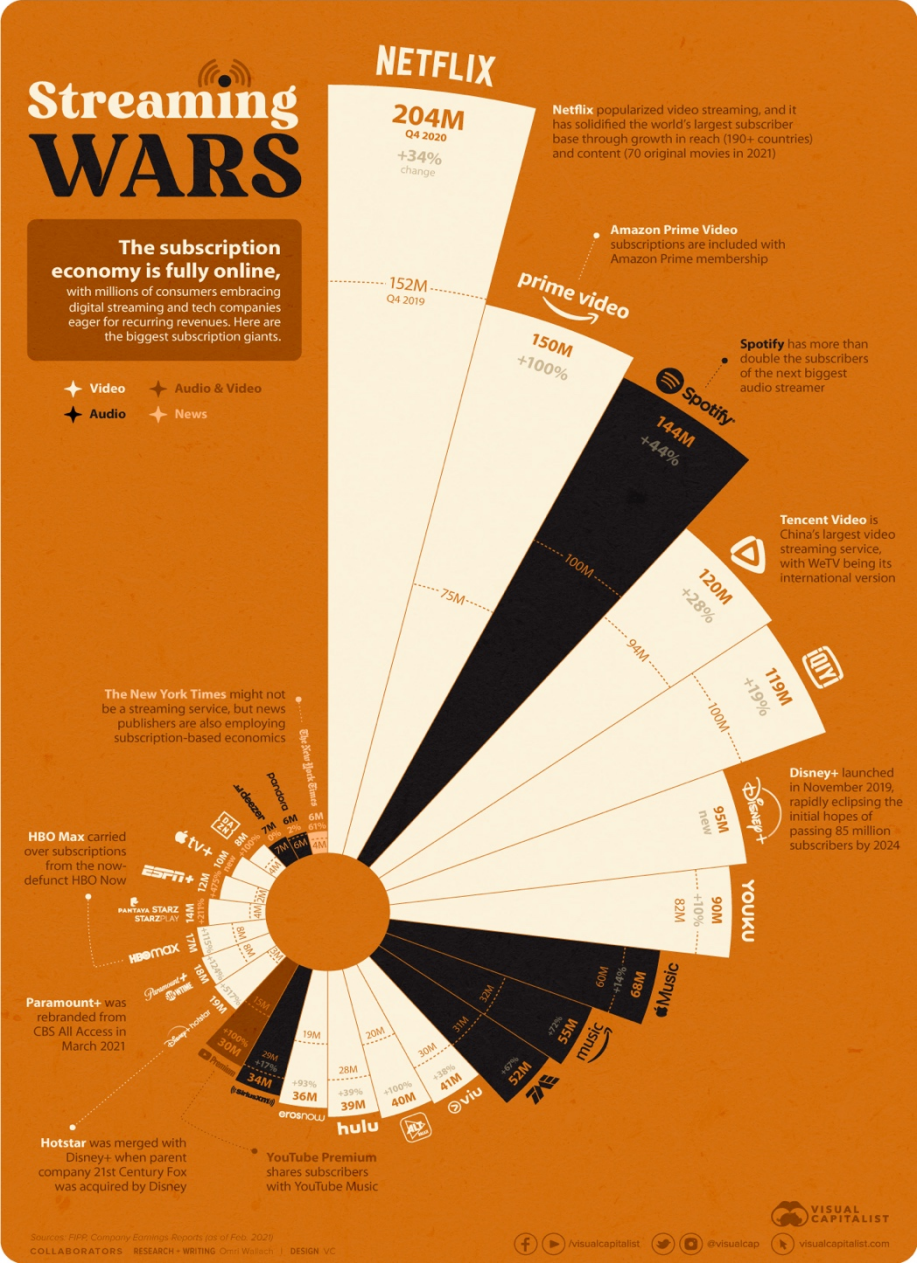


Figure 2.11
 “Which Streaming Service Has the Most Subscriptions” by Omri Wallach (2021).

because they represent a large amount of information in one static image, while displaying them with unique components like a spiraling pie chart and “brick” pieces.

They draw the user's attention and represent a large amount of information in one graphic. Visualizations like the ones shown in these figures are effective because they follow a set of important guidelines.

The goal of effective data visualization is to present comprehensive depictions of data that help a user understand the information without having to view a raw data file or read long strings of text. Edward Tufte, one of the creators of human-centered visualization methods, writes in *The Visual Display of Quantitative Information* that “Often the most effective way to describe, explore, and summarize a set of numbers—even a very large set—is to look at pictures of those numbers. . . of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful” (Tufte, 1985). Tufte makes the argument that it is often most effective to describe a data set with a visual representation; however, it is critical that these models meet his set of requirements to be high quality visualizations.

Tufte argues that visualizations satisfying *graphical excellence* will induce the viewer to think about the data rather than the design, avoid distorting what the data says, present many numbers in a small space, make large data sets coherent, encourage the eye to compare information, serve a purpose, and be closely integrated with the statistical description of the data set (1985). Many statistics, psychology, and design articles refer to these guidelines to support the quality of their visualizations (Venables & Ripley, 2001; Kress & Theo Van Leeuwen, 1996; Mayer, 2002). Unfortunately, it is far too common that designers and data scientists do not follow these guidelines, and

the resulting visualizations are difficult to understand. This goes against the motivation for producing visualizations, and works against the user trying to understand them.

Producing visualizations is not an easy task. There is a long process of finding the information, compiling and cleaning the data sets, using the right software and tools to create the visuals, and designing a graphic that will portray the information accurately, succinctly, and in a way that the viewers will understand it. Therefore, it is wasteful to produce imperceptive visualizations that do not appeal to the audience. Based on the default graphics offered in Excel, I assumed that these types of visualizations would fall into a default design category.

2.2 Common Mistakes in the Default Design Process to Visualize Data

In order to describe the default methods of generating data visualizations, I must contrast the guidelines that Tufte recommends in section 2.1, and illustrate common faults in the design process that produce ineffective visuals.

Visualizations created without a methodological approach, and that therefore do not meet Tufte's guidelines, are how I defined a default data visualization. I defined the default approach this way because it is the inverse of a human-centered design.

Visualizations that follow Tufte's principles are inherently human-centered because they incorporate the science of human perception and comprehension. If the designer wants to meet the goal of producing an effective graphic, then they need to think about the user viewing their design. Often the default approach does not include intentional thought about human perception, which will result in an ineffective visualization.

There are some famous examples of confusing and unnecessary visualizations that highlight some of these faults. For instance, in Figure 2.20 below, the viewer can

see causes of untimely death in regards to their annual change from 2005 to 2010.

Although Bill Gates voted this the best graph of the year in 2013 (Wonkborg), it has

been criticized by data visualists and designers alike. There are a few reasons this chart

is problematic. While it is eye-catching, the use of a three-dimensional block is

unnecessary and can lead to confusion. This goes against Tufte’s outlines for graphics

of making large data sets coherent, and doesn’t encourage the eye to compare different

pieces of data.

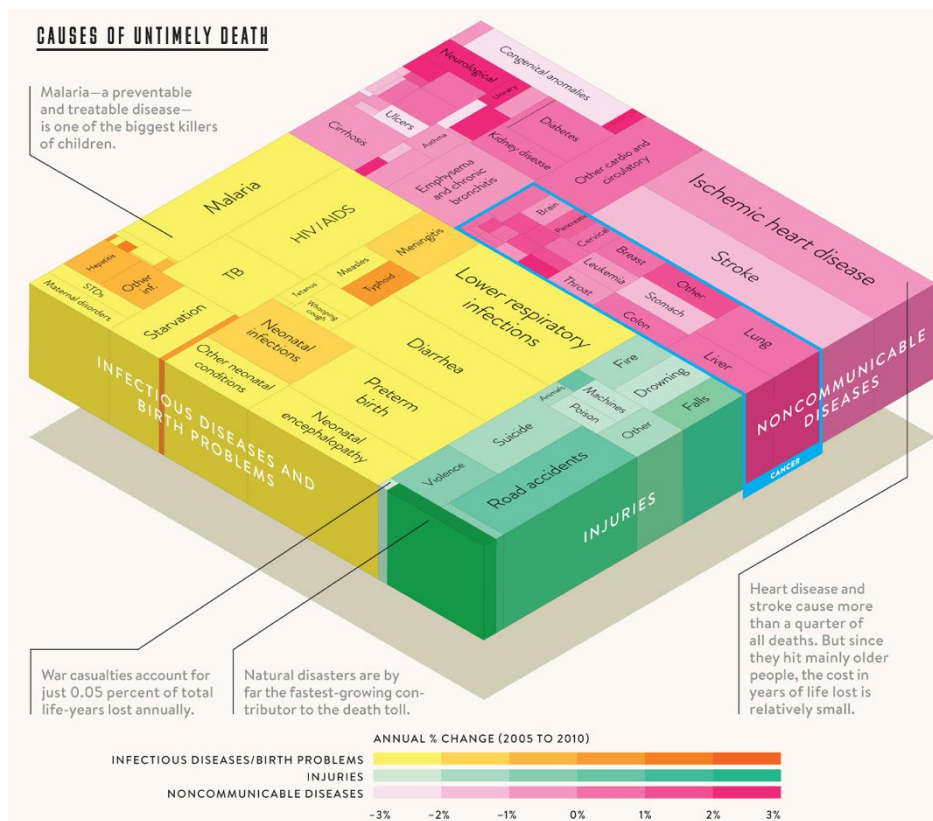


Figure 2.20

Bill Gates’s graph of the year in 2013 (Wonkborg, 2013).

2.3 An Argument for a Human-Centered Approach

A common saying in the U.S. Air Force is “*Perception is reality.*” My dad, an Air

Force fighter pilot, would say this phrase to me often as a child. Although this

expression reflects the focus necessary for American servicemen to work towards a common mission, it accurately represents how we can measure the effectiveness of data visualizations. If the person viewing the chart misunderstands the data being displayed, then the chart is ineffective. The viewer's perception of the chart determines their comprehension of the data being portrayed. This is why a human-centered design is so critical to ensuring understanding of the important information being illustrated in these visualizations.

2.4 Current Methods in Human-Centered Design

Two increasingly influential leaders in the human-centered design of data visualizations are Edward Tufte and Tamara Munzner. Each of them provides a set of guidelines that have motivated my design of the human-centered visualizations, and dispute the errors of the default visualization.

Edward Tufte is a Professor Emeritus of Statistics and Computer Science at Yale University, and has been said to be “The Leonardo da Vinci of data” by the New York Times (Tufte, 1985). He proposed a set of guidelines for graphical excellence. He defined excellence in graphics to include displays that show the data, induce the viewer to think about the substance rather than design of the visualization, avoid distorting what the data says, present many numbers in a small amount of space, make large data sets coherent, encourage the eye to compare different pieces of data, highlights the data at several levels of detail, serve a clear purpose, and are closely integrated with the descriptions of the data set (Tufte, 1985). While it is difficult to maintain all of these rules in one visualization, I used them to guide my creation of the human-centered graphics.

Tamara Munzner is a Professor of Computer Science at the University of British Columbia, and is an expert in information visualization. Munzner outlines the “Rules of Thumb” that guide the generation of effective visualizations. She asserts that these rules of thumb are meant to characterize which idioms are appropriate for tasks in data visualization design (Munzner, 2015). These guidelines are as follows: No Unjustified 3D; No Unjustified 2D; Eyes Beat Memory; Resolution over Immersion; Overview First, Zoom and Filter, Detail on Demand; Responsiveness is Required; Get It Right in Black and White; and Function First, Form Next (Munzner, 2015). I will highlight an important characteristic of these guidelines that have influenced my human-centered design methodology.

Munzner argues in her section about No Unjustified 3D that designers should carefully avoid tilted text due to its illegibility to most readers. She writes, “text fonts have been very carefully designed for maximum legibility when rendered on the grid of pixels that makes up a 2D display. . . . As soon as a text label is tilted in any way off of the image plane, it typically becomes blocky and jaggy” (Munzner, 2015). This is an interesting argument because, as I explain further in chapter three, Excel’s default design for many charts includes slanted text on the x-axis. When designing my human-centered visualization, I positioned all text to be flat to avoid this misstep. Her rules of thumb were central to my design of the human-centered visualizations.

Chapter 3: Methodology

In this section I describe the data sets used in this study and how I gathered them, how I designed the visualizations and the survey, how I selected participants and conducted the experiment, and how I analyzed the results. This methodology provides an overview of my research process from April 2020 until time of writing in April 2021.

3.1 Data Collection

After developing the questions that would drive this research, the next step of the process was to find relevant data sets to use in creation of the charts. I wanted to ask participants questions about their social-distancing and mask-wearing behavior, so the data needed to be evocative of their decision to mask and stay six feet apart from others. I looked for data sets related to case counts, masking, social-distancing, or any information connected with the pandemic.

There were several challenges in finding a relevant and usable set. The sets needed to be publicly available so I would not need a license to visualize the data. They also needed to be in raw form so I could prepare the information for the visualization. Additionally, at the time of data collection in the early months of the pandemic, it was difficult to find publications related to COVID-19. However, I eventually found and investigated the provenance of a few data sets.

In August, I found the New York Times' "covid-19-data" repository on Github. This repository includes a "mask-use" dataset that contains estimates of mask usage by county in the United States based on results of a survey conducted by global data and survey firm Dynata (New York Times, 2020). Over 250,000 participants were asked, "How often do you wear a mask in public when you expect to be within six feet of

another person?” with the available responses as *Never*, *Rarely*, *Sometimes*, *Frequently*, and *Always*. These responses were obtained from July 2 to July 14, 2020.

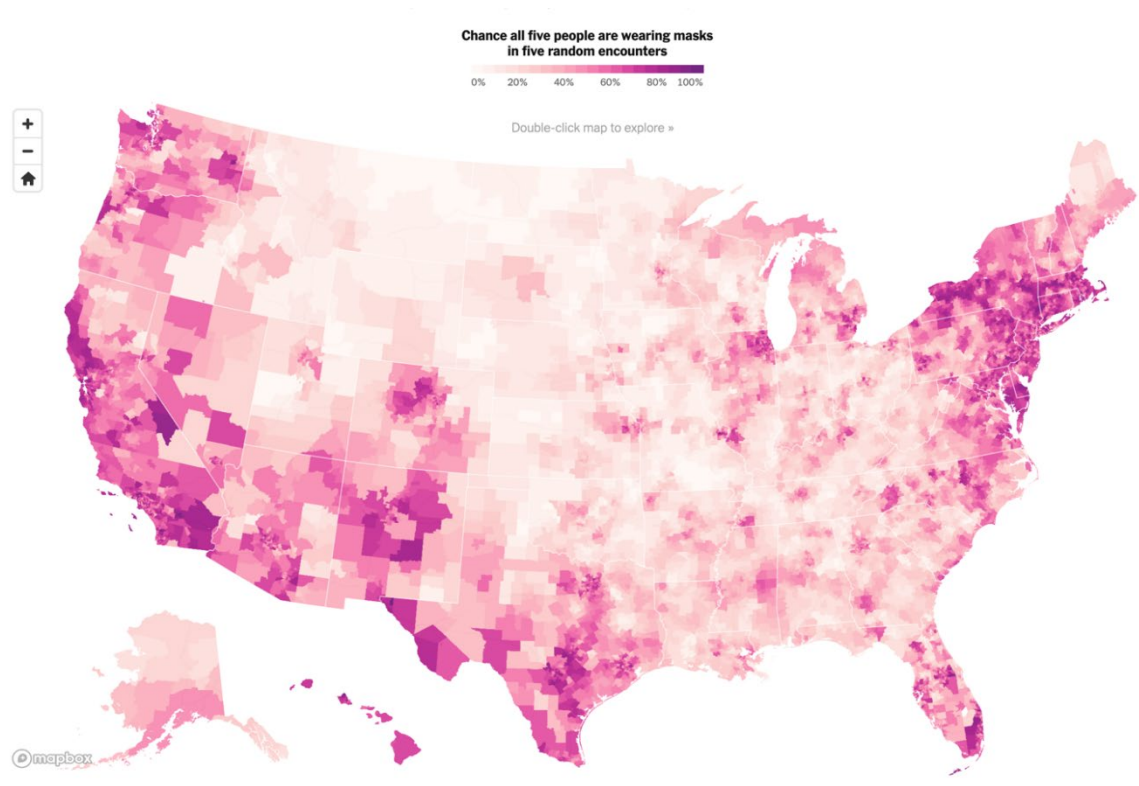


Figure 3.10

“A Detailed Map of Who is Wearing Masks in the U.S.” by Josh Katz, Margot Sanger-Katz, and Kevin Quely (2020).

These data were freely available for public research use but showed surprising results. The data reflected high mask usage in places where cases were rising rapidly. For instance, on July 14 during the time of conducting the survey, Florida was entering its highest peak since the pandemic began with 9,100 new reported cases (New York Times, 2021), but the self-reported mask usage data reflected that the average probability one would encounter someone wearing a mask in Florida was 79.8% (New York Times, 2020). Public health experts would argue that with a large majority of people wearing masks, this would help reduce the spread rather than amplify it.

I computed the average probability of encountering someone wearing a mask in Florida by assigning 0%, 25%, 50%, 75%, and 100% to *Never*, *Rarely*, *Sometimes*, *Frequently*, and *Always*, respectively, and adding the product of the frequency of each response with these percentages. The raw data and the visualization resulting from computing the probability of encountering someone wearing a mask in each county of Florida can be viewed in Figure 3.11 and 3.12 below.

	B	C	D	E	F	G	H	I
1	STATE	COUNTYFP	NEVER	RARELY	SOMETIMES	FREQUENTLY	ALWAYS	PROBABILITY_WEARING_25_50_75
2	Alabama	1001	0.053	0.074	0.134	0.295	0.444	0.75075
3	Alabama	1003	0.083	0.059	0.098	0.323	0.436	0.742
4	Alabama	1005	0.067	0.121	0.12	0.201	0.491	0.732
5	Alabama	1007	0.02	0.034	0.096	0.278	0.572	0.837
6	Alabama	1009	0.053	0.114	0.18	0.194	0.459	0.723
7	Alabama	1011	0.031	0.04	0.144	0.286	0.5	0.7965

Figure 3.11

Raw data from the mask usage file with each county FIPs code, response, and probability computation, sourced from the New York Times mask-usage data set (2020), and combined with the U.S. Census Bureau's 2018 FIPs Codes to align each row with the corresponding state (2020).

Probability of Encountering Someone Wearing a Mask in Florida in July 2020

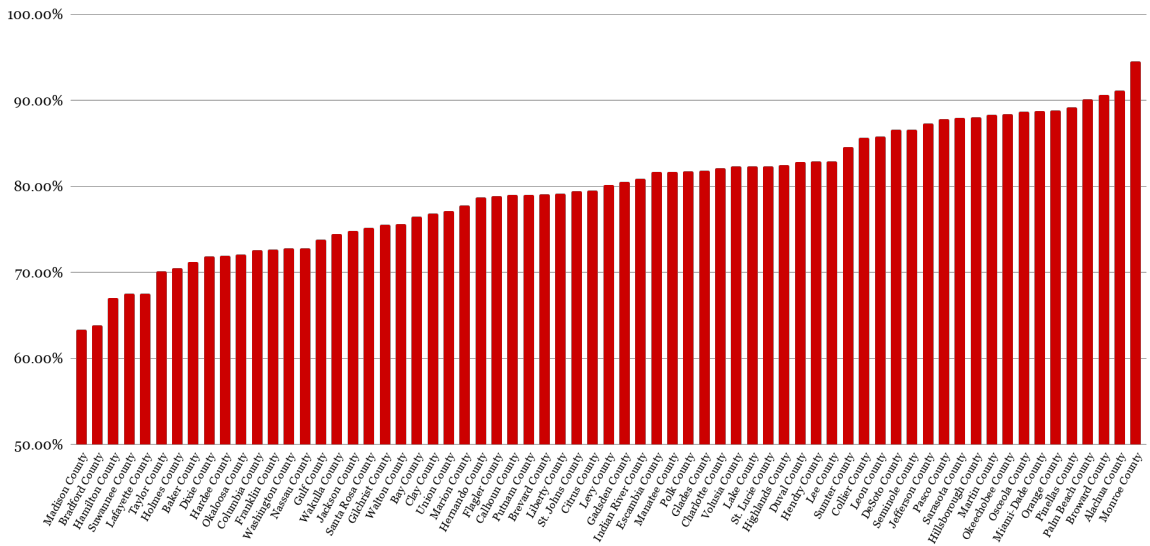


Figure 3.12

The probability of encountering someone wearing a mask in Florida in July 2020. Data sourced from the New York Times “covid-19-data” repository (New York Times, 2020), and visualization created using Google Charts.

These results led me to believe that respondents were self-reporting higher mask usage than their true mask-wearing behavior. The visualizations I produced using these data portrayed mask-wearing to have little or no effect on the rising cases of COVID-19 across the nation, so I decided to find another set that would reflect a more compelling argument to practice social-distancing and mask-wearing protocols.

I decided to use the Centers for Disease Control and Prevention’s (CDC) “United States COVID-19 Cases and Deaths by State Over Time” case surveillance data (CDC Case Task Force, 2021). This set includes the state, total cases, total deaths, new deaths, and the day the counts were recorded for each county in the U.S. This data provided a solution to each of my research components: it is available for public

research use, in raw form, and addressed the effects of social-distancing and mask-wearing.

3.2 Creating Visualizations

The next step after finding this set was to clean and organize the information into a format ready for visualization. The data file contains over 27,800,000 rows and 15 columns of information about case rate and death count. I separated the data into two sets: a daily overview of cases, and a cumulative overview of cases. I then separated the sets by state, and generated the following two visualizations: for the daily chart, I used information regarding submission date and new cases, whereas for the cumulative chart I used submission date and total cases. For both data sets, I narrowed in on Florida as the state to represent. I selected Florida because it had some of the highest reported cases in the country, and many of the outbreaks were happening on college campuses. I hoped a visualization of the rapid spread would make the most compelling argument to social distance.

Once I organized the data, I was prepared to generate the visualizations. I initially started designing the visualizations using Google Charts, as seen in Figure 3.11 above, but ultimately used Microsoft Excel for its diversity of design elements, the quality of PNG exports for charts, and the default colors, fonts, and line sizes that provide a common representation of data visualizations. I designed four total visualizations to test my hypothesis: human-centered daily and cumulative visualizations, and default daily and cumulative visualizations. I wanted to provide two charts for each design method to provide additional validity for each design's efficacy.

I designed the default charts for the daily and cumulative set first, and then used that design as a starting point to build the equivalent human-centered chart. For the default visualization, my approach was to use the default settings for most of the chart's design. I used Georgia for font since it is one of the most widely used fonts on the web (Rawsthorn, 2006), set the line width at 1.5 point for the daily chart and 3 point for the cumulative chart, and left the line color as the default light green. The numbers representing case counts were automatically placed to the left of the y-axis, so I left them there as well. Excel automatically places dates at a slanted angle, so I left them in this position.

My approach to designing the human-centered charts was to improve or remove the aspects of the default chart that have been shown to negatively impact the comprehension and aesthetic appeal of the visualization. Based on Tufte's guidelines for graphical excellence and Munzner's rules of thumb, I made several changes to the font, color, lines, text, and positioning in the visuals. I widened the lines from 1.5 to 3 point in the daily visual, and from 3 point to 6 point in the cumulative visual. I made the line a deep blue so that the graph would look the same if it were converted to black and white. This is so that any user, including those who are colorblind, could read the lines on the graph, and is a strategy suggested by Munzner in "Get it Right in Black and White" (2015). Furthermore, I flattened the dates on the x-axis, and put the timeline descriptions in black boxes to separate the marking lines from the text. I eliminated the tilted dates on the x-axis based on one of Munzner's Rules of Thumb that argues "Tilted Text Isn't Legible" (2015). I also placed the digits along the y-axis to put them closer to the most recent dates on the x-axis. Lastly, I changed the font to Arial Bold because

of a study supporting its readability on computer screens in a user survey with serif and sans serif typefaces (Wilson, 2001). All of these design choices were supported by Munzner and Tufte's guidelines about generating effective visuals.

3.3 Designing the Survey

My goal for designing the survey was to ask questions that would measure the effectiveness of the visualization's ability to represent the data, determine which visualizations were most pleasing to view, and gauge the respondent's willingness to socially-distance and wear a mask before and after viewing the visualizations. I structured the questions around the four visuals (shown below), and organized the flow of the survey to present half of the participants one design of a visual with two total graphs, and the other half would view the alternative design with two total graphs. This meant that the participants would have a 50% chance of viewing the human-centered visuals or the default visuals. Before viewing the visualizations, they were asked questions about their social distancing behavior, during viewing they were asked about components of the data to measure their understanding, and after viewing they were asked about future social distancing behavior and demographic questions (see Appendix).

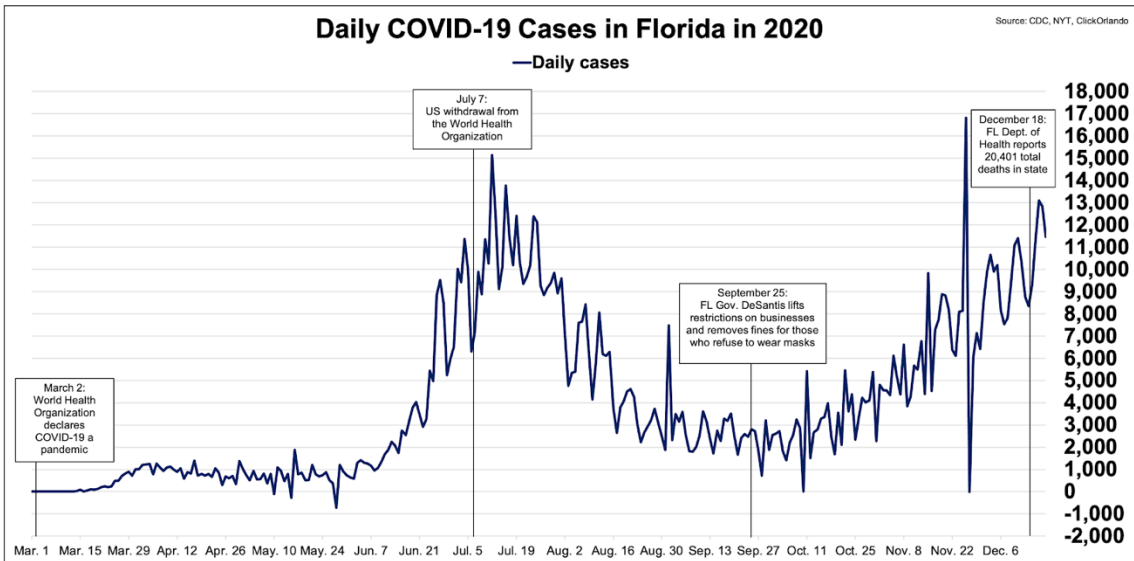


Figure 3.3.1: Human-Centered Daily Cases Visualization

Daily cases of COVID-19 across the state of Florida over March to December of 2020, created with a Human-Centered Design methodology.

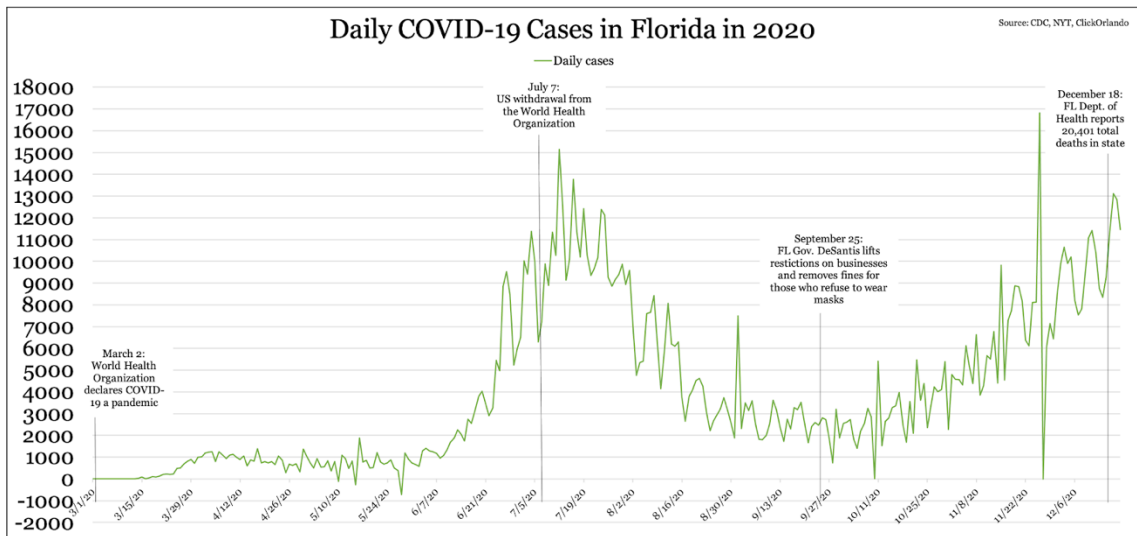


Figure 3.3.2: Default Daily Cases Visualization

Daily cases of COVID-19 across the state of Florida over March to December of 2020, created using a default approach.

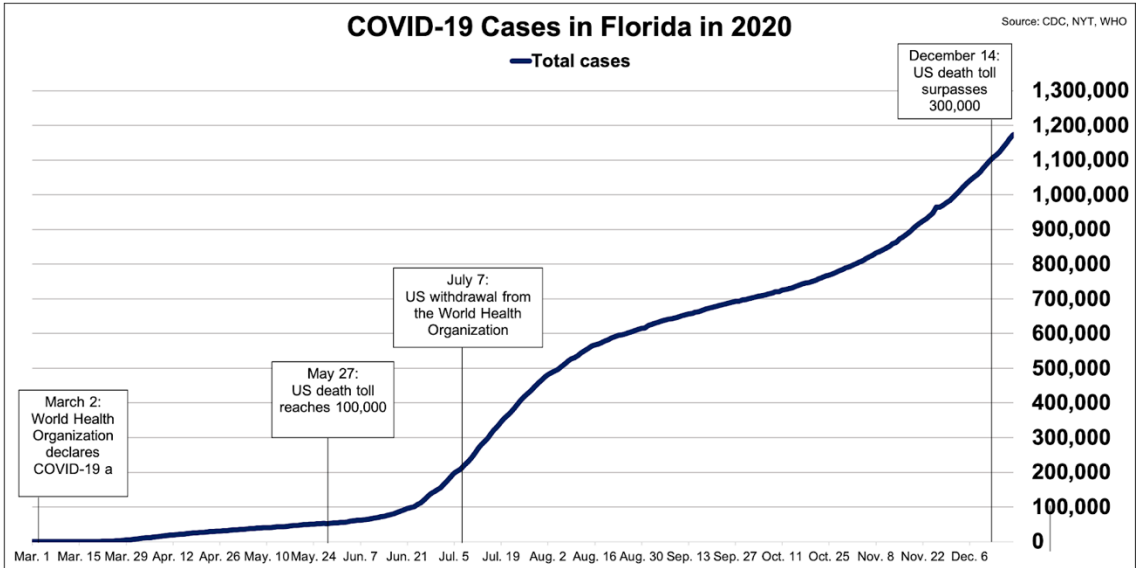


Figure 3.3.3: Human-Centered Cumulative Cases Visualization

Cumulative cases of COVID-19 across the state of Florida from March to December of 2020, created with a Human-Centered Design methodology.

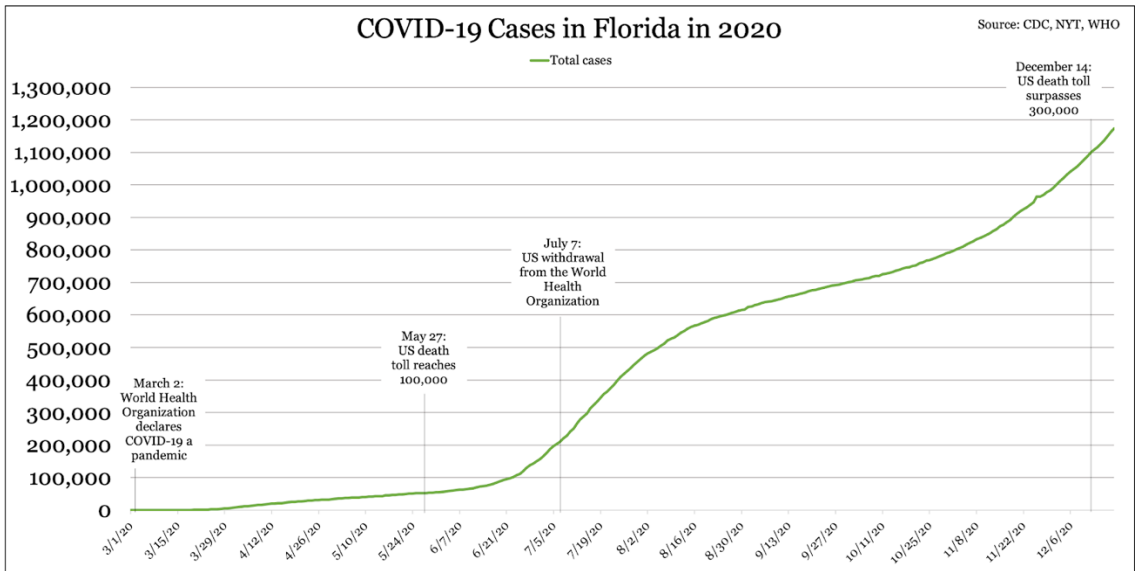


Figure 3.3.4: Default Cumulative Cases Visualization

Cumulative cases of COVID-19 across the state of Florida from March to December of 2020, created using a default approach.

3.3.1 Structure and Question Design

I decided to use Qualtrics for the survey platform for its professional design, data analysis and report capabilities, and familiarity with University of Oregon affiliates. I organized the questions in the following structure: survey information and consent statement, questions about social-distancing and mask behavior, identical questions to measure comprehension of the cumulative data for both the human-centered and default visualizations, identical questions to measure comprehension of the daily data for both the human-centered and default visualizations, follow-up questions gauging the respondent's future masking and social-distancing behavior, demographic questions, and then a debriefing form.

I made a few key choices in constructing the response options in each question. For the appeal and behavior questions, I decided to offer participants a Likert scale with only four choices so that they must make a polar decision. This is to avoid respondents choosing to "sit on the fence" and select a null answer such as "Neutral" or "Unsure". This allowed me to compare the responses between questions and charts, and compute significance of comprehension, appeal, and predictions about future behavior.

Duration of the survey was an important consideration for how I posed each question. I kept all questions as multiple choice to keep the participant engaged and moving through the survey as fast as possible. The survey duration was limited to spans of 15-minute intervals due to the 0.25, 0.5, 0.75, and 1.0 credits offered to student researchers in the UO human subjects pool, so I needed to expedite all processes in the questionnaire with the way I framed the response choices. With all questions as multiple choice, the Qualtrics survey rating system, ExpertReview, predicted that the

questionnaire would take approximately 10.3 minutes to complete. I extended the duration to 15 minutes so that participants would have ample time to finish the survey.

All questions and survey materials can be found in the appendix for reference while reading this section.

3.3.2 Social Distancing and Mask Behavior Questions

My intentions for the mask usage and social-distancing questions were for the participant to actively reflect on their own behavior as they moved through the survey, and to measure their willingness to change their social-distancing behavior. For these questions, I asked about which behaviors the respondent has participated in over the last 24 hours, for which of those activities they decided to wear a mask, and their comfortability participating in a variety of indoor and outdoor activities. I duplicated each of these questions so the participant would first reflect on their current behavior and later make a decision about their future behavior. To measure their willingness to change their behavior, I positioned these questions before and after the participant viewed the visualizations. An example of one set of these parallel questions regarding social-distancing behavior can be seen in figures 3.30 and 3.31 below.

To what extent are you intentionally avoiding contact with other people?

All of the time: I always avoid contact with others.

Most of the time: I only leave home to purchase food and other essentials.

Some of the time: I have reduced the amount of times I am in public spaces, social gatherings, or at school.

None of the time: I am often in physical contact with others.

Figure 3.30

Question seven in the survey asks participants about how often they are physically interacting with others. After viewing the visualizations, participants are asked the same question again regarding their behavior over the next three days in question twenty-two.

In the next 3 days, to what extent will you intentionally avoid contact with other people?

All of the time: I will always avoid contact with others.

Most of the time: I will only leave home to purchase food and other essentials.

Some of the time: I will reduce the amount of times I am in public spaces, social gatherings, or at school.

None of the time: I will be in physical contact with others often.

Figure 3.31

Question twenty-two asks participants about their social distancing behavior over the next three days to mirror question seven from before viewing the visualizations.

For the participant to fully reflect on their behavior, I asked about several aspects of their daily life regarding COVID-19 safety protocols. The survey asked the participant for their estimation of how many people they had direct contact with outside of their own household, which included a definition of direct contact: “*‘Direct contact’ means a conversation lasting more than 5 minutes with a person who is closer than 6 feet from you without either person wearing a mask.*” I also asked for an estimation of how many people *around the participant* decided to wear a mask in public spaces when social distancing was not possible over the past three days, with a definition of social distancing: “*‘Social distancing’ means staying 6 feet away from those around you.*” After answering these questions, the participant would view the randomly selected design of visualization and answer the associated comprehension questions, and then view the same social-distancing questions regarding their future behavior afterwards. Mirroring the precursory and subsequent behavior questions allowed for the results to simulate the participant’s choice to social distance and mask in the future.

3.3.3 Measuring Comprehension Questions

To measure the effectiveness of the visualizations, I asked questions that would gauge the respondent’s comprehension of the data. I structured the questions to ask for identical information in both charts so I could compare participant comprehension rates. For instance, after viewing the human-centered or default visualization for daily case rates, the survey participant would be asked: “Which state is being represented?”, “What is the single highest recorded number of daily cases in Florida?”, and “On September 25, approximately how many daily cases of COVID-19 were recorded in Florida?”

For the human-centered and default visualization for cumulative case rates, I asked for the same information but referred to a different date in the third question. These questions can be seen below in figures 3.32 and 3.33.

On **July 7th**, approximately how many cases of COVID-19 were recorded in Florida?

210,000

150,000

400,000

350,000

Figure 3.32

Participants would first view the cumulative data set visualization, and then answer questions about which state is represented, the single highest number of recorded cases, and how many cases were recorded on July 7th.

On **September 25**, approximately how many daily cases of COVID-19 were recorded in Florida?

4,300

1,200

2,800

4,100

Figure 3.33

After seeing the cumulative cases visualization, the participant would view the daily cases visualization and answer questions about which state is represented, the single highest number of recorded cases, and how many cases were recorded on September 25.

Using the results of these questions, I could determine whether the respondent was truly participating in the survey, and whether they understood the information being displayed in the visual.

3.3.4 Visual Appeal Questions

After being asked about specific pieces of data in the visuals, I asked the participants about their personal preference and experience viewing the visualizations. I asked, “*Complete the sentence: ‘Reading this chart was...’*” with the Likert scale response of *Very easy, Easy, Difficult, and Very difficult*. I also asked, “*Respond to the statement: ‘I enjoyed looking at this chart.’*”, and “*Respond to the statement: ‘I understand the information being portrayed in this graph,’*” with the responses *Strongly agree, Somewhat agree, Somewhat disagree, and Strongly disagree*. With these results, I was able to determine which visualizations were preferred and most pleasing to view.

3.3.5 Demographic Questions

In order to measure the effectiveness, appeal, and persuasiveness of my visualization across demographics, I asked questions referring to many areas of one’s identity. I asked each participant for their current ZIP code, if they have any known color blindness, their gender and age, if they are of Hispanic, Latino, or Spanish origin, their race, and the highest degree or level of school completed. The results of these questions allowed me to compare comprehension rates among younger versus older participants, female and male participants, and those with colorblindness and those without. Furthermore, knowing which demographic responded best to each visualization allowed me to predict how each graphic could be improved for specific age groups, genders, colorblindness, and other identity traits.

3.4 Participant Criteria

The survey was distributed through three main channels: the UO Human Subjects pool to students taking Psychology courses, posting the survey on my personal LinkedIn, Facebook, and GroupMe accounts, and by use of Amazon's Mechanical Turk. Figure 3.41 shows the recruitment statement used to distribute the survey to UO Computer Science students. I used a sample size calculator (Creative research systems, 2012) to determine the number of people I needed to participate in the survey. With a confidence level of 95% and a confidence interval of 5, I needed to reach 384 participants. After the recruitment period through all three channels, I was able to gather 427 responses.

Credit allocation affected the number of people I was able to recruit to take the survey. The UO human subjects pool (UOHSP) participants taking a psychology course must participate in 5 credit hours of research. Since I described the survey to have a duration of 15 minutes, I offered 0.25 credits to each student. While I predicted the short duration would increase the number of participants, the UOHSP only contributed one participant to my overall sample. Many of the surveys and experiments posted on the UOHSP site offer more credits to students, which may have negatively impacted my ability to recruit respondents. I had greater success recruiting participants using my personal social media and offering payment to workers on Mechanical Turk.

Using my personal social media as a distribution method was a successful tactic in recruiting participants. I posted a statement asking viewers to take the survey on my Facebook, LinkedIn, and sent a private message to the Oregon Consulting Group, of which I am a member. After splitting the sample from the Mechanical Turk workers and

UOHSP, the survey results show that I was able to recruit 111 participants over the span of a month using this method. This is an interesting finding because I offered no reward for taking the survey, which leads me to believe that familiarity with the person distributing the survey has a positive effect on participant size. I continued this positive recruitment trajectory when I distributed the survey on Mechanical Turk.

I was able to recruit the largest number of participants from Amazon's Mechanical Turk platform. Amazon's Mechanical Turk offers researchers access to a pool of workers that receive small payments for each Human Intelligence Task ("HIT") they complete. These HITs are often transcribing receipts, reading documents, and taking surveys. Thanks to funding from my advisor, I was able to offer participants \$3.75 for completing the single HIT of my 15-minute survey—an equivalent of \$15 an hour. I decided to offer this amount after reading an article from the Atlantic (Semuels, 2018) about how little Mechanical Turk workers earn. The worker interviewed in the article shared that she made an average of \$4 to \$5 an hour from her participation in these tasks (Semuels, 2018). With this payment, I was able to receive 275 responses within 9 hours.

3.5 Data Analysis

My null hypothesis is that the design method used in generating the visualizations has no effect on comprehension of data, aesthetic appeal of the chart, or willingness to social distance. My alternative hypothesis is that participants who view charts generated using a human-centered approach will show higher rates of data comprehension, aesthetic appeal, and willingness to social distance than those who viewed the default charts. I used these hypotheses to determine whether my human-centered design (HCD)

was more effective in each demographic of respondents. For all of my analyses, I computed a confidence interval of the difference between proportions. This means that I compared the difference between each response coming from pre- and post-visualization questions, and determined if that difference were larger than the critical value associated with p-value of .05. If it was larger than .05, then the result was statistically significant, and I could reject the null hypothesis and accept the alternative hypothesis for that question.

In order to determine how many participants of each demographic were viewing the HCD visualizations, I filtered the results based on whether every default visualization question was empty. The set of visualization questions associated with the default visualizations would only be empty if the participant had not been presented with them. This leaves the set of participants who were presented the opposite visualizations—in this example, the human-centered visualizations. After splitting the respondents based on this filter, I had a new set of respondents and needed to determine the new confidence interval for the size of the set using an interval calculator (Creative research systems, 2012) for a confidence level of 95%.

In order to measure the effectiveness of each visualization, I analyzed the rates of correct responses, the rates of appeal based on the Likert scale responses, and the difference in social-distancing behavior and masking after viewing the visualizations. I started by computing the confidence interval for each demographic sample size, and compared the responses between the two designs. For instance, the entire sample contained 427 participants which corresponds to a confidence interval of +/- 4.74% with a confidence level of 95%, whereas the sample size of the colorblind participants was

56 which corresponds to a confidence interval of +/- 13.1% with the same confidence level (Creative research systems, 2012).

For Likert scale responses, I added the rates of “Very comfortable” and “Comfortable”; “Very uncomfortable” and “Uncomfortable”; “Strongly agree” and “Agree”; “Strongly Disagree” and “Disagree”; “Very easy” and “Easy”; and “Very difficult” and “Difficult” together to compare the results using the confidence interval for each set. Each of the categories of responses sum to 100%; e.g., the response rates of “Very comfortable”, “Comfortable”, “Very uncomfortable”, and “Uncomfortable” totaled to 100%.

Chapter 4: Results

This research was motivated by the questions posed in Chapter One: Why use a human-centered design approach to creating visualizations? Can visualizations be clear, and convincing enough to change people’s behavior after looking at them? What is the best way to present visualizations of data for maximum human-understanding?

Ultimately, my study answered the question: which method of visualizing COVID-19 case data will result in the highest number of responses showing a willingness to continue social-distancing and mask-wearing behavior, or change their behavior if not doing so? In this chapter I will provide a breakdown of the survey results, including an analysis of each sample of participants. I will go through each section of the survey to address my research questions, and determine if my hypothesis was supported. The legend showing the categories of each of the questions is shown below in figure 4.10.

Question Legend								
Question Categories	Questions							
Social-Distancing and Masking	Q62	Q63	Q97	Q130				
Pre-Visualization Behavior	Q93	Q98	Q131					
Cumulative Cases Default Visualization	Q153	Q70	Q50	Q135	Q134	Q66	Q145	Q139
Daily Cases Default Visualization	Q176	Q157	Q165	Q161	Q163	Q170	Q171	Q172
Cumulative Cases HCD Visualization	Q154	Q121	Q122	Q136	Q137	Q124	Q146	Q138
Daily Cases HCD Visualization	Q159	Q158	Q166	Q162	Q164	Q173	Q174	Q175
Post-Visualization Behavior	Q71	Q100	Q142					
Demographic	Q52	Q88	Q54	Q55	Q56	Q657	Q58	

Figure 4.0: Question Legend

Legend shows each category associated with the questions. The numbers associated with each question correspond to the order of creation during survey design.

4.1 Participant Demographic

The demographic breakdown of the 392 survey participants is based on a few categories: ZIP code, color blindness, age, gender, ethnic background, and education. Analysis of the ZIP code entries show that participants, at the time of taking the survey, are residing in 38 states across the U.S., including Alaska and Hawaii.



Figure 4.11

Participants were asked to enter their current ZIP code, which resulted in codes across 38 states.

4.1 Overview of Entire Sample

There were 392 responses for the entire sample of survey participants. The participants can be split into two groups based on the visualizations they viewed: those who were presented the HCD designs, and those who were presented the default

designs. The sample was evenly split with 196 HCD participants and 196 default participants.

The social-distancing and masking questions, Q62, Q63, Q97, and Q130, were posed in order for the user to reflect on their own masking and social-distancing behavior, and prime them for future questions about COVID-19. The results of these questions show that respondents are going to a grocery store, a pharmacy, work, or school, spending time with people the respondent does not live with, or exercising outdoors most frequently. Of those activities, they are wearing a mask most often when they go to a market or pharmacy--28.15% of the time. Somewhat surprisingly, respondents are wearing a mask most infrequently on public transportation--5.53% of the time. On average, participants reported having direct contact with 8.58 people at work or school, 8.9 people at a grocery store or pharmacy, 8.94 people at social gatherings, and 7.93 people during other activities. Across many demographics including this set of all participants, results did not show increases in comfortability for exercising, whether indoors or outdoors. All these activities regarding comfortability are related to how the respondent feels participating in these activities in general, with no assumption of wearing or not wearing a mask, to measure whether there was an impact of viewing the visualizations on their willingness to participate in these activities. Respondents reported that the people around them chose to wear masks “All of the time” at a rate of 34.78%, and “Most of the time” at a rate of 37.92%.

The pre-visualization behavior questions, Q93, Q98, and Q131, all ask the respondent about their current mask-wearing and social-distancing behavior. The post-visualization behavior questions, Q71, Q100, and Q142 ask the same questions but refer

to future masking and social-distancing behavior. These questions are used to measure the participant's willingness to change their behavior after viewing the visualizations. Results from the questions prior to viewing the visualizations show that 63.63% of respondents reported that they felt comfortable going to work or school in the past two weeks. After viewing the visualizations, participants reported that they feel 71.51% comfortable in the same setting when every person is wearing a mask—an increase of 7.88%.

Results from the remaining four pre- and post- visualization questions also show significant increases in comfortability across activities. Participants said they felt comfortable going to a bar or restaurant in the past two weeks 47.38% of the time, but reported they feel comfortable 56.01% of the time when every person is wearing a mask—an increase of 8.63%. Similarly, respondents showed an increase in their comfortability rating in every activity aside from exercising outdoors.

For the set of comprehension questions asked while viewing the visualizations, the results show that for Q161 and Q162, there is a difference between the accuracy of responses for the human-centered design and default charts, with the daily default chart receiving 81.38% correct responses, and the daily HCD receiving 72.82% correct responses—a difference of 8.56%. The remaining three questions in this block have participants rate readability and preference of the visualization they viewed. On Q66 and Q124, the cumulative default chart was rated as easiest to read, with a response rate of 89.8% for the cumulative default visual, and 81.63% for the cumulative HCD visual. For questions 145 and 146, 83.16% of participants who viewed the cumulative default

visual said they enjoyed looking at the chart, whereas only 75.39% of those who viewed the cumulative HCD chart said they enjoyed it.

4.2 Age

For this demographic, I split the age groups of participants on those who are 18 to 34 years old, and those who are 35 years or older.

4.2.1 Ages 18-24

There were 221 responses for the set of 18-to-34-year old's. Recall that the pre-visualization behavior questions, Q93, Q98, and Q131, all ask the participant about their current mask-wearing and social-distancing behavior. The post-visualization behavior questions, Q71, Q100, and Q142 ask identical questions but refer to future masking and social-distancing behavior. Results from Q93 and Q71 reflect an increase of 7.54% in overall comfortability going to work or school after viewing the visualizations.

Similarly, respondents showed an increase in comfortability when going to a market or pharmacy; a bar, restaurant, or café; attending an event with more than ten people; using public transit; and exercising in a fitness facility. For question 98 and 100, which ask about how often the participant will wear a mask, the respondents reported they were “always” wearing one with a frequency of 53.39% before the visualization, and a frequency of 63.80% afterwards—an increase of 10.41%.

For the set of comprehension questions asked while viewing the visualizations, results show that for Q135, Q161, Q136, and Q162, there is a statistically significant level of difference between the accuracy of responses for the human-centered design and default charts, with the daily default chart receiving 82.88% correct responses over the daily HCD chart that received 73.64% correct responses. Inversely, the cumulative

HCD chart received 83.64% correct responses, a jump over the cumulative default chart that received 68.47% correct responses.

For the questions regarding readability and visual appeal, results reflect that participants enjoyed looking at the cumulative default chart most frequently. When asked “Respond to the statement: ‘I enjoyed looking at this chart.’”, 85.59% of participants who viewed the cumulative default chart responded “agree”, whereas only 78.18% of participants responded “agree” for the cumulative HCD chart.

4.2.2 Ages 35+

For the group of participants 35 years or older, there were 148 responses. The pre-visualization behavior questions reflect an increase in level of comfortability for participants when going to work or school; going to a market or pharmacy; going to a bar, restaurant, or café; spending time with someone they are not staying with; attending an event with more than ten people; and using public transit. There was no difference between comfortability levels in exercising in a fitness facility or exercising outdoors before and after viewing the visualizations. For Q98 and Q100 about mask usage, respondents reported that they would wear a mask “frequently” 15.54% of the time prior to viewing the visualization, which increased to a rate of 25.00% after viewing the visualizations.

For the set of comprehension questions asked while viewing the visualizations, the results show that respondents correctly answered the daily HCD chart most frequently, but rated the cumulative default chart as easiest to read. In Q164, the daily HCD visualization received 65.38% correct responses, whereas the daily default visualization only received 55.71% correct responses. In regards to readability,

participants rated the cumulative default chart as “easy” with 90% frequency, whereas the cumulative HCD chart only received “easy” for 75.64% of the responses.

4.3 Gender

The survey results reflect a gender split of 38% female, 61% male, and 1% other participants. In this section, I will report the results of the female and male demographics.

4.3.1 Female Participants

There were 140 responses from female participants. Results from the pre-visualization behavior questions reflect that female participants were less persuaded by the visualizations regarding comfortability in comparison to the demographics analyzed above. Participants showed an increase in comfortability before and after viewing the visualizations in regards to going to a market or pharmacy; spending time with someone they are not staying with; attending an event with more than ten people; and using public transit. There was no meaningful difference reported in levels of comfortability regarding going to work or school; going to a bar, restaurant, or café; exercising in a fitness facility; or exercising outdoors. For questions 98, 100, 131, and 142, there was no meaningful difference between willingness to wear a mask or intentionally avoid contact with others before and after viewing the visualizations.

The results of the comprehension questions for female participants show a mix of both the default and HCD visualizations in terms of comprehension and readability. For Q161 and 162, there is a level of difference between the accuracy of responses for the human-centered design and default charts, with the daily default chart receiving 81.54% correct responses, whereas the daily HCD chart received 70.67% correct

responses. Furthermore, the cumulative HCD visualization received 77.33% correct responses in Q136, where the cumulative default chart received 66.15% correct answers in Q135. The other comprehension questions in this block did not show a meaningful difference in accuracy of responses. In terms of readability, participants rated the cumulative default chart as “easy” 90.77% of the time, whereas the cumulative HCD visualization only received 74.67% “easy” responses. The other questions regarding enjoyment of looking at the chart and understanding the information being portrayed did not reflect meaningful differences between the HCD and default charts.

4.3.2 Male Participants

There were 228 responses from male participants. The pre-visualization behavior questions reflect that male respondents reported increased levels of comfortability in more activities after viewing the visualizations in comparison to the female participants. The results of Q93 and Q71 reflect that male participants showed increases in comfort levels when asked about going to work or school; going to a market or pharmacy; going to a bar, restaurant, or café; attending an event with more than ten people; and using public transit. Results reflected no meaningful differences between the level of comfortability spending time with someone they are not staying with; exercising in a fitness facility; or exercising outdoors.

For question 98 and 100, which ask about how often the participant will wear a mask, the respondents reported they were “always” wearing one 47.37% of the time before viewing the visualization, but reported being willing to wear one “always” 53.95% of the time after viewing the visualizations—an increase of 6.58%. Results

reflected no meaningful differences between the extent to which participants avoid contact with others before and after viewing the visualizations.

The results of the comprehension questions for the male participants show that the default charts had the highest rates of comprehension, readability, and visual appeal. For Q161 and Q162, there is a level of difference between the accuracy of responses for the human-centered design and default charts, with the daily default chart receiving 82.76% correct responses, where the HCD received 75.89% correct responses. On Q163 and Q137, the cumulative default chart received 68.10% correct responses, whereas the cumulative HCD chart received 60.71% correct responses—a difference of 7.39%. On Q134 and Q137, participants answered the questions associated with cumulative default chart correctly at a rate of 68.10%, whereas the participants who viewed the cumulative HCD chart responded correctly at a rate of 60.71%. Respondents who viewed the cumulative default chart responded “agree” at a rate of 85.35% to the question asking about enjoyment of looking at the chart, whereas the respondents who viewed the cumulative HCD selected “agree” at a rate of 74.77%. Similarly, participants agreed that the daily default visualization was enjoyable to look at with a rate of 79.31%, where the daily HCD received “agree” for only 71.98% of the responses. There was no meaningful difference between comprehension of the state being represented, or ratings of understanding the information being portrayed or readability of the visualization.

4.4 Colorblindness

For this demographic of respondents, I separated the group into two sections: those with any form of color-blindness, and those with no colorblindness. I will provide an analysis of both samples below.

4.4.1 Colorblind Participants

There were 56 responses from individuals who responded that they were Red-Green, Blue-Yellow, or completely colorblind.

The results of the pre-visualization behavior questions reflect little differences between those who viewed the HCD visualization and the default visualization. For Q93 and Q71 asking about activities the respondent is comfortable participating in, there were no meaningful differences between levels of comfortability before and after viewing the visualizations. In Q98, which asks about how often the individual wears a mask before viewing the visualization, respondents answered “sometimes” at a rate of 37.50%, whereas after viewing the visualizations, they reported “sometimes” at a rate of 21.43%—a decrease of 16.07%. In Q131 and Q142, participants selected that they intentionally avoid contact with others “most” of the time at a rate of 37.5% before viewing the visualizations, and at a rate of 55.36% after viewing the visualizations—an increase of 17.86%.

The results of the comprehension questions for the colorblind (RG, BY, and complete) participants show that on Q135 and Q136, the cumulative HCD chart received 60% correct responses and the cumulative default chart received 30.77% correct responses—a difference of 29.23%. Furthermore, on question 163 and 164, the daily default chart received 38.46% correct responses, whereas the daily HCD chart received 53.33% correct responses.

For the questions about readability, visual appeal, and understanding information being portrayed, the default visualization was rated highest overall. The daily default chart was rated “easy” to read at a rate of 100% in comparison to the daily

HCD chart that received “easy” at a rate of 86.67%. Regarding visual appeal, participants reported that they enjoyed looking at the daily default visualization at a rate of 96.15%, whereas the daily HCD visualization received 79.31% of responses corresponding to enjoyment. In response to the questions 139 and 138 that ask about understanding the information being portrayed, participants who viewed the cumulative default visualization selected “agree” 100% of the time, whereas participants who viewed the cumulative HCD visualization selected agree 86.66% of the time. Although the comprehension rates for the HCD designed visualizations are significantly higher than the default visualizations, the colorblind participants rated the default designs as more appealing and easier to comprehend.

4.4.2 Non-Colorblind Participants

There were 313 participants who responded that they have no form of colorblindness. The results of the pre-visualization behavior questions reveal that participants reported increased levels of comfortability in nearly every activity before and after viewing the visualizations. Respondents showed an increase in levels of comfort when going to work or school; going to a market or pharmacy; going to a bar, restaurant, or café; spending time with someone they are not staying with; attending an event with more than ten people, and using public transit. Like many other demographics in this chapter, there was no meaningful difference between levels of comfortability when exercising in a fitness facility or outdoors before and after viewing the visualizations.

Results from the questions about social-distancing and mask behavior reflect an increased willingness to wear a mask before and after viewing the visualizations. Before looking at the visual, respondents reported that they “always” wore a mask at a rate of

56.23%, but responded that they would “always” wear a mask in the next three days at a rate of 62.30%. Furthermore, participants also responded that they are intentionally avoiding contact with other people “all” of the time at a rate of 23.32% before viewing the visualization, but responded they would avoid contact with others “all” of the time in the next three days at a rate of 29.71%.

The results of the comprehension questions reflect that the default visualizations received the highest rate of correct responses overall. In question 161 and 162, participants viewing the daily default visualization answered correctly at a rate of 85.81%, while those who viewed the daily HCD visualization answered correctly at a rate of 77.85%. The results of the remaining comprehension questions reflect that there is no meaningful difference between the comprehension rates of participants who viewed the HCD visuals and those who viewed the default visuals.

Participants reported that the cumulative default visualization was easiest to read and most enjoyable to look at overall. In question 66, participants rated the cumulative default chart as “easy” to read at a rate of 89.03%, whereas the cumulative HCD visualization was rated “easy” to read at a rate of 80.38%. Furthermore, participants reported that they enjoyed looking at the cumulative default chart at a rate of 81.94%, where participants responded that they enjoyed viewing the cumulative HCD visualization at a rate of 73.89%.

Chapter 5: Discussion and Recommendations

I will now discuss the above results, provide recommendations for producing effective visualizations, and offer ways to improve this research in future work.

5.1 Producing Effective Visuals

The results of the survey do not provide sufficient evidence to reject the null hypothesis, and suggest that there is room for improvement in the human-centered design process for producing visualizations. Across participant demographics, the human-centered designs were not shown to be significantly more effective or compelling than the default designs. In the cases of female, male, 18-34, 35+, and non-colorblind participants, the default designs had higher rates of comprehension, readability, and visual appeal.

There are a few potential factors as to why participants did not respond to the HCD visualizations with a higher rate of comprehension than the default visualizations. One factor is the tool I used to generate the visualizations. I decided to produce the visualizations using Microsoft Excel for its diversity of design elements, the quality of PNG exports for charts, and the default colors, fonts, and line sizes that provide a common representation of data visualizations. While this did provide an adequate foundation for producing the default visualizations, it was not effective in creating a human-centered design distinct from the default visualization. Excel uses slanted labels on the x-axis as the default design for their charts, despite Munzner's specific guidance against this positioning, yet this did not seem to prevent users who were presented these labels from successfully reading the graphs. The elements of the visualizations were not altered enough for the participants to show a meaningful change in their understanding

of the data. If I had chosen to generate the visualizations using more extensive technology, like D3.js, R, ArcGIS, or Python's datascience library, I could have produced interactive visualizations supported by Tufte and Munzer's recommendations that may have commanded the attention of my participants more effectively.

Another factor that contributed to these results is the testing environment of this survey. Due to this being a COVID-19 constrained virtual study, I could not enforce the device that they used to view the visualizations. Although my consent statement asked participants to use a laptop, desktop computer, or tablet, many of them could have been taking the survey on their phone. This variation of device sizes contributes to variation in the participant's ability to comprehend the visualizations. If this were an in-person study, all participants would take the survey on the same device so I could ensure that they viewed identical visualizations.

5.2 Changing Public Behavior

One component that affected willingness to change behavior is the timeliness of the datasets. My visualizations were produced in late December 2020, and the survey was not approved by the IRB until February 18, 2021. In December, the US was reaching some of its highest counts of new reported cases of the entire pandemic (The New York Times, 2021). By the time participants were taking the survey in February 2021, COVID cases had dropped and the risk of infection may have been perceived as diminishing. The rate of decline in cases could have impacted the way participants view their own behavior, and their decision to wear a mask and socially-distance in the future.

5.3 Future Research

It is important to note some ways that this research could be improved for future study.

There are several components of my research process that I would have changed if I had the opportunity to conduct this study again: generating interactive visualizations, dividing participants based on different data sets, designing the survey for use with Mechanical Turk, and completing more comprehensive statistical analysis.

In future work, I suggest researchers use more effective technology to produce interactive visualizations. As I mentioned in Chapter One, the visualization that inspired this research was an interactive article from the New York Times about the spread of water droplets from a person's sneeze. My lack of expertise in dynamic visualization generation and decision to use Excel limited my ability to produce these types of visuals. It would be interesting to see how D3.js, Python libraries, and R could be used to generate interactive visualizations that meet more of Tufte and Munzer's guidelines. This would result in visualizations that are significantly different than the default designs I produced with Excel, and may lead to higher rates of comprehension and willingness to change behavior.

It would be valuable to modify this survey to divide participants based on the ordering of the HCD and default designs, rather than the daily and cumulative data sets. I would show the participant the default images first, ask them questions to measure their understanding of the data, and then show them the human-centered design with the same questions from the first block. The results from this modified survey would show a difference in accuracy that would measure the readability of each design. If I included

the preference questions from the original survey, then it would also be a more accurate measure of which visuals were more pleasing to view.

Researchers should take advantage of Amazon's Mechanical Turk to distribute their online surveys, but be very intentional about the way they design the question set to capture and verify completion. Mechanical Turk recommends that those posting their surveys with Qualtrics should track their responses by either generating a unique code for the worker to upload to Amazon, or by entering their WorkerID directly into the survey itself. I chose to have the survey generate a random ID number to each participant. Although I followed the appropriate steps suggested by MTurk, several workers entered false ID numbers into Amazon and did not truly complete the survey. Because I didn't ask for any information from the workers, verifying that each worker had completed the survey involved comparing every random ID from Qualtrics with the random ID from Amazon. Furthermore, the survey was not closed for distribution in other sources, so there were several participants presented with a random ID that were not associated with Amazon. The verification process took a significant amount of time and prevented me from filtering the results on those from Mechanical Turk and those who were not. If I had asked the participants to enter their WorkerID, then I could have verified their response and successfully filtered results based on this value. My recommendation for survey designers is to ask participants for their WorkerID to avoid a lengthy verification process.

Future work in this area should include more comprehensive analysis of the survey results using an ANOVA test, Pearson's chi-squared test, and a two-sample t-test. My lack of familiarity with high level statistical analysis and time prevented me

from completing these analyses, but would highlight variation within the data that I did not present. The ANOVA test would determine whether the survey results are significant, especially among the questions that involved Likert scale responses. Pearson's chi-squared test provides an evaluation of categorical data with likelihood that any observed difference between the sets arose by chance. This would provide a more in-depth analysis of the comprehension questions to see which visualizations are more effective. The t-test determines if there is a significant difference between the means of two groups, and would be most effective in analyzing the questions about comfortability in participating in different activities.

With these statistical tests, I propose that researchers continuing this work also take advantage of the many demographic questions asked in this survey, and separate the participant pool using other filters. If I had time, I would conduct analyses of college and non-college educated, white participants and participants of color, Mechanical Turk workers and those not working for Mechanical Turk, and participants by region of the US based on ZIP code.

Conclusion

In summary, the results of this survey do not reject the null hypothesis. This means that there is no difference between the effectiveness of my human-centered design approach and the default design approach. However, one could draw from the results of this study that presenting participants with visualizations related to public health have the power to compel them to reconsider their future behavior. There are many ways to expand this research to determine the best way to visualize data for it to be more persuasive, effective, and pleasing to view, but this study provides a foundation for measuring the effectiveness of human-centered data visualization methods.

Appendix

Glossary

Amazon Mechanical Turk: Amazon Mechanical Turk is a crowdsourcing website for businesses to hire remotely located "crowdworkers" to perform discrete on-demand tasks that computers are currently unable to do. It is operated under Amazon Web Services, and is owned by Amazon. (Wikipedia Contributors, 2019)

Comprehension: the capacity for understanding fully (*Definition of COMPREHENSION*, n.d.)

Data visualization (data vis): Models that provide visual representations of datasets designed to help people carry out tasks more effectively (Munzner, 2015).

Direct contact: A conversation lasting more than five minutes with a person who is closer than 6 feet from you without either person wearing a mask.

Graphical excellence: Tufte's guidelines for producing effective visualizations as written in *The Visual Display of Quantitative Information*. See chapter 2.4 about current methods in human-centered design.

Human-Centered Design (HCD): An approach to problem solving, commonly used in design and management frameworks that develops solutions to problems by involving

the human perspective in all steps of the problem-solving process (Wikipedia Contributors, 2019).

Human-subjects research: Research involving "a living individual about whom an investigator (whether professional or student) conducting research: obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens (Definition of Human Subjects Research | grants.nih.gov, n.d.).

Institutional Review Board (IRB): Under FDA regulations, an Institutional Review Board is a group that has been formally designated to review and monitor biomedical research involving human subjects. In accordance with FDA regulations, an IRB has the authority to approve, require modifications in (to secure approval), or disapprove research. This group review serves an important role in the protection of the rights and welfare of human research subjects (Center for Drug Evaluation and Research, 2019).

Repository: This is a folder where files are stored inside a project on Github, a version control platform.

Social distancing: The practice of staying 6 feet away from those around you.

Survey Materials

Question Breakdown

In the last 24 hours, have you done any of the following? Please select all that apply.

Physically gone to work or school outside the place where you are currently staying

Gone to a market, grocery store, or pharmacy

Gone to a bar, restaurant, or cafe

Spent time with someone who isn't currently staying with you

Attended an event with more than 10 people

Used public transit

Exercised in a fitness facility

Exercised outdoors

None of the above

Prefer not to say

Figure SM1: Q62

During which activities **in the past 24 hours** did you wear a mask? Please select all that apply.

Gone to work or school outside the place where you are currently staying
Gone to a market, grocery store, or pharmacy
Gone to a bar, restaurant, or cafe
Spent time with someone who is currently staying with you
Exercised in a fitness facility
Exercised outdoors

Figure SM2: Q63

In the past 24 hours, with how many people have you had direct contact with **outside of your household**? Your best estimate is fine.

"Direct contact" means a conversation lasting more than 5 minutes with a person who is closer than 6 feet from you without either person wearing a mask.

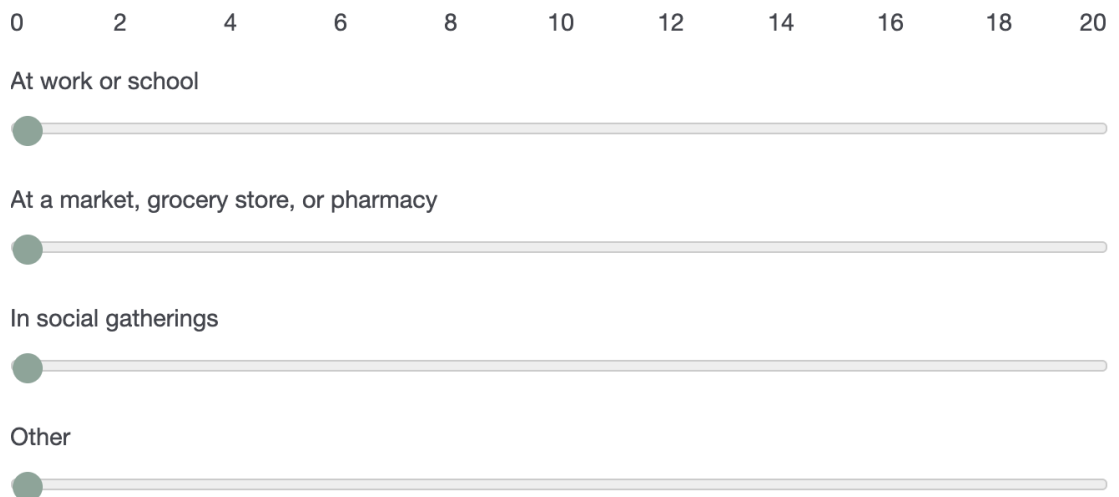


Figure SM3: Q97

In the past 3 days, when out in public spaces where social distancing is not possible, about how many people wore a mask?

"Social distancing" means staying 6 feet away from those around you.

All of the people were wearing masks
Most of the people were wearing masks
Some of the people were wearing masks
A few of the people were wearing masks
None of the people were wearing masks
I have not been in public places in the past 3 days

Figure SM4: Q130

Complete the following sentence: "In the past two weeks, I feel..."

	Very comfortable	Somewhat comfortable	Somewhat uncomfortable	Very uncomfortable	Not applicable
Going to work or school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Going to a market, grocery store, or pharmacy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Going to a bar, restaurant, or cafe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spending time with someone who I'm not staying with	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Attending an event with more than 10 people	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Using public transit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exercising in a fitness facility	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exercising outdoors	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure SM5: Q93

In the past 3 days, how often did you wear a mask when in public spaces where social distancing is not possible?

"Social distancing" means staying 6 feet away from those around you.

Always
Frequently
Sometimes
Rarely
Never
I have not been in public spaces in the past 3 days

Figure SM6: Q98

To what extent are you intentionally avoiding contact with other people?

All of the time: I always avoid contact with others.
Most of the time: I only leave home to purchase food and other essentials.
Some of the time: I have reduced the amount of times I am in public spaces, social gatherings, or at school.
None of the time: I am often in physical contact with others.

Figure SM7: Q131

Please use the information above to answer the following questions.

Which state is being represented?

Figure SM8: Q50 / Q165 / Q122 / Q166

What is the single highest recorded number of cases in Florida?

2,100,000

1,520,000

1,170,000

750,000

Figure SM9: Q135 / Q136

On **July 7th**, approximately how many cases of COVID-19 were recorded in Florida?

210,000

150,000

400,000

350,000

Figure SM10: Q134 / Q137

Complete the sentence: "Reading this chart was..."

Very easy

Easy

Difficult

Very difficult

Figure SM11: Q66 / Q170 / Q124 / Q173

Respond to the statement: "I enjoyed looking at this chart."

Strongly agree

Somewhat agree

Somewhat disagree

Strongly disagree

Figure SM12: Q145 / Q171 / Q146 / Q174

Respond to the statement: "I understand the information being portrayed in this graph."

Strongly agree

Somewhat agree

Somewhat disagree

Strongly disagree

Figure SM13: Q139 / Q172 / Q138 / Q175

What is the single highest recorded number of **daily cases** in Florida?

18,000
16,800
12,500
11,300

Figure SM14: Q161 / Q162

On **September 25**, approximately how many daily cases of COVID-19 were recorded in Florida?

4,300
1,200
2,800
4,100

Figure SM15: Q163 / Q164

Complete the following sentence: "When **every** person is wearing a mask, I feel..."

	Very comfortable	Somewhat comfortable	Somewhat uncomfortable	Very uncomfortable	Not applicable
Going to work or school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Going to a market, grocery store, or pharmacy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Going to a bar, restaurant, or cafe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spending time with someone who I'm not staying with	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Attending an event with more than 10 people	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Using public transit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exercising outdoors	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exercising in a fitness facility	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure SM16: Q71

In the next 3 days, how often do you plan to wear a mask?

Always
Frequently
Sometimes
Rarely
Never
I will not go in public in the next 3 days

Figure SM17: Q100

In the next 3 days, to what extent will you intentionally avoid contact with other people?

All of the time: I will always avoid contact with others.
Most of the time: I will only leave home to purchase food and other essentials.
Some of the time: I will reduce the amount of times I am in public spaces, social gatherings, or at school.
None of the time: I will be in physical contact with others often.

Figure SM18: Q142

What is your current ZIP code?

--

Figure SM19: Q52

Do you have any known color blindness?

Yes, red-green
Yes, blue-yellow
Yes, complete
No
Prefer not to say

Figure SM20: Q88

What is your gender?

Female
Male
Non-binary
Prefer to self-describe
Prefer not to say

Figure SM21: Q54

What is your age?

18-24 years
25-34 years
35-44 years
45-54 years
55-64 years
65-74 years
75 years or older
Prefer not to say

Figure SM22: Q55

Are you of Hispanic, Latino, or Spanish origin?

Yes
No, not of Hispanic, Latino, or Spanish origin
Prefer not to say

Figure SM23: Q56

What is your race?

American Indian or Alaska Native
Asian
Black or African American
Native Hawaiian or other Pacific Islander
White
Some other race
Prefer not to say

Figure SM24: Q57

What is the highest degree or level of school you have completed?

Less than high school
High school graduate or equivalent (GED)
Some college
2 year degree
4 year degree
Professional degree
Doctorate
Prefer not to say

Figure SM25: Q58

For MTurk Workers:

This is your ID: 6031

Copy this value to paste into MTurk.

Figure SM26: Random ID

Recruitment Statement

The following statement was posted during the recruitment stage of my survey distribution. I sent the statement with a link to the survey on LinkedIn, Facebook, and through the UO computer science undergraduate and graduate student listserv. It was

not attached to the distribution post on Mechanical Turk because Amazon's system put the link in their catalog independently.

Hello! My name is Stephanie Schofield, and I'm a senior Computer Science student in the Clark Honors at the University of Oregon. I am conducting research about the effectiveness of data visualization methods using COVID-19 case data. The research is conducted in a survey that should take about 15 minutes. This study has been reviewed by an independent institutional review board, and is not expected to pose a risk to you. You are eligible to participate in the survey if you are at least 18 years old, and have the ability to take the survey on a laptop, desktop computer, or tablet. If you choose to participate, your responses will remain strictly anonymous. Any data which is presented will be reported in aggregate.

Link to survey:

https://oregon.qualtrics.com/jfe/form/SV_dmbWguXGatb5Ekl

Your input is very important, and will help us understand how to create visualizations that are clear, impressionable, and have lasting effects on a person's behavior. If you have any questions, please feel free to contact me: sschofie@cs.uoregon.edu. Thank you for your time!

Survey Information Statement

The following statements were attached at the beginning of the survey and introduced the survey's purpose and scope. There were two statements prepared: one for Mechanical Turk participants, and the other for UO students and those not associated with Amazon. For live or in-person human-subject research, a consent statement is required in order to maintain ethical research practices according to the Declaration of Helsinki (World Medical Association, 2014). However, remote survey human-subject research requires a survey information statement. Writing and including this statement is a key factor in human-subject research, and was necessary in order to receive

approval from the UO Institutional Review Board (IRB). The statement for UO students and non-MTurk Workers reads:

This voluntary survey is part of a research study led by Stephanie Schofield at the Robert D. Clark Honors College and University of Oregon Computer Science department. Your responses will help us understand how to produce effective visualizations for public health data.

This survey is intended to take roughly 15 minutes to complete, and **must be taken on a laptop, desktop computer, or tablet**. The survey questions do not request any unique or personally identifiable information about you and your answers to all questions will remain confidential.

Your responses may be shared with other researchers studying data visualization and human-computer interaction. We may also publish aggregate tables of results for public research use. Results published will be in aggregate and will not identify individual participants of their responses. There are no foreseeable risks in participating and no compensation is offered.

If you are taking this survey as a worker on Amazon's Mechanical Turk, you will be paid \$3.75 for your participation. The survey should not take longer than 15 minutes.

If you are taking this survey as a SONA participant, you will be awarded 0.25 credits for your participation. If you choose to discontinue participation in this survey at any point after clicking through this consent page, you will receive 1/4 credit for each 15 minutes of participation, rounded up to the next 15 minutes. For example, if you complete 1-15 minutes you will receive 1/4 credit, if you complete 16-30 minutes you will receive 1/2 credit, and so on. If you discontinue participating in the middle of the study, contact the listed researcher to receive partial credit. This research was reviewed by the University of Oregon Institutional Review Board.

Your participation is voluntary. Your decision whether or not to participate will not affect your relationship with the UO Psychology Department or the UO Linguistics Department. If you decide to participate, you are free to withdraw your consent and discontinue

participating at any time without penalty. The Psychology and Linguistics Departments have established alternative assignments for students who do not wish to participate as research subjects. Please see your instructor if you would rather complete an alternative assignment.

If you have any questions, please contact Stephanie Schofield at sschofie@cs.uoregon.edu.

The statement for MTurk Workers reads:

This voluntary survey is part of a research study led by Stephanie Schofield at the Robert D. Clark Honors College and University of Oregon Computer Science department. Your responses will help us understand how to produce effective visualizations for public health data.

This survey is intended to take roughly 15 minutes to complete, and must be taken on a laptop, desktop computer, or tablet. The survey questions do not request any unique or personally identifiable information about you and your answers to all questions will remain confidential.

If you are taking this survey as a worker on Amazon's Mechanical Turk, you will be paid \$3.75 for your participation. If you decide to participate, you are free to withdraw your consent and discontinue participating at any time without penalty.

Your responses may be shared with other researchers studying data visualization and human-computer interaction. We may also publish aggregate tables of results for public research use. Results published will be in aggregate and will not identify individual participants of their responses. There are no foreseeable risks in participating. This research was reviewed by the University of Oregon Institutional Review Board.

If you have any questions, please contact Stephanie Schofield at sschofie@cs.uoregon.edu

Charts

There were four visualizations that participants could potentially view while taking the survey, which were created using two datasets: daily cases and cumulative cases of

COVID-19 across the state of Florida from March 1 to December 19, 2020. The two datasets were represented in the form of two separate line charts. Each of the two charts for the respective datasets were created using a Human-Centered Design (HCD) methodology, and a default methodology. The four different charts can be seen below.

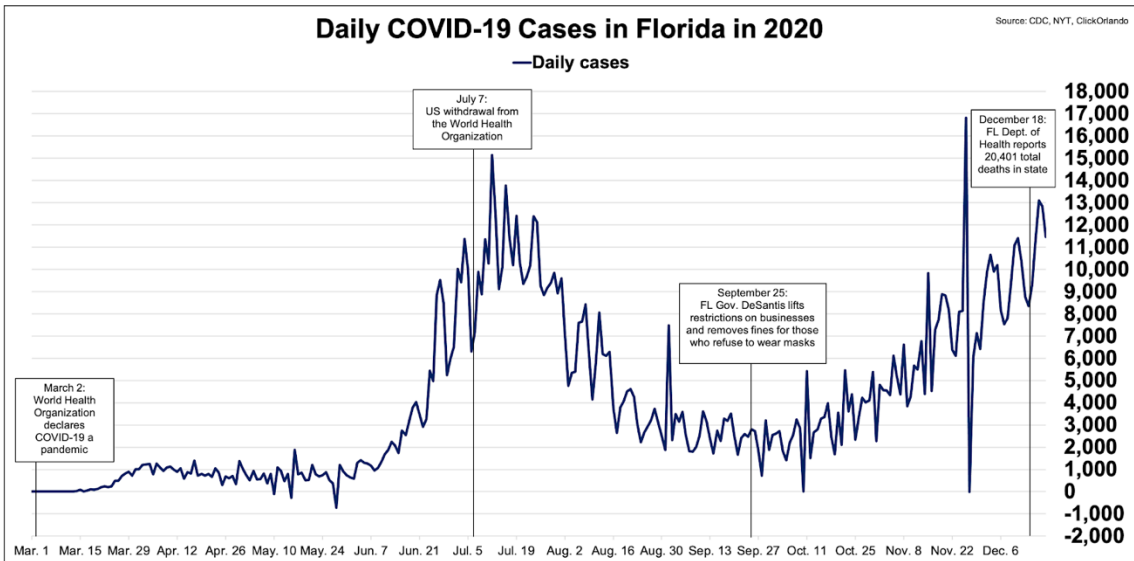


Figure C1: Human-Centered Daily Cases Visualization

Daily cases of COVID-19 across the state of Florida over March to December of 2020, created with a Human-Centered Design methodology.

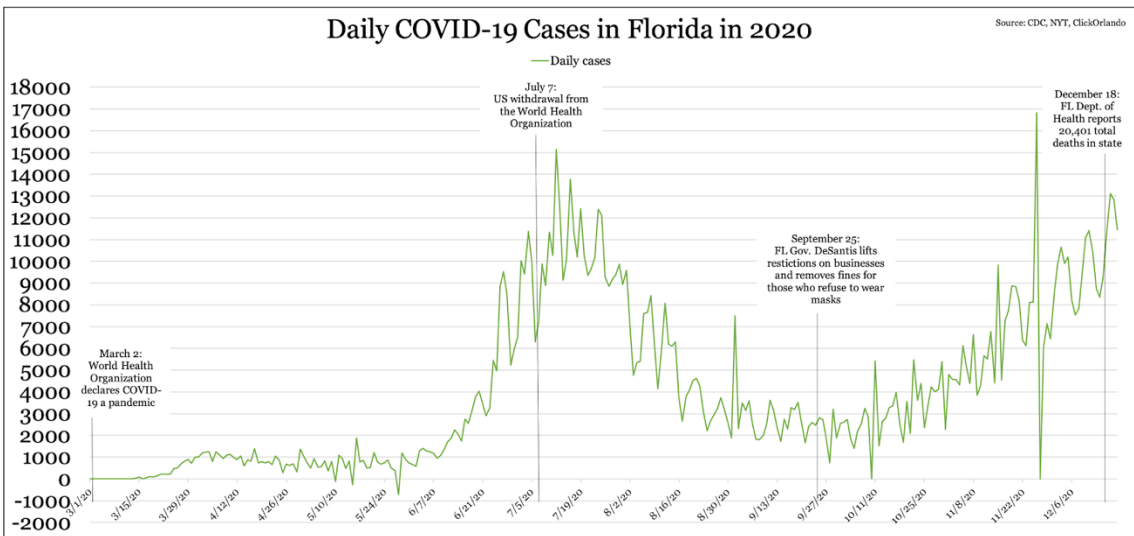


Figure C2: Default Daily Cases Visualization

Daily cases of COVID-19 across the state of Florida over March to December of 2020, created using a default approach.

Participants had a 50% chance of viewing either the HCD or the default visualizations, but were asked the same questions regardless of which chart they viewed.

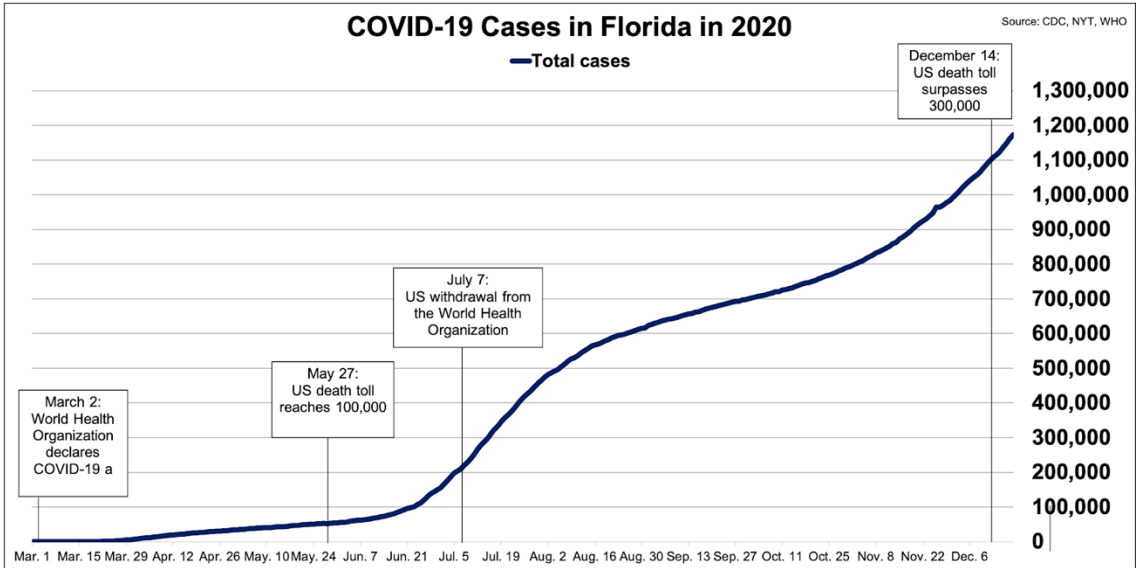


Figure C3: Human-Centered Cumulative Cases Visualization

Cumulative cases of COVID-19 across the state of Florida from March to December of 2020, created with a Human-Centered Design methodology.

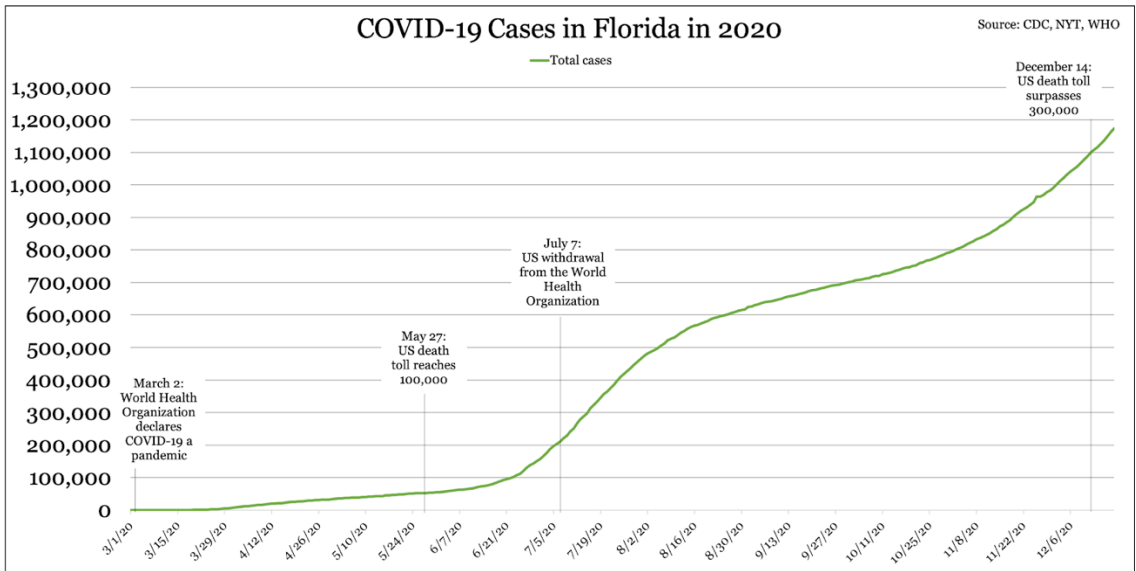


Figure C4: Default Cumulative Cases Visualization

Cumulative cases of COVID-19 across the state of Florida from March to December of 2020, created using a default approach.

Debriefing Statement

The following statement was attached to the last block of the survey to inform the participant of what the survey is measuring. This debriefing is another important aspect of conducting ethical human-subjects research.

Thank you for your participation!

Background: Today's study examined the correlation between comprehension of data visualizations and willingness to wear a mask and social distance. While a lot of research is done into the efficacy of charting public health data, there isn't as much research about how these charts affect people's behavior after viewing them.

Purpose of this study: We hope to determine which kinds of charts are most impressionable, clear to understand, and have lasting effects on a person's behavior. These results will help data visualization researchers create more effective graphics for people to understand large, important data sets. Furthermore, we hope this survey showed participants the risks posed from not wearing a mask, and leads to a higher rate of social-distancing and mask-wearing behavior in young adults most prone to spreading Coronavirus.

Your part: The part you play in this research is very important! Just giving us an idea of which visualizations were difficult or easy to read, some of your social distancing behavior, and your willingness to continue or change those habits will provide great insight. Thank you for contributing your valuable time to this survey today.

Feedback and further information: If you have additional questions about this study, please feel free to ask the experimenter, Stephanie Schofield at sschofie@cs.uoregon.edu, 541-704-8191, or her advisors, Nicole Dudukovic, ndudukov@uoregon.edu, Department of Psychology; Joe Sventek, jsventek@uoregon.edu, Department of Computer Science. If you have any questions concerning your rights as a research participant, please contact Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, 541-346-2510, or email ResearchCompliance@uoregon.edu. You can also email the Human Subjects Coordinator for psychology and linguistics research at hscoord@uoregon.edu.

Notice of Amendment Review and Exempt Determination



DATE: February 18, 2021 **IRB Protocol Number: 01052021.005**

TO: Stephanie Schofield, Principal Investigator
Computer Science

RE: Protocol entitled, "Perception is Reality: Measuring the Effectiveness of Human-Centered Data Visualization Methods in Young Adults"

Notice of Review and Exempt Determination

The above protocol has been reviewed and determined to qualify for exemption. The research is approved to be conducted as described in the attached materials. Any change to this research will need to be assessed to ensure the study continues to qualify for exemption, therefore an amendment will need to be submitted for verification prior to initiating proposed changes.

For this research, the following determinations have been made:

- This study has been reviewed under the 2018 Common Rule and determined to qualify for exemption under Title 45 CFR 46.104(d)(2).

Approval period: February 18, 2021 - February 28, 2022

If you anticipate the research will continue beyond the approval period, you must submit a Progress Report at least 45-days in advance of the study expiration. **Without continued approval, the protocol will expire on February 28, 2022 and human subject research activities must cease.** A closure report must be submitted once human subject research activities are complete. Failure to maintain current approval or properly close the protocol constitutes non-compliance.

You are responsible for the conduct of this research and adhering to the Investigator Agreement as reiterated below. You must maintain oversight of all research personnel to ensure compliance with the approved protocol.

The University of Oregon and Research Compliance Services appreciate your commitment to the ethical and responsible conduct of research with human subjects.

Sincerely,

A handwritten signature in black ink, appearing to read 'Russell Melia'.

Russell Melia
Research Compliance Administrator

CC: Nicole Dudukovic Kuhl

COMMITTEE FOR THE PROTECTION OF HUMAN SUBJECTS • RESEARCH COMPLIANCE SERVICES
677 E. 12th Ave., Suite 500, 5237 University of Oregon, Eugene OR 97401-5237
T 541-346-2510 F 541-346-5138 <http://rcs.uoregon.edu>

An equal-opportunity, affirmative-action institution committed to cultural diversity and compliance with the Americans with Disabilities Act

Figure N1: IRB Exempt Determination

Initial IRB Exemption for the study was received on February 18, 2021, and allowed me to move forward with distributing the survey to participants.



DATE: April 14, 2021 **IRB Protocol Number: 01052021.005**
TO: Stephanie Schofield, Principal Investigator
Department of University Housing
RE: Protocol entitled, "Perception is Reality: Measuring the Effectiveness of Human-Centered Data Visualization Methods in Young Adults"

Notice of Amendment Review and Exempt Determination

The amendment submitted on April 05, 2021 to the above protocol has been reviewed and determined to continue to qualify for exemption. Any other change to this research will need to be assessed via a separate amendment to ensure the study continues to qualify for exemption. The research is approved to be conducted as described in the attached materials.

The purpose of this Amendment is to:

- Add MTurk as a recruitment method.
- Associate a new funding to the protocol.

For this research, the following determinations have been made:

- This study has been reviewed under the **2018 Common Rule** and determined to qualify for exemption under Title 45 CFR 46.104(d)(2).

Approval period: April 14, 2021 - February 28, 2022

If you anticipate the research will continue beyond the approval period, you must submit a Progress Report at least 45-days in advance of the study expiration. **Without continued approval, the protocol will expire on February 28, 2022 and human subject research activities must cease.** A closure report must be submitted once human subject research activities are complete. Failure to maintain current approval or properly close the protocol constitutes non-compliance.

You are responsible for the conduct of this research and adhering to the *Investigator Agreement* as reiterated below. You must maintain oversight of all research personnel to ensure compliance with the approved protocol.

The University of Oregon and Research Compliance Services appreciate your commitment to the ethical and responsible conduct of research with human subjects.

Sincerely,

Yubin Cho
Research Compliance Administrator

CC: Nicole Dudukovic Kuhl

Figure N2: IRB Amendment Review

The IRB approved my amendment application to use Amazon’s Mechanical Turk on April 14, 2021, and allowed me to distribute the survey to Mechanical Turk workers

References

- CDC Case Task Force. (2021, April 29). *United States COVID-19 Cases and Deaths by State over Time | Data | Centers for Disease Control and Prevention*. Data.cdc.gov; Centers for Disease Control and Prevention. <https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36>
- Center for Drug Evaluation and Research. (2019). *(IRBs) and Protection of Human Subjects*. U.S. Food and Drug Administration. <https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/institutional-review-boards-irbs-and-protection-human-subjects-clinical-trials>
- Clark, J. (2011). Let's reflect: what is the point? *British Journal of General Practice*, 61(593), 747–747. <https://doi.org/10.3399/bjgp11x613232>
- Creative research systems. (2012). Sample Size Calculator - Confidence Level, Confidence Interval, Sample Size, Population Size, Relevant Population - Creative Research Systems. Surveysystem.com. <https://www.surveysystem.com/sscalc.htm>
- CSSLOhioStateU. (2014). Your Survey Closed, Now What? Quantitative Analysis Basics. In *YouTube*. <https://www.youtube.com/watch?v=4q5UZwwidRI>
- Definition of COMPREHENSION. (n.d.). Wwww.merriam-Webster.com. <https://www.merriam-webster.com/dictionary/comprehension>
- Definition of Human Subjects Research | grants.nih.gov. (n.d.). Grants.nih.gov. <https://grants.nih.gov/policy/humansubjects/research.htm>
- Desjardins, J. (2019, April 17). *How much data is generated each day?* World Economic Forum. <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/>
- etanrioven. (2015, March 8). *Bill Gates Graph of the Year – Causes of Untimely Death – Class comments*. Nehirege. <https://nehirege.wordpress.com/2015/03/08/bill-gates-graph-of-the-year-causes-of-untimely-death-class-comments/>
- Katz, J., Sanger-Katz, M., & Quealy, K. (2020, July 17). A Detailed Map of Who Is Wearing Masks in the U.S. *The New York Times*. <https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html>
- Koop, A. (2021, March 27). *Visualized: The Richest Families in America*. Visual Capitalist. <https://www.visualcapitalist.com/visualized-the-richest-families-in-america/>

- Kress, G. R., & Theo Van Leeuwen. (1996). *Reading images: The Grammar of Visual design*. Routledge.
- Moraes Bueno Rodrigues, A., Diniz Junqueira Barbosa, G., Côrtes Vieira Lopes, H., & Diniz Junqueira Barbosa, S. (2021). What questions reveal about novices' attempts to make sense of data visualizations: Patterns and misconceptions. *Computers & Graphics*, 94(0097-8493), 32–42. ScienceDirect. <https://doi.org/10.1016/j.cag.2020.09.015>
- Munzner, T. (2015). *Visualization Analysis & Design*. Crc Press, Taylor & Francis Group.
- New York Times. (2020, July 28). *nytimes/covid-19-data*. GitHub. <https://github.com/nytimes/covid-19-data/tree/master/mask-use>
- New York Times. (2021, April 15). Florida Coronavirus Map and Case Count. *The New York Times*. <https://www.nytimes.com/interactive/2021/us/florida-covid-cases.html>
- Parshina-Kottas, Y., Saget, B., Patanjali, K., Fleisher, O., & Gianordoli, G. (2020, April 14). This 3-D Simulation Shows Why Social Distancing Is So Important. *The New York Times*. <https://www.nytimes.com/interactive/2020/04/14/science/coronavirus-transmission-cough-6-feet-ar-ul.html>
- Rawsthorn, A. (2006, July 9). Quirky serifs aside, Georgia fonts win on Web - Style - International Herald Tribune. The New York Times. <https://www.nytimes.com/2006/07/09/style/09iht-dlede10.2150992.html>
- Semuels, A. (2018, January 23). *The Online Hell of Amazon's Mechanical Turk*. The Atlantic; The Atlantic. <https://www.theatlantic.com/business/archive/2018/01/amazon-mechanical-turk/551192/>
- The New York Times. (2021, May 16). Coronavirus in the U.S.: Latest Map and Case Count. The New York Times. https://www.nytimes.com/interactive/2021/us/covid-cases.html?pageType=LegacyCollection&collectionName=Maps+and+Trackers&label=Maps+and+Trackers&module=hub_Band@ion=inline&template=storyline_band_recirc
- Tufte, E. R. (1985). *The Visual Display of Quantitative Information*. Graphics Press.
- US Census Bureau. (2020, May 20). *2018 FIPS Codes*. The United States Census Bureau. <https://www.census.gov/geographies/reference-files/2018/demo/pepest/2018-fips.html>

- Venables, W. N., & Ripley, B. D. (2001). *Modern applied statistics with S- PLUS*. Springer, Cop.
- Wallach, O. (2021, March 3). *Which Streaming Service Has the Most Subscriptions?* Visual Capitalist. <https://www.visualcapitalist.com/which-streaming-service-has-the-most-subscriptions/>
- Wikipedia Contributors. (2019, January 31). *Human-centered design*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Human-centered_design
- Wikipedia Contributors. (2019, March 8). *Amazon Mechanical Turk*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Amazon_Mechanical_Turk
- Wilson, Dr. R. F. (2001, March 1). *HTML E-Mail: Text Font Readability Study*. Practical Ecommerce. <https://www.practicalecommerce.com/html-email-fonts>
- Wonkborg. (2013, December 23). *Bill Gates's graph of the year*. *Washington Post*. <https://www.washingtonpost.com/news/wonk/wp/2013/12/27/bill-gatess-graph-of-the-year/>
- World Medical Association. (2014). *WMA - The World Medical Association-WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*. Wma.net; *WMA - The World Medical Association-WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>