DATA SCIENTISTS' EPISTEMIC THINKING FOR CREATING AND INTERPRETING VISUALIZATIONS

CHARLOTTE A. BOLCH
Midwestern University
cbolch@midwestern.edu

KENT J. CRIPPEN
University of Florida
kcrippen@coe.ufl.edu

ABSTRACT

The purpose of the study was to understand the experiences of data scientists regarding common skills and strategies for interpreting and creating data visualizations. In this study, the participants were researchers in data science. The Delphi method was used to gather common processes of data visualization through three rounds of surveys called Delphi panels where responses from the previous panel were used to frame the questions on the next panel. Skills and strategies were identified after Delphi Panel 1 and then brought back to the participants in Delphi Panel 2 to rate the level of importance they attributed to those skills/strategies. Consensus was determined using a cut-off for the interquartile range for each skill/strategy, and overall group ratings were presented to researchers in Delphi Panel 3 for them to adjust their ratings as desired. This study provided empirical evidence for a consensus set of skills/strategies that data scientists engage in when interpreting and creating visualizations.

Keywords: *Statistics education research; Data visualization; Epistemic thinking; Data science*

1. INTRODUCTION AND BACKGROUND LITERATURE

1.1. DATA SCIENCE AND DATA VISUALIZATION

Data science is an evolving field that encompasses the use of data to extract value and meaning from complex and rich datasets that have been amassed from science, industry, or government (De Veaux et al., 2017; Takemura, 2018). The new research area, data science, was first introduced by Cleveland (2001) in an action plan for the field of statistics, which described how the technical areas of statistics needed to be redesigned for practicing analysts to learn from data by using various tools that have a direct benefit while understanding statistical theories that have an indirect benefit. In recent years, technological advances have had a universal impact on the ability of various industries and research fields to collect, store, visualize, and analyze large amounts of data (Eilam, 2015; Forbes et al., 2014). As a result, visualization software tools have developed over the past ten years to present complex data and information in a more succinct way so as to foster interpretation and discovery of new relationships or patterns (Forbes et al., 2014; Keim et al., 2010; Mirel et al., 2016). The use of visualization methods and techniques to make sense of large datasets has become an essential way to represent and interpret information (Figueiras, 2013). Data visualization has further developed as a process for representing information to facilitate understanding, identify trends and patterns, and make inferences about data (Kapler & Wright, 2004). This has created the need to understand the components of data visualization literacy and encourage universal data visualization literacy in society (Börner et al., 2019).

The basic idea of data visualization, namely, making sense of graphical representations of data, has been well studied across statistics, mathematics, and science education in multiple ways using various names and constructs (Cooper & Shore, 2010; Pfannkuch, 2006; Maltse et al., 2015; Mirel et al., 2016; Roth & Bowen, 2001). One prominent and pertinent approach from the perspective of statistical literacy

Statistics Education Research Journal, 21(2), Article 11. https://doi.org/10.52041/serj.v21i2.21

emphasizes graph comprehension (Curcio, 1987; Friel et al., 2001). Research of this type is typically based upon the rationale that standard statistical graphs such as line, bar, and histogram, are typically taught in introductory courses for statistics, mathematics, and science, and thus are viewed as fundamental displays that people should be able to read and understand in order to be literate citizens. For example, Curcio (1987) introduced three levels of graphical comprehension, which he defined as "read the data", "read between the data", and "read beyond the data". Friel et al. (2001) furthered Curcio's framework and identified skills at each level. The authors also brought together many perspectives and literature that identified critical cognitive behaviors that influence student graphical comprehension, including the concept of graph sense and proposed instructional implications based on those behaviors. Graph sense is defined as a combination of the process of graph comprehension and the importance of emerging technology on allowing graphical displays to be created in dynamic environments. The behaviors associated with graph sense are closely aligned with what is recognized as data visualization in that they encompass the reasoning and interpretation of a graphical display as cognitive abilities in ways that are not captured with the contrast of graphical comprehension.

Graphical comprehension has also been studied from other perspectives such as a focus on cognitive ability (Lem et al., 2015; Lowrie et al., 2012; Nicolaou et al., 2007), the graph as a singular entity (Lima & Selva, 2010; Shah & Hoeffner, 2002), or a student's ability to translate information between multiple different modes (Carrión Pérez & Espinel Febles, 2006; Kosslyn, 1985). For example, Roth and Bowen (2001) investigated graph-related practices based on scientists reading and interpreting familiar and unfamiliar graphs and found that they used their understanding of the phenomena and context represented by the graph to interpret signs or patterns in the graph. Strategies have also been identified for assisting primary and secondary school teachers in developing their reasoning and thinking about how to communicate purposeful comparative reasoning from graphs focused on the statistical inquiry cycle (Pfannkuch et al., 2010).

In contrast to graphical comprehension, more recent studies using the construct of data visualization have been varied in focus and approach. For example, Azzam and Evergreen (2013) define data visualization as a process of creating a representation that consists of the following three criteria: (a) the data must be qualitative or quantitative, (b) the raw data is accurately represented and important information is not omitted, and (c) the data can be explored, examined, and communicated. Boy and colleagues (2014) used the term and concept for the actual display of data that is produced from the process while Börner and colleagues (2016) defined data visualization as a process of making sense about data by applying higher-order thinking, such as reasoning, synthesis, or evaluation. In a similar way, Laina and Wilkerson (2016) concluded from a case study analysis that reasoning with complex data visualizations may require learners to be fluid about their interpretations requiring reorganizing data in interactive displays and adjusting interpretations to consider multiple patterns within the data. These various perspectives and studies all share a common recognition, to one degree or another, for the central role of creation and interpretation of a representation, often identified as a data visualization or simply a visualization, as the two central thinking processes that encompass data visualization literacy. This situation, in the context of data science emerging and evolving as a field, served as the impetus for a study on the perspective of data scientists on their thinking processes.

The purpose of this study was to use the consensus perspective of data scientists as experts to construct a rich definition of data visualization literacy by capturing the skills and strategies of the thinking processes as they both create and interpret visualizations. Accordingly, the study was framed by the following two research questions:

1) What common skills and strategies define interpreting visualizations based on the experiences from data scientists?
2) What common skills and strategies define creating visualizations based on the experiences from data scientists?

1.2. THEORETICAL FRAMEWORK

The problem of understanding how data scientists interpret and create visualizations was approached through the framework of epistemic thinking, which is also recognized as epistemic cognition. The terms epistemic or epistemological are derivations of the construct epistemological beliefs, which refers to a person's "beliefs about the nature of knowledge and the process of knowing" (Mason, 2016, p. 375). Suggesting that something is epistemic implies that it involves a person's personal beliefs and assumptions about the source and/or justification of knowledge (Mason, 2016). With regards to the process of knowing, cognition has been defined as the "mental action or process of acquiring knowledge and understanding through thought, experience, and the senses" (Cambridge Cognition, 2015, p. 1). Cognition can also be understood as the mental processes that an individual has as they learn and store information along with how that information is used to influence their behavior (Cambridge Cognition, 2015).

The field of study about the nature of knowledge is personal epistemology that focuses on how individuals reflect on and perceive the aspects of knowledge and knowing (Barzilai & Zohar, 2012). The term epistemology derives from the Greek words, "episteme" and "logos", meaning the "study of knowledge". In the past, the research about epistemology was focused on defining what knowledge is and differentiating the term from personal opinions and perspectives (Sandoval et al., 2016). Traditional epistemology research centered on understanding how individuals would have gained knowledge not just by chance or luck, but by justifying for their knowledge. It was not until the late 1990's that the field of personal epistemology was developed, influenced by an awareness that individuals' discernments of using and applying knowledge claims were complex (Hofer & Pintrich, 1997; Sandoval et al., 2016). The study of personal epistemology has been researched from three primary perspectives: as a cognitive developmental process, as a system of beliefs, and as a set of resources (Barzilai & Zohar, 2012, 2014; Hofer, 2004).

The combination of epistemic and cognition into epistemic cognition or epistemic thinking indicates a focus on how individuals think about what they have come to understand as knowledge, as opposed to believing or doubting that something is acceptable. Epistemic cognition involves mental structures related to knowledge and acquiring true beliefs and understanding (Barzilai & Eilam, 2018). Epistemic cognition also focuses on understanding how an individual acquires, justifies, and uses the knowledge that they possess (Hofer, 2016). Understanding how individuals evaluate the degree of commitment to the information and claims, as well as the sources of that information combined with the strategies and processes for reasoning about specific information, are all components of epistemic cognition (Barzilai & Zohar, 2016). An example of epistemic cognition would be an individual examining how reliable a source of information is as a means to justify and achieve acceptable belief of that information.

Epistemic thinking can operate at both the cognitive and metacognitive levels, but the distinction between the two aspects have not been well defined in the field of personal epistemology (Barzilai & Zohar, 2012, 2014). The multifaceted framework of epistemic thinking by Barzilai and colleagues (2014) makes a distinction between cognition and metacognition when exploring learners' epistemic thinking. The framework further delineates among the knowledge, skills, and experiences within epistemic metacognition. There is a gap within the data visualization literacy literature that this study addressed by using the framework of epistemic metacognitive knowledge to better understand the skills and strategies data scientists engage in when interpreting and creating a visualization.

The process of data visualization involves epistemic thinking because the individual is using mental structures to process and understand the visualization as well as producing beliefs about the topic of information the data represents. Understanding the mental structures of data scientists regarding data visualization will assist in defining the skills and strategies of the thinking processes that individuals on the high-end of the data visualization literacy continuum use when interpreting and creating visualizations. By asking experts to make their epistemic thinking explicit, to reflect upon it individually and in relation to the thinking of other experts, we specifically target what is recognized as epistemic metacognitive knowledge (Barzilai & Zohar 2012, 2014).

1.3. DATA VISUALIZATION LITERACY

Data visualization has been researched in a variety of disciplines, including Library Science, Engineering, Information Science, STEM education, and Computer Science/Human-Centered Computing. Consequently, data visualization literacy has been interpreted and defined in many ways relative to the context in which it was studied. In the area of library science, data visualization is defined as "enabl[ing] the creation of engaging, aesthetically [pleasing] representations of data and aggregate statistics" (Bouquin & Epstein, 2015, p. 350). This research focused on providing librarians with data visualization tips and techniques that would assist them in creating marketing tools and to help them feel more comfortable interacting with various topics.

In the research area of computer graphics and visualization in engineering, there have been a few studies that have focused on the test development process for assessing data visualization literacy. Boy et al. defined data visualization literacy as “the ability to use well-established data visualizations (e.g., line graphs) to handle information in an effective, efficient, and confident manner” (2014, p. 1963). The focus of Boy and colleagues’ study was to describe a method to assess data visualization literacy that concentrates on conventional representations of statistical graphics, such as line graphs, bar charts, and scatterplots. In contrast, the purpose of the current study was to have a way to identify participants’ data visualization literacy levels when conducting online research studies.

Another study by Lee et al. (2017) focused on the development of a visualization literacy assessment test (VLAT) for non-expert users to define data visualization literacy in engineering. The authors indicated that the definition of data visualization literacy is not consistent among the field but went on to define it as “the ability and skill to read and interpret visually represented data in and to extract information from data visualizations” (p. 552). Lee and colleagues’ used other research studies to construct their definition of data visualization literacy. The data visualization tasks used in the VLAT were identified from task taxonomies defined in the field of information visualization. In the field of computer science engineering, Borner et al. (2019) defined a typology of a data visualization literacy framework, but the research article only discussed the definitions within the engineering and computer graphics field.

Data visualizations have been discussed in the field of information science, particularly in the research areas of information literacy and data literacy practices (Philip et al., 2016). In these areas, data visualizations were defined as narratives that are comprised of an interchange between the author’s intention and the user’s comprehension or understanding of the visualization (Hullman & Diakopoulos, 2011; Mackinlay & Kosara, 2013; Philip et al., 2016; Segel & Heer, 2010). From this perspective, the interpretation and use of data visualizations require the user to have information literacy skills. In information science, information literacy involves identifying the need for information, understanding how to access the information, critically evaluating the validity and quality of information, identifying the purpose for the information, and understanding the social, legal, and economic policies and possible consequences of using the information (Philip et al., 2016). Data literacy is a part of information literacy that involves understanding how to use data and the appropriate data representations to support evidence-based thinking that aims to communicate solutions to authentic problems (Calzada Prado & Marzal Miguel, 2013; Vahey et al., 2012). This definition of data literacy with the focus on representations of data to facilitate the communication of ideas about a problem or issue closely resembled definitions of data visualization literacy from library science and engineering.

In science education, Maltese et al. (2015) developed a data visualization literacy assessment to investigate how novices to experts in STEM fields interpret data. The purpose of the study was “understanding the ability of individuals to read and interpret graphs as well as the application of those skills in graph-construction tasks” (Maltese et al., 2015, p. 85). However, the study lacked a formal definition for data visualization literacy and an articulation of the skills that the researchers planned on assessing among their sample of participants.

Finally, in computer science and human-centered computing, research concerning information visualization focuses on defining taxonomies or typologies of visualization tasks that users employ when working with information visualization tools (Amar et al., 2005; Brehmer & Munzner, 2013). In a similar fashion to research in other disciplines, there is a lack of definitions for key constructs and a large focus on providing information about tasks that users of visualization tools engaged in to provide feedback to system designers. Another aspect of information visualization in this field was evaluation scenarios to assist practitioners in identifying goals when assessing visualization tools (Lam et al., 2012).

The definitions offered for data visualization across the studies from different fields all included processes that occur by the creator of the visualization to make an engaging and effective visualization. The articles then used their definition to describe tips and techniques for creating visualizations but did so with minimal to no supporting empirical evidence. The articles mainly focused on personal views regarding how visualizations should be constructed. Thus, there exists a large gap in empirical research about the thinking processes that creators employ and consequently consumers of the visualizations engage in to achieve what is often indicated as the "perfect visualization".

For this study, the definition of data visualization literacy was understood from the reconceptualized four resources model for literacy of visual and multi-modal texts (Serafini, 2012). A visualization that is composed of more than one mode to represent and express ideas in a text, such as images, videos, sounds, music, and graphic designs, is considered a multi-modal text. The reconceptualization of the four resources model focuses on the roles being constructed by the reader in the context of the multi-modal text and not as a predetermined set of cognitive skills for decoding printed texts. The four roles that the reader has when interacting with a multi-modal text are: reader as navigator, reader as interpreter, reader as designer, and reader as interrogator. The four resources model allows data visualization literacy to be understood from the perspective of consuming and producing visualizations. A consumer of a visualization assumes the roles of navigator and interpreter, while the roles of designer and interrogator are taken on by the producer of a visualization.

2. METHODS

2.1. RESEARCH DESIGN

The Delphi method—defined as a group facilitation technique that uses multiple iterations or revisions of a survey administered to a group of experts in the field until consensus is achieved (Hasson et al., 2000)—was used to gather information about skills and strategies regarding data visualization from researchers in the field of data science that have various experiences. The use of the Delphi method allowed for a rich understanding of their experiences and facilitated a consensus of how data scientists interpret and create visualizations. Participants responded in a electronic form to a set of questions about their experiences in the field of data science and their views of data visualization. The responses from the panel were analyzed, and the results were used to frame the next round of questions. This process of analysis and using the results to inform the next round of questions continued for a total of three rounds, which allowed for the opportunity to clarify areas of disagreement and develop a common set of skills and strategies regarding data visualization. The questions used in the Delphi panels were focused on understanding the thinking processes of study participants through the theoretical framework of epistemic thinking. For example, the strategies and skills of interpreting a data visualization was setup by having participants in the first panel describe types of visualizations in their field and then respond with skills or strategies that they think people use when interpreting a data visualization to start collecting a variety of skills and strategies to find common skills among participants in the subsequent Delphi panels. The theoretical framework was used to frame the research design of the study and how the Delphi panels were constructed. The framework of the four resources model was used as a guideline for interpreting the analysis of the qualitative data as the model aligns with epistemic thinking.

In this study, the Delphi panel participants were intentionally grouped by the first author and given the survey in a electronic form for round 1. The author acted as a discussion facilitator for rounds 2 and 3 by presenting a synthesis of results from previous rounds and soliciting further response to those results. This process allowed the first author to be fully embedded in each of the rounds to systematically collect and analyze the responses (Miller & Pasley, 2012). This study was approved by the University of Florida [IRB 201902965].

2.2. SAMPLING/DESCRIPTION OF PARTICIPANTS

The participants were researchers in the field of data science or researchers whose projects involved components of data science. All participants were from the same large research extensive, land-grant university in the southern United States. Selection of participants involved meeting all three inclusion criteria and not meeting any of the exclusion criteria (Table 1). A participant with a title of Lecturer was excluded from this study because the lecturer title at this university indicates a teaching only position without research expectations. These criteria were developed by reviewing the online biographical information for faculty working in data science at a selection of universities.

Table 1. Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Affiliation with a Data Science Institute or Initiative or a department/school of Statistics, Biostatistics, Computer Science, Mathematics, Engineering, Biomedical Informatics, Applied Sciences, Economics, Biology, or Agricultural Sciences	No mention in research biography about modeling, analysis, or processing of data
Research focus involves working with high-dimensional data or big datasets	Job position title is Lecturer or job position is focused only on teaching
Research involves application of quantitative methods to understand data such as advanced statistical techniques, machine learning techniques, or text/language processing	

Purposeful sampling was used to identify and select participants based on their knowledge and experience with data visualization (Patton, 2002). In particular, the strategy of homogeneity was employed because of the focus on a particular subgroup of experts in the field of data science that use visualizations to communicate results of data analysis (Palinkas et al., 2015). In addition to the participants' knowledge and experience, it was also important to consider their willingness to participate and their ability to communicate their experiences in a reflective manner (Bernard, 2002; Spradley, 1979).

A total of 57 potential participants were identified and invited to participate in the study for the first Delphi panel. Previous studies suggested that a consensus panel typically consists of seven to 30 participants from the same discipline (Keeney et al., 2011; Linstone, 1978), where representation is assessed by the quality of experts participating on the panels rather than the size of the panel.

Each potential participant was sent an email invitation through *Qualtrics* (Qualtrics, Provo, UT), a survey software system, stating the purpose of the study, their anticipated involvement, and time commitment. If the potential participants consented to be in the study, the participants were directed in Qualtrics to Delphi Panel 1. Substantial loss of participants between panels can jeopardize the validity of Delphi panels. Therefore, the informed consent document stated that consent to participate would include a total of three Delphi panels over the course of two to three months. Sixteen respondents (28%) out of the total 57 potential participants provided informed consent and became the participants for the study. The demographics of the sample are provided in Table 2.

Table 2. Demographics of the sample

Demographic attribute	Count
Department/School	
Agricultural and Biological Engineering	2
Astronomy	1
Biology	1
Biomedical Engineering	2
Computer and Information Science and Engineering	1
Electrical and Computer Engineering	1
Epidemiology	1
Forest Resources and Conservation	1
Human Development and Organizational Studies in Education	2
Microbiology & Cell Science	1
Special Education, School Psychology, and Early Childhood Studies	1
Statistics	2
Job Title	
Assistant Professor	7
Associate Professor	5
Professor	3
Scientist	1
Would you describe yourself as a data scientist?	
Yes	12
No	4

2.3. DATA COLLECTION

A visualization of the three Delphi panels is provided in Figure 1. The first survey (Delphi Panel 1) was comprised of 11 questions, which was emailed to participants after informed consent was obtained. Table A1 shows the framework, which focused on characterizing the experiences of the data scientists in terms of the data used in their research and their experiences creating and interpreting visualizations. The questions focused on understanding the research experiences of the participants and gathering their comments and thoughts on the skills/strategies they wanted people to use when interpreting and creating a data visualization. A definition of a data visualization from the literature was provided to ensure that all participants had a similar reference point when responding to the questions.

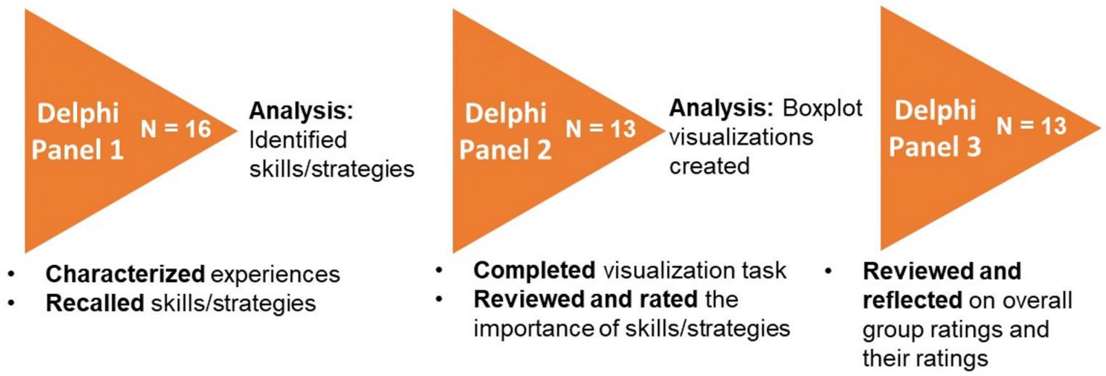


Figure 1. Visualization of Delphi method

In the second survey (Delphi Panel 2), the participants were sent an email that reminded them of their participation in the three-survey study, informed them of the results from the qualitative analysis from Delphi Panel 1, and prepared them for the tasks in Delphi Panel 2. The results of Delphi Panel 1 included a list of emerging skills/strategies for interpreting and creating visualizations that was provided in a reference document for participants to use for completing Delphi Panel 2, which was comprised of two parts: (a) interpreting an interactive visualization, and (b) thinking about creating a visualization. Participants were asked to interpret an interactive visualization about flu trends from 2012 until 2020. The visualization showed the percentage of flu symptoms in the United States as reported by the US Centers for Disease Control and Prevention (CDC) and self-reported flu symptoms from a crowd-sourcing website, Flu Near You (FNY) (https://www.healthmap.org/flutrends/#). The first part of Delphi Panel 2 included open-ended response questions that asked about the participants' overall interpretation of the visualization, what they thought the purpose of the visualization was from the perspective of the creator of the visualization, and what skills/strategies they used when interpreting the visualization. The next two questions were regarding the skills/strategies of interpreting a visualization that were identified during the analysis of Delphi Panel 1. The first question asked participants to select all the skills/strategies from the list from Delphi Panel 1 that they used when interpreting the visualization about flu trends in the previous question. Then, the second question asked the participants to rate the selected skills/strategies according to the level of importance that they attributed to each skill when interpreting the visualization. The Likert-type scale used for rating each skill/strategy was the following: 1 = Not Important, 2 = Somewhat Important, 3 = Moderately Important, 4 = Very Important, and 5 = Extremely Important (Brill et al., 2006; Thangaratinam & Redman, 2005). This scale was used as a proxy for the order of the skills/strategies in terms of importance that the participants believe should be followed when interpreting a data visualization.

The second part of Delphi Panel 2 focused on the skills/strategies for creating a visualization. The first question asked the participants to recall a visualization they had created recently and to describe that visualization in a few sentences. Next, the participants were asked to select all the skills/strategies from the list (same list as provided to them via email) that they used when creating a visualization. Then, the participants were asked to rate the selected skills/strategies according to the level of importance when creating a visualization using the same Likert-type scale as the interpretation question. Table A2 is the framework for Delphi Panel 2.

The purpose of the third and final Delphi panel was to provide participants the ratings for all skills/strategies from the analysis of Delphi Panel 2 and allow them the opportunity to adjust their ratings on any skills/strategies. This third panel was important for the Delphi method because the study goal was to reach a consensus among the diverse panel of experts in the study about their opinions regarding the skills/strategies for interpreting and creating a visualization (Ritchie & Earnest, 1999). In order to reach consensus among the participants on the level of importance for each of the skills/strategies for interpreting and creating a visualization, the overall group ratings for the skills/strategies were presented to the participants in boxplot visualizations.

Delphi Panel 3 followed a similar layout to Delphi Panel 2 with first the skills/strategies for interpreting a visualization and then the skills/strategies for creating a data visualization. For each part, the participants were first asked to review and reflect on the overall ratings for the skills/strategies and review a table of their own individualized ratings from Delphi Panel 2. An individualized survey in Qualtrics was created for each participant because a table specific to each participant was needed. The overall ratings for all study participants were provided in a boxplot visualization with the count of participants that selected each skill/strategy on the y-axis label. The order of the skills/strategies in the boxplot visualization indicated the level of consensus among the participants from Delphi Panel 2.

Next, the participants were asked to answer a question about whether they would like to adjust any of their ratings. If "yes", the participants were directed to the next screen that showed all the skills/strategies that they had selected from Delphi Panel 2. Participants were asked to move the slider for any skill/strategy for which they would like to adjust their rating. If the participants indicated that they did not want to change their ratings, they were directed to the last question in that part of the survey that asked them if they had any additional comments about their individual ratings for each skill/strategy. Table A3 is the framework for Delphi Panel 3.

2.4. DATA ANALYSIS

The open-ended survey response data that were collected from Delphi Panel 1 and 2 were analyzed using the thematic analysis approach by Braun and Clarke (2006) using MAXQDA software (VERBI Software, 2019). This approach included coding for epistemic thinking focused on epistemic metacognitive knowledge as well as the roles participants assumed when engaged with interpreting and creating visualizations from the literacy framework of the four resources model. Thematic analysis is a method of "identifying, analyzing, and reporting patterns (themes) within data" (Braun & Clarke, 2006, p. 79). The six phases of a thematic analysis are (1) familiarizing yourself with your data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report (Braun & Clarke 2006, p. 87). For Delphi Panel 2 and 3, the rating scores for each skill/strategy were analyzed using descriptive statistics to determine whether consensus had been achieved using the interquartile range (middle 50% of the responses).

Delphi Panel 1 analysis. The total number of participants that completed Delphi Panel 1 was 16. After an initial reading of all responses, the first part of the analysis consisted of focusing on questions 5, 6, and 7 from Table A1. The initial coding performed was in reference to the research questions, so all responses were read and coded as either referring to interpreting or creating visualizations. The definitions that were defined about "interpreting" and "creating" based on the framework of data visualization were referenced.

The next part of thematic analysis involved rereading the responses that were coded as interpreting and creating and generating sub-codes based on the four resources model literacy framework of navigator, interpreter, designer and interrogator. Then, a codebook was developed with definitions, and instructions about when to use/not use the codes in order to refine and edit the analysis. Overall, the focus of the thematic analysis (Braun & Clark, 2006) with the four resources model literacy framework (Serafini, 2012) was to ensure that the coding focused on active engagement in the interpretation and creation of visualizations, which aligned with the research questions.

The final part of the thematic analysis for Delphi Panel 1 consisted of integrating all parts of the analysis to identify skills and strategies for interpreting and creating visualizations for Delphi Panel 2. The process consisted of reading all the coded responses about interpreting and identifying skills and strategies. The next step was identifying the skills/strategies that could be combined where appropriate, such as combining *dynamic visualizations* and *interacting with the visualization* into the skill/strategy, *exploring data by interacting with the visualization*. The same process was carried out for the coded responses about creating visualizations.

Delphi Panel 2 analysis. A total of 13 of the 16 participants responded to the second survey. An initial read through of all the data, focusing on the open-ended response questions (question 1, 2, 3, and 8 from Table A2) was conducted first. Then, descriptive statistics (total count, mean, standard deviation, minimum, maximum, 1st quartile, 3rd quartile, median, and count missing) for the ratings of the 15 skills/strategies for interpreting and 21 skills/strategies for creating were calculated. Interquartile range (*IQR*) was used to assess consensus in the responses because the rating scale used was an ordinal measurement (Hasson et al., 2000). The *IQR* is the absolute difference between the 3rd quartile and the 1st quartile, which measures the middle 50% of the distribution of ratings. A smaller *IQR* indicates a higher degree of consensus (Persai et al., 2016). The cut-off points for the *IQR* to determine consensus among participants in this study for a skill/strategy required (a) a minimum sample size of 5 given the small sample for the panel (n = 13), and that all participants did not rate all skills/strategies, and (b) an $IQR \leq 1.2$ (Alexander, 2008; Baker, 2005; Hussein, 2010). The use of the *IQR* as an objective measure of the stability of participants' responses rather than the use of percentage measures is important for consensus building (Hsu & Sandford, 2007). Two boxplot visualizations were created to show the distribution of ratings for each skill/strategy for interpreting and creating a data visualization, and provided to the participants in Delphi Panel 2.

Delphi Panel 3 analysis. All 13 participants responded to the third and final survey. The descriptive statistics were calculated for Delphi Panel 3 for any ratings of skills/strategies for interpreting and creating a visualization that were changed by participants. The same criteria cut-off points for the *IQR* from the analysis of Delphi Panel 2 was used to determine fnal consensus among participants. Boxplot visualizations were created for the final ratings of skills/strategies for interpreting and creating a visualization with skills that achieved consensus based on the criteria colored as green. Mean values were indicated on the boxplots for each skill by black diamonds (see Figures 2 and 3).

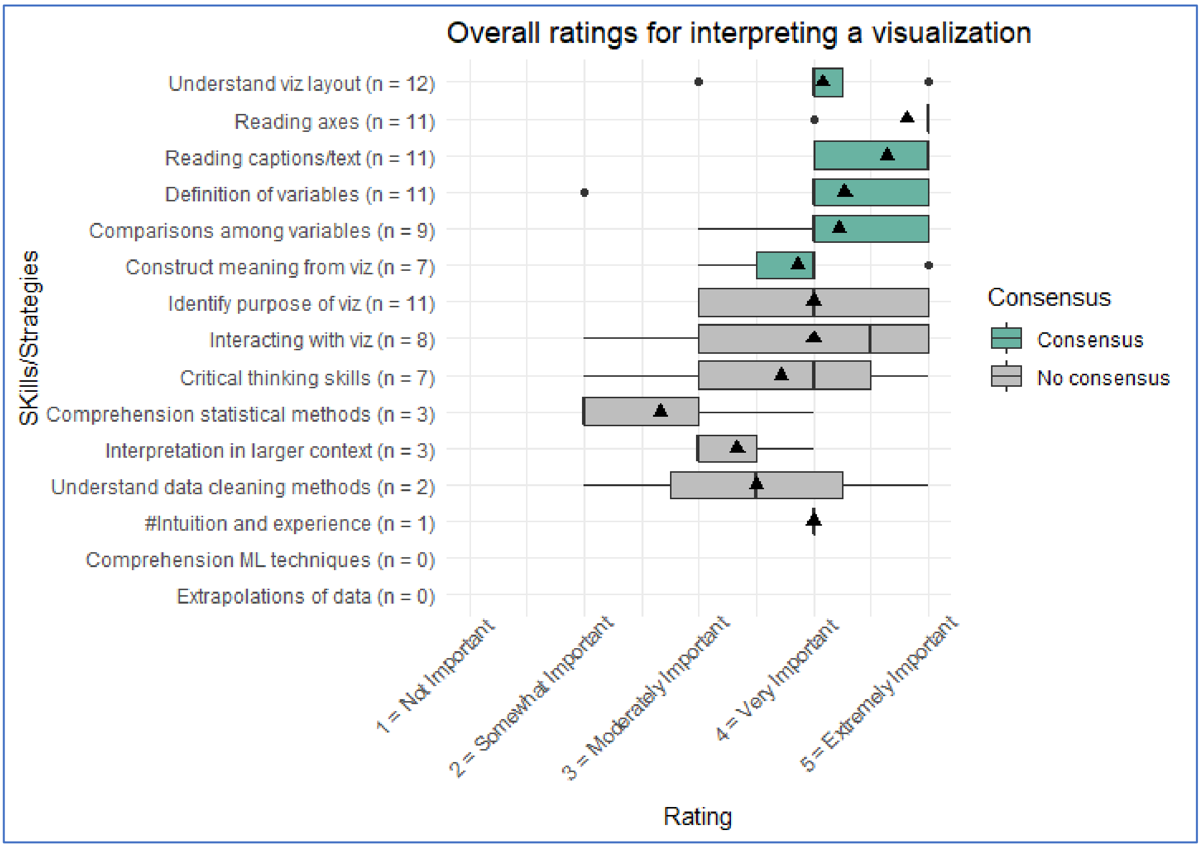
3. RESULTS

3.1. RESEARCH QUESTION 1: INTERPRETING A VISUALIZATION

A visualization showing the adjusted overall ratings for the skills/strategies for interpreting a visualization from Delphi Panel 3 is provided in Figure 2. The table of descriptive statistics are provided in the Appendix Table A4. The two criteria for whether consensus was achieved was (1) a minimum sample size (N) of 5 and (2) an $IQR \leq 1.2$. The results from Delphi Panel 3 indicated that the final consensus was 6 skills/strategies for interpreting a visualization. The skills/strategies that achieved consensus were *understanding the layout of the visualization* (n = 12, median = 4.00, IQR = 0.25) *reading axes* (n = 11, median = 5.00, IQR = 0.00*), constructing meaning from the visualization/gaining insight* (n = 7, median = 4.00, IQR = 0.50), *drawing comparisons among variables* (n = 9, median = 4.00, IQR = 1.00), *reading captions*/text (n = 11, median = 5.00, IQR = 1.00), *understanding the definition/meaning of variables displayed* (n = 11, median = 4.00, IQR = 1.00). All of these skills/strategies maintained a median rating of 4 to 5 from Delphi Panel 2 to 3, which meant that 50% of study participants rated these skills/strategies as "very important" to "extremely important".

Overall, the 6 skills/strategies for interpreting a visualization that achieved consensus did not change between Delphi Panel 2 and 3. There were only two participants that adjusted their ratings: one participant changed their rating of *reading captions/text* from 3 to 4 and the other participant changed their rating of *reading axes* from 4 to 5. These changes in ratings resulted in no change in the IQR for the skill/strategy of *reading captions/text* but the IQR for *reading axes* decreased from 0.50 to 0.00. It is notable to mention that the minimum value for *reading captions/text* increased from 3.00 to 4.00 with the participant's change in their rating of that skill/strategy.

The three interpreting skills/strategies that had not achieved consensus from Delphi Panel 2 remained consistent after Delphi Panel 3 because no participants adjusted their ratings for *critical thinking skills*, *exploring data by interacting with the visualization*, and *identifying a purpose of the visualization*. The skills of *exploring data by interacting with the visualization* and *identifying a purpose of the visualization* can be considered has having high importance but no consensus with each skill having a mean score above 4 ("very important"). The skills/strategies that had not reached consensus after Delphi Panel 3 because only a few participants selected them (N < 5) were *understanding the methods of data cleaning and data staging*, *comprehension of statistical methods*, *interpreting the visualization within a larger context within the field*, *intuition and experience*, *comprehension of machine learning techniques*, and *making extrapolations of the data*.

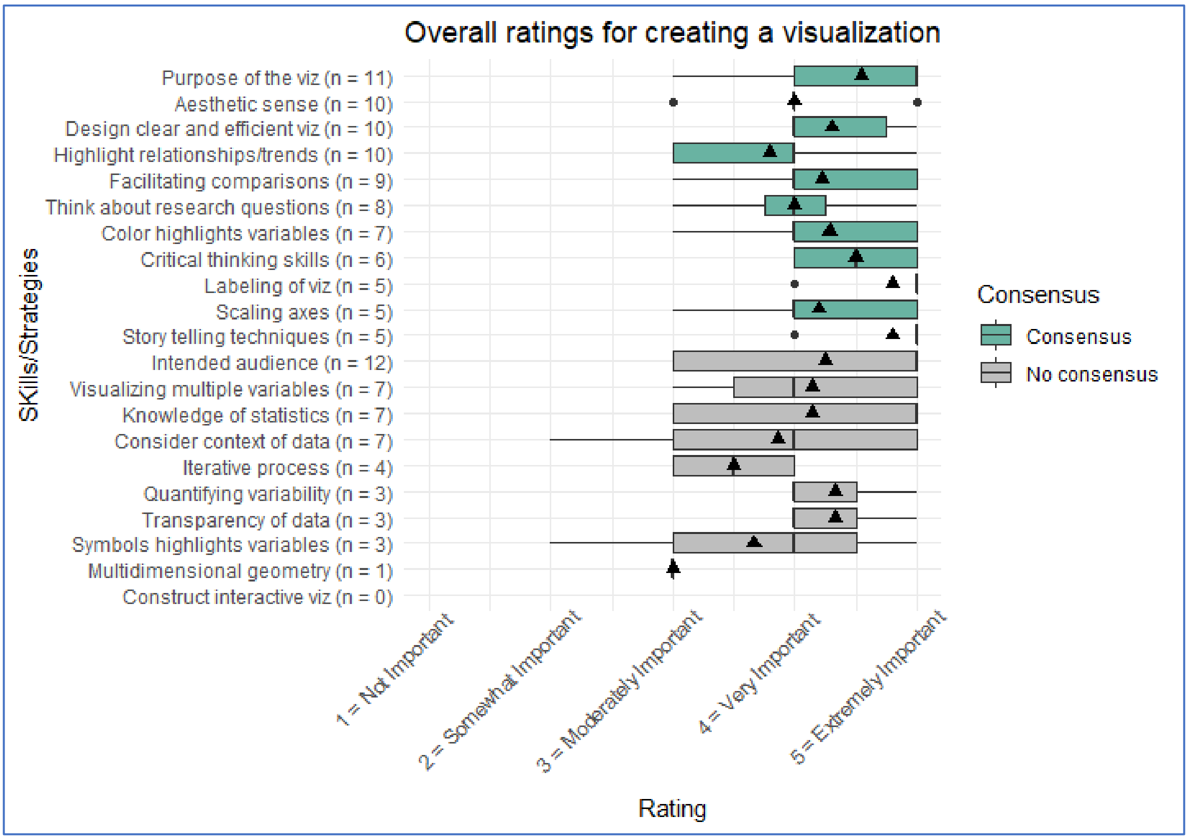


Note: ▲ indicates mean value for each skill/strategy; # indicates an additional skill/strategy that was added during Delphi Panel 2; "viz" is visualization; "ML" is machine learning.

Figure 2. Visualization of Delphi Panel 3 results for interpreting skills/strategies

3.2. RESEARCH QUESTION 2: CREATING A VISUALIZATION

A visualization showing the adjusted overall ratings for the skills/strategies for creating a visualization from Delphi Panel 3 is provided in Figure 3. The table of descriptive statistics are provided in the Appendix Table A5. The results from Delphi Panel 3 demonstrated that the final consensus among participants was that 11 skills/strategies were important for creating a visualization. The skills/strategies that achieved consensus were *aesthetic sense* (*n* = 10, median = 4.00, *IQR* = 0.00), *thinking about the research questions from the study/experiment* (*n* = 8, median = 4.00, *IQR* = 0.50), *designing visualizations with clear and efficient meaning* (*n* = 10, median = 4.00, *IQR* = 0.75), *labeling all aspects of the visualization (axes, legends, etc.)* (*n* = 5, median = 5.00, *IQR* = 0.00), *scaling axes appropriately* (*n* = 5, median = 4.00, *IQR* = 1.00), *facilitating comparisons among graphs in a visualization* (*n* = 9, median = 4.00, *IQR* = 1.00), *critical thinking skills* (*n* = 6, median = 4.50, *IQR* = 1.00), *defining the purpose of the visualization* (*n* = 11, median = 5.00, *IQR* = 1.00), *highlighting main points/patterns* (relationships/trends) (*n* = 10, median = 4.00, *IQR* = 1.00), *using color to highlight multiple variables* (*n* = 7, median = 4.00, *IQR* = 1.00), and *using story telling techniques* (*n* = 5, median = 5.00, *IQR* = 0.00). All these skills/strategies had a median rating of 4 to 5, which remained constant from Delphi Panel 2. Therefore, 50% of study participants rated these creating skills/strategies as "very important" to "extremely important".



Note: ▲ indicate mean values for each skill/strategy; "viz" is visualization.

Figure 3. Visualization of Delphi Panel 3 results for creating skills/strategies

Overall, the consensus of 11 skills/strategies for creating a visualization did not change between Delphi Panel 2 and 3. There were a total of three participants who decided to adjust their ratings of the creating skills/strategies. One participant changed their rating for *aesthetic sense* from 3 to 4. This change resulted in the 1st quartile increasing so the *IQR* of *aesthetic sense* decreased from 0.75 to 0.00. This same participant also decided to change their rating of the skill *transparency of information about the data* from a rating of 3 to 4. Another participant changed their rating from 4 to 3 for *engaging in an iterative process when creating the visualization*. This skill/strategy did not meet the sample size criteria with only four participants that selected it and the change in rating resulted in the *IQR* increasing from 0.25 to 1.00. The third participant adjusted their rating of *using story telling techniques* from 4 to 5, which resulted in the *IQR* decreasing from 1.00 to 0.00.

Four creating skills/strategies had not achieved consensus with *IQR*s ranging from 1.50 to 2.00 from Delphi Panel 2. These four skills/strategies remained outside of the *IQR* cut-off after Delphi Panel 3 as well because no ratings for those skills/strategies were adjusted by participants (*visualizing multiple variables at once*, *thinking about the intended audience of the visualization*, *knowledge of foundational statistics concepts*, and *considering the context of the data*). The skills of *visualizing multiple variables at once*, *thinking about the intended audience of the visualization*, *knowledge of foundational statistics concepts* can be considered has having high importance but no consensus with each skill having a mean score above 4 ("very important"). After Delphi Panel 3, the skills/strategies that had not reached consensus because only a few participants selected them ($n < 5$) were *engaging in an iterative process when creating the visualization*, *quantifying variability*, *transparency of information about the data*, *using symbols to highlight multiple variables*, *multidimensional geometry*, and *constructing dynamic/interactive visualizations*.

4. DISCUSSION

Data visualization literacy has been interpreted in various ways throughout the library science, STEM education, information science, and computer science/human-centered computing research fields. The definition of data visualization literacy and the strategies that are involved when creating and interpreting a visualization have not been well-defined. However, one commonality among these varied definitions is that the creator of the visualization is involved in a strategic process to create an engaging and effective visualization. There is a lack of empirical research that examines and defines the strategic thinking processes that consumers and producers engage in when interpreting and creating visualizations. There is a need to better understand and clearly define data visualization literacy.

The use of epistemic thinking as the theoretical framework for this study allowed the thinking processes that data scientists utilize when creating and interpreting visualizations to be explored and for the strategies to be identified. The comprehension or interpretation of visualizations requires complex thinking in terms of understanding the data displayed, making meaning from those data values, reflecting on the purpose of the visualization, and how the information that is interpreted can be extrapolated or applied to the field of research (Kirk, 2016). The process of creating visualizations requires data scientists to make decisions about how the visualization will be interpreted in terms of the audience and the intended purpose of the visualization.

4.1. INTERPRETING STRATEGIES

The results indicated a consensus for six strategies as important during the process of interpreting a visualization, which were: *understanding the layout of the visualization*, *reading axes*, *constructing meaning from the visualization/gaining insight*, *drawing comparisons among variables*, *reading captions/text*, and *understanding the definition/meaning of variables displayed*. These strategies were not passive actions and consisted of thinking processes that are more than just "reading" and basic graph interpretation.

In a previous study about students' epistemic thinking that examined their epistemic metacognitive knowledge about both persons and strategies and tasks regarding evaluating and integrating online sources, the researchers stated it was difficult for students to identify epistemic metacognitive knowledge about strategies and tasks when they were not actively engaged in those strategies (Barzilai & Zohar, 2012). The aspect of having the participants engaged in interpreting a visualization was essential in this study because the participants were able to identify more epistemic metacognitive knowledge about strategies in Delphi Panel 2 in terms of strategies selected compared to their open-ended responses in Delphi Panel 1 when they were not engaged in a cognitive task. Epistemic metacognitive knowledge about persons was less emphasized in the strategies about interpreting a visualization that reached consensus among the participants. In the study by Barzilai and Zohar (2012), their results suggested that epistemic metacognitive knowledge about strategies "is an important aspect of epistemic metacognition and a crucial link between students' epistemic metacognitive knowledge about persons (EMKP) about the nature of knowledge and knowing and how they go about creating and justifying knowledge" (p. 73).

In a study by Barzilai and Eilam (2018), seventh and eighth grade students' epistemic metacognitive knowledge about epistemic criteria for evaluating scientific visual representations (VR) resulted in three types of epistemic criteria which were communicative, representational, and epistemic aim affordance. The communicative criterion of organization, which was described as "a good VR is well-organized or well-arranged" (Barzilai & Chinn, 2018, p. 142), aligns with the strategy of *understanding the layout of the visualization*. The organization or layout of a visualization needed to be clear and well-defined for the interpreter of the visualization. The epistemic aim affordance criteria of understanding defined as "a good VR can be well understood by the viewer", inquiry described as "a good VR enables inquiry or research", and learning which was defined as "a good VR enables learning about the referent" (Barzilai & Chinn, 2018, p. 142) line up with the strategies of *constructing meaning from the visualization/gaining insight*, *drawing comparisons among variables*, and *understanding the definition/meaning of variables displayed*. The ability of the interpreter of a visualization or VR to takeaway information from a visualization that enables them to better understand the data or further their research is necessary and critical in the field of data science.

Middle school students' epistemic criteria for good models were investigated to understand the extent to which student-generated epistemic criteria align with the criteria used by practicing scientists (Pluta et al., 2011). Examples of criterion elements identified by the students were diagrams, labels, arrows, visuals, and titles. The "responses in the model constituents category specify features of models that are needed to help communicate what the models represent" (Pluta et al., 2011, p. 496). Skills and strategies that achieved consensus in the present study were *reading captions/text* and *reading axes.* These similar skills focus on helping the consumer orient themselves to the important features of the visualization.

An aspect of graphical comprehension is the ability to enhance students' understanding of statistical concepts such as variability or inference. Students' ability to understand and identify the variability in data can be improved through seeing and interacting with different graphical displays and seeing visual comparisons (Cooper & Shore, 2010). In addition, use of pictorial visualizations helps students grasp concepts like statistical inference (Forbes et al., 2014). The understanding of variability is a part of statistical thinking. However, students are not necessarily taught about graphical displays with the connection to statistical concepts, such as center and spread or meaning of axes when learning about different graph types (Cooper & Shore, 2010). The strategies of *constructing meaning from the visualization/gaining insight* and *drawing comparisons among variables* align with students' interpretation of variability as represented in graphical displays and data visualizations.

4.2. CREATING STRATEGIES

The results indicated a consensus for 11 strategies for creating visualizations, which included: *aesthetic sense*, *thinking about the research questions from the study/experiment*, *designing visualizations with clear and efficient meaning*, *labeling all aspects of the visualization (axes, legend, etc.)*, *scaling axes appropriately*, *facilitating comparisons among graphs in a visualization*, *critical thinking skills*, *defining the purpose of the visualization*, *highlighting main points/patterns (relationships/trends)*, *using color to highlight multiple variables*, and *using story telling techniques*. These strategies were all active actions that the particpants engaged in when creating a visualization and were more in number compared to those indicated for interpreting a visualization. Barzilai and Zohar (2012) found that students engaged in higher levels of evaluating sources and integration of multiple sources were associated with increased epistemic metacognitive knowledge about evaluation and integration. In a similar way, strategies that participants engaged in when thinking back on the process of creating a visualization may have resulted in increased epistemic metacognitive knowledge, hence, the increased number of strategies that reached consensus compared to the strategies for interpreting a visualization. Additional research directly comparing strategies for interpreting and creating visualizations is needed to better understand if creating visualizations results in higher epistemic metacognitive knowledge.

In the study by Barzilai and Zohar (2012), one of the important educational implications was the development of complex epistemic thinking to improve students' abilities to construct knowledge. The researchers identified that it was valuable to have students recognize explicitly the connection between their comprehension of the nature of knowledge and knowledge construction strategies. In a parallel way, it is likley beneficial for the general public to recognize the connection in data visualization literacy between the knowledge strategies for interpreting and for constructing a visualization. The ability to understand the strategies of thinking processes that experts use may assist learners at all levels on the data visualization literacy learning progression.

Barzilai and Eliam (2018) identified communicative criteria about students' epistemic metacognitive knowledge about epistemic criteria for evaluating scientific VRs such as clarity, simplicity, aesthetic appeal, and color. The criterion of clarity, which was defined as "a good VR is clear or presents information in a clear way" (Barzilai & Chinn, 2018, p. 142) related to the skill/strategy of *designing visualizations with clear and efficient meaning*. A visualization that can be distinctly created to communicate the purpose of showcasing the data is an important criterion for a good VR and an aspect of the process of creating a visualization that needs to be kept at the forefront. The criterion also supported the skill/strategy of *labeling all aspects of the visualization (axes, legend, etc.)* because a clear labelling of a visualization facilitates the information being understood correctly and without any confusion. For example, labels to identify what the y-axis or legend represents from the data. The criterion of simplicity which was described as “a good VR is simple or not too overloaded” (Barzilai & Chinn, 2018, p. 142) aligned to the skill/strategy of *facilitating comparisons among graphs in a visualization*. When creating a visualization, it is critical to balance presenting detailed information to the viewer, but also to make sure that the viewer is not cognitively overloaded with information. The skill/strategy of *facilitating comparisons* when creating a visualization allows the viewer to be assisted in what their eyes should focus on and helps direct their attention.

The criterion of aesthetic appeal was defined as “a good VR is aesthetically pleasing or attractive” (Barzilai & Chinn, 2018, p. 142) which corresponded to the skill/strategy of *aesthetic sense*. The value of having a visualization that is appealing to look at is a significant factor that needs to be considered during the process of creating a visualization. The use of a color palate that is unappealing or font choice that is hard to read can impact how a viewer will interpret and evaluate the visualization. The criterion of color was described as “a good VR is colorful” (Barzilai & Chinn, 2018, p. 142) paralleled the skill/strategy of *using color to highlight multiple variables*. This skill/strategy when creating a visualization has the purpose of using color to bring out certain variables in terms of categories or levels, so the intention is to not make a pretty looking graphical display. Instead, the color used in a visualization has a purpose and intention.

The representational criteria that were identified in the Barzilai and Eilam (2018) study that corresponded to the strategies identified for creating a visualization were accuracy/precision and importance of information. The criterion of accuracy/precision was defined as “a good VR is accurate/precise or presents information accurately/precisely” (Barzilai & Chinn, 2018, p. 142) corresponded to the strategy of *scaling axes appropriately*. The scaling of axes correctly when creating a visualization ensures that the data represented is accurate and not skewed incorrectly. The criterion of importance of information was described as “a good VR includes important information about the referent” (Barzilai & Chinn, 2018, p. 142) which aligned to the strategy of *highlight main points/patterns (relationships/trends)*. When creating a visualization, making sure that information important to the visualization is emphasized is vital to having the intended purpose be understood by the viewer. The strategies for creating a visualization that achieved consensus could have been positioned within the first dimension of the investigative cycle of the statistical thinking model (Wild & Pfannkuch, 1999). The investigative cycle “concerns the way one acts and what one thinks about during the course of a statistical investigation” (Wild & Pfannkuch, 1999, p. 225). The investigative cycle includes the components of problem, plan, data, analysis, and conclusions. The strategies of *thinking about the research questions from the study/experiment* and *defining the purpose of the visualization* can be labeled under the problem component in the cycle because both skills/strategies focus on defining the problem. The strategies of *facilitating comparisons among graphs in a visualization*, *highlighting main points/patterns (relationships/trends)*, *using color to highlight multiple variables*, *scaling axes appropriately*, *critical thinking skills*, and *labeling all aspects of the visualization (axes, legends, etc.)* all fall under the component of analysis because the strategies are focused on exploring the data and understanding how to display the data in an effective way. The final strategies of *designing visualizations with clear and efficient meaning*, *using story telling techniques*, and *aesthetic sense* can be understood through the conclusions component of the cycle because the strategies focus on forming inferences and communicating information to an audience.

4.3. LIMITATIONS

One limitation of the study was that all the experts that were selected to participate were from the same large research university. The reason for this decision was that the researchers wanted to ensure that consistent contact could be maintained between the researchers and the participants to ensure that they were engaged in all three rounds of the Delphi panels. The survey method was a limitation in that each participant’s epistemic cognition could not be more deeply investigated with probing questions as the participant was engaged in an activity. However, the Delphi method using electronic surveys was the best method to efficiently collect data from participants to answer the research questions of this study. Another limitation was the use of the *IQR* criteria as a way to determine consensus. This criterion did not allow for consensus to be established for some strategies that had a spread of responses from 3 to 5 on the Likert-type scale but a mean around 4. Other criteria could be used to determine consensus, which are subject to interpretation and varies between Delphi studies (Hsu & Sandford, 2007; Giannarou & Zervas, 2014).

The definition of what an expert is could also have been a potential limitation of this Delphi study since the identification of experts has been a major issue of deliberation and some researchers have criticized the use of the title of expert (Keeney et al., 2001; Linstone & Turoff, 1975). However, Delphi participants should meet the following four criteria to meet the expertise requirements based on Adler and Ziglio (1996): (1) knowledge and experience with the phenomena being studied, (2) willingness and ability to participate, (3) adequate time to participate in all rounds of the Delphi panels, and (4) effective communication skills. Therefore, the participants in this study met all four expertise requirements so for the purpose of the study they were considered experts in data visualization.

4.4. IMPLICATIONS FOR RESEARCH

Overall, the results reported here have implications for future research involving data visualization literacy. Think aloud interviews with data scientists as they interpret and create visualizations may provide additional detail about these processes. The participants in this study voiced some difficulty in having to think about their personal knowledge and processes of knowing when answering the open-ended response questions on the Delphi panels. Think aloud interviews would allow for the researchers to have participants engage in the task of interpreting and creating a visualization and have the participants verbalize their thoughts, insights, wonderings, and processes for completing the tasks. This would afford the researchers the opportunity to probe participants more deeply about strategies they use when creating and interpreting visualizations. This methodological approach has also been promoted as having high potential for exploring epistemic metacognitive knowledge (Hofer, 2004). Some studies have been conducted that looked at data analysis processes and data driven decision making done by data scientists using semi-structured interviews, but the focus has not been on cognitive skills when creating and interpreting visualizations (Kandel et al., 2012; Mosca et al., 2019).

The development of new curriculum materials for data visualization literacy is another area for future research. These findings provide a basis for learning modules as well as a starting point for an assessment rubric. Finally, think aloud interviews with college students about their epistemic cognition when interpreting and creating a visualization could assist in defining the variability in data visualization literacy based upon background and level of experience. These results could then be used in defining a learning progression of college-level students' data visualization literacy.

5. CONCLUSION

The purpose of this study was to understand the experiences of data scientists regarding common strategies of data visualization to define the expert skills and strategies of a learning progression of data visualization literacy. The findings include six strategies for interpretation and eleven strategies for creation of a data visualization that reached consensus and were complimentary. These evidence-based strategies indicate what data scientists consider important when interpreting and creating a visualization. The findings specifically can be used to support and define a learning progression of data visualization literacy. The future implications of this research are additional think aloud studies with data scientists to further explore the strategies that achieved consensus along with investigating the variability of data visualization literacy of college-level students given that a high-end anchor has now been defined.

REFERENCES

Adler, M., & Ziglio, E. (1996). *Gazing into the oracle: The Delphi method and its application to social policy and public health*. Kingsley.

Alexander, D. R. (2008). *A modified delphi study of future crises and required leader competencies.* [Doctoral dissertation, University of Phoenix].

Amar, R., Eagan, J., & Stasko, J. (2005). Low-level components of analytic activity in information visualization. In J. Stasko & M. Ward (Eds.), *Proceedings of the IEEE Symposium on Information*

Visualization (InfoVis05), Minneapolis, MN, October 23–25, (pp. 111–117). https://doi.ieeecomputersociety.org/10.1109/INFVIS.2005.1532136

Azzam, T., & Evergreen, S. (2013). *Data visualization, Part 1: New directions for evaluation*. Jossey-Bass.

Baker, K. (2005). *A model for leading online K–12 learning environments* [Doctoral dissertation, University of Phoenix].

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences, 27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Eilam, B. (2018). Learners' epistemic criteria and strategies for evaluating scientific visual representations. *Learning and Instruction, 58*, 137–147.

Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources *Cognition and Instruction, 30*(1), 39-85. https://doi.org/10.1080/07370008.2011.636495

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist, 49*(1), 13–35. https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., & Zohar, A. (2016). Epistemic (meta)cognition: Ways of thinking about knowledge and knowing. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition*. Routledge.

Bernard, H. R. (2002). *Research methods in anthropology: Qualitative and quantitative approaches* (3rd ed.). Alta Mira Press.

Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences, 116*(6), 1857–1864. https://doi.org/10.1073/pnas.1807180116

Börner, K., Maltese, A., Balliet, R. N., & Heimlich, J. (2016). Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization, 15*(3), 198–213.

Bouquin, D., & Epstein, H.-A. B. (2015). Teaching data visualization basics to market the value of a hospital library: An infographic as one example. *Journal of Hospital Librarianship, 15*(4), 349–364. https://doi.org/10.1080/15323269.2015.1079686

Boy, J., Rensink, R. A., Bertini, E., & Fekete, J.-D. (2014). A principled way of assessing visualization literacy. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 1963–1972. https://doi.org/10.1109/TVCG.2014.2346984

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brehmer, M., & Munzner, T. (2013). A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics, 19*, 2376–2385.

Brill, J. M., Bishop, M. J., & Walker, A. E. (2006). The competencies and characteristics required of an effective project manager: A web-based Delphi study. *Educational Technology Research and Development, 54*(2), 115–140.

Calzada Prado, J., & Marzal Miguel, Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123–134. https://doi.org/10.1515/libri-2013-0010

Cleveland, W. S. (2001). Data science: An action plan for expanding the technical areas of the field of statistics. *International Statistical Review/Revue Internationale de Statistique, 69*(1), 21–26. https://doi.org/10.2307/1403527

Cambridge Cognition. (2015). What is cognition? https://cambridgecognition.com/blog/entry/what-is-cognition

Carrión Pérez, J. C., & Espinel Febles, M. C. (2006). *An investigation about translation and interpretation of statistical graphs and tables by students of primary education.* In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education. Proceedings of the 7th International Conference on Teaching Statistics* (ICOTS7), Salvador, Bahia, Brazil, July. 2–7. ISI. https://iase-web.org/documents/papers/icots7/C332.pdf?1402524967

Cooper, L. L., & Shore, F. S. (2010). The effects of data and graph type on concepts and visualizations of variability. *Journal of Statistics Education*, *18*(2), 1–16. https://doi.org/10.1080/10691898.2010.11889487

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, *18*(5), 382–393.

De Veaux, R. D., Agarwal, M., Averett, M., Baumer, B. S., Bray, A., Bressoud, T. C., . . . Ye, P. (2017). Curriculum guidelines for undergraduate programs in data science. *Annual Review of Statistics and Its Application*, *4*(1), 15–30. https://doi.org/10.1146/annurev-statistics-060116-053930

Eilam, B. (2015). *Teaching, learning, and visual literacy: The dual role of visual representation*. Cambridge University Press.

Figueiras, A. (2013). A typology for data visualization on the web. In the *Proceedings of the 17th International Conference on Information Visualisation*, London, England, July 16–18 (pp. 351–358). https://doi.ieeecomputersociety.org/10.1109/IV.2013.45

Forbes, S., Chapman, J., Harraway, J., Stirling, D., & Wild, C. (2014). Use of data visualizations in the teaching of statistics: A New Zealand perspective. *Statistics Education Research Journal*, *13*(2), 187–201. https://doi.org/10.52041/serj.v13i2.290

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*(2), 124–158.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*, 1008–1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist*, *39*(1), 43–55. https://doi.org/10.1207/s15326985ep3901_5

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 19–38). Routledge.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88–140. https://doi.org/10.3102/00346543067001088

Hsu, C. C., & Sandford, B. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, *12*(10), Article 10. https://doi.org/10.7275/pdz9-th90

Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2231–2240. https://doi.org/10.1109/TVCG.2011.255

Hussein, M. M. (2010). Corporate social responsibility: Finding the middle ground. *Social Responsibility Journal*, *6*(3), 420–432.

Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2012). Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2917–2926.

Kapler, T., & Wright, W. (2004). *GeoTime information visualization.* In M. Ward & T. Munzer (Eds.), *Proceedings of the IEEE Symposium on Information Visualization (InfoVis04),* Washington, DC, October 10–12. https://doi.org/10.1109/INFVIS.2004.27

Keeney, S., Hasson, F., & McKenna, H. P. (2001). A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies*, *38*(2), 195–200. https://doi.org/10.1016/S0020-7489(00)00044-4

Keeney, S., Hasson, F., & McKenna, H. P. (2011). *The Delphi technique in nursing and health research*. Wiley-Blackwell.

Keim, D. A., Kohlammer, J., Mansmann, F., May, T., & Wanner, F. (2010). Visual analytics. In D. A. Keim, J. Kohlhammer, & G. Ellis (Eds.), *Mastering the information age: Solving problems with visual analytics.* http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.363.2661

Kirk, A. (2016). *Data visualisation: A handbook for data driven design*. SAGE Publications.

Kosslyn, S. M. (1985). Graphics and human information processing: A review of five books. *Journal of the American Statistical Association*, *80*(391), 499–512.

Laina, V., & Wilkerson, M. (2016). Distributions, trends, and contradictions: A case study in sensemaking with interactive data visualizations. In C. K. Looi, J. L., Polman, U. Cress, & P. Reimann (Eds.). *Proceedings of the International Society of the Learning Sciences,* Singapore, June 20-24). https://repository.isls.org//handle/1/347

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., & Carpendale, S. (2012). Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, *18*(9), 1520–1536.

Lee, S., Kim, S.-H., & Kwon, B. C. (2017). VLAT: Development of a visualization literacy assessment test. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 551–560. https://doi.org/10.1109/TVCG.2016.2598920

Lem, S., Kempen, G., Ceulemans, E., Onghena, P., Verschaffel, L., & Van Dooren, W. (2015). Combining multiple external representations and refutational text: An intervention on learning to interpret box plots. *International Journal of Science and Mathematics Education, 13*(4), 909–926. https://doi.org/10.1007/s10763-014-9604-3

Lima, I., & Selva, A. (2010). Youth and adults students interpreting bar and line graphs. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the 8th International Conference on Teaching Statistics* (ICOTS8), Ljubljana, Slovenia, July 11–16. https://iase-web.org/documents/papers/icots8/ICOTS8_C134_LIMA.pdf?1402524973

Linstone, H. A. (1978). *The Delphi technique. Handbook of futures research*. Greenwood.

Linstone, H. A., & Turoff, M. (1975). *The Delphi method: Techniques and applications*. Addison-Wesley.

Lowrie, T., Diezmann, C. M., & Logan, T. (2012). A framework for mathematics graphical tasks: The influence of the graphic element on student sense making. *Mathematics Education Research Journal*, *24*(2), 169–187. https://doi.org/10.1007/s13394-012-0036-5

Mackinlay, J., & Kosara, R. (2013). Storytelling: The next step for visualization. *Computer, 46,* 44–50.

Maltese, A. V., Svetina, D., & Harsh, J. A. (2015). Data visualization literacy: Investigating data interpretation along the novice-expert continuum. *Journal of College Science Teaching, 45*(1), 84–90.

Mason, L. (2016). Psychological perspectives on measuring epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 375–392). Routledge.

Miller, B., & Pasley, J. (2012). What do we know and how well do we know it? Identifying practice-based insights in education. *Evidence and Policy, 8*(2), 193–212.

Mirel, B., Kumar, A., Nong, P., Su, G., & Meng, F. (2016). Using interactive data visualizations for exploratory analysis in undergraduate genomics coursework: Field study findings and guidelines. *Journal of Science Education and Technology*, *25*, 91–110.

Mosca, A., Robinson, S., Clarke, M., Redelmeier, R., Coates, S., Cashman, D., & Chang, R. (2019). Defining an analysis: A study of client-facing data scientists. In J. Johansson, F. Sadlo, & G. E. Marai (Eds.), *Proceedings of the EuroVis 2019: 21th EG/VGTC Conference on Visualization, Short Papers*, Porto, Portugal, June 3–7 (pp. 73–77). https://doi.org/10.2312/evs.20191173

Nicolaou, C. T., Nicolaidou, I. A., Zacharia, Z. C., & Constantinou, C. P. (2007). Enhancing fourth graders' ability to interpret graphical representations through the use of microcomputer-based labs implemented within an inquiry-based activity sequence. *The Journal of Computers in Mathematics and Science Teaching, 26*(1), 75–99.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*, 533–544.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). SAGE Publications.

Persai, D., Panda, R., Kumar, R., & Mc Ewen, A. (2016). A Delphi study for setting up tobacco research and practice network in India. *Tobacco Induced Diseases*, *14*, 4–4. https://doi.org/10.1186/s12971-016-0067-x

Pfannkuch, M. (2006). Comparing box plot distributions: A teacher's reasoning. *Statistics Education Research Journal*, *5*(2), 27–45.

Pfannkuch, M., Regan, M., Wild, C., & Horton, N. J. (2010). Telling data stories: Essential dialogues for comparative reasoning. *Journal of Statistics Education*, *18*(1), Article 10. https://doi.org/10.1080/10691898.2010.11889479

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition & Instruction*, *34*(4), 361–388. https://doi.org/10.1080/07370008.2016.1210418

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, *48*(5), 486–511. https://doi.org/10.1002/tea.20415

Ritchie, D., & Earnest, J. (1999). The Future of Instructional Design: Results of a Delphi Study. *Educational Technology, 39*(1), 35–42.

Roth, W. M., & Bowen, G. M. (2001). Professionals read graphs: A semiotic analysis. *Journal for Research in Mathematics Education*, *32*(2), 159–194.

Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge. *Review of Research in Education, 40*(1), 457–496.

Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 1139–1148. https://doi.org/10.1109/TVCG.2010.179

Serafini, F. (2012). Expanding the four resources model: Reading visual and multi-modal texts. *Pedagogies: An International Journal*, *7*(2), 150–164.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, *14*(1), 47–69. http://dx.doi.org/10.1023/A:1013180410169

Spradley, J. P. (1979). *The ethnographic interview*. Holt, Rinehart & Winston.

Takemura, A. (2018). *New undergraduate departments and programs of data science in Japan.* In M. A. Sorto & E. Paparistodemou (Eds.), *Looking back, looking forward. Proceedings of the 10th International Conference on the Teaching of Statistics* (ICOTS10), Kyoto, Japan, July 8–14. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_1I2.pdf?1531364187

Thangaratinam, S., & Redman, C. W. (2005). The Delphi technique. *The Obstetrician & Gynaecologist*, *7*(2), 120–125. https://doi.org/10.1576/toag.7.2.120.27071

Vahey, P., Rafanan, K., Patton, C., Swan, K., van ’t Hooft, M., Kratcoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics, 81*(2), 179–205. https://doi.org/10.1007/s10649-012-9392-z

VERBI Software. (2019). MAXQDA 2020 [computer software]. VERBI Software. www.maxqda.com

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–248. https://doi.org/10.2307/1403699

Charlotte A. Bolch
Midwestern University
19555 N 59th Ave.
Glendale, AZ 85308

Table A1. Framework for Delphi Panel 1

Question	Rationale
1. What is your job title at the university?	Focus on tenure track professors that are using data visualization in their research; not teaching only positions.
2. Please select from the list below the department/school where your primary job appointment is located.	Participants with research experiences from multiple different fields, datasets, and forms of visualizations.
3. In one sentence, please describe your research interests as you would to a layman.	Prepare participants to start thinking about their research.
4. In a brief sentence, please explain in your own words one dataset that you have used in your research.	Prepare participants to start thinking about data visualization by having them reflect on a dataset they have used in their research.
5. Based on the definition of data visualization provided above, what form (types of representations) does data visualization take in your research field?	Understand what visualizations participants are used to creating and interpreting in their research field.
6. Based on the definition of data visualization above, what skills or strategies do you want people to use when *interpreting* a data visualization?	Thinking of strategies and skills as processes of data visualization; Collecting variety of skills and strategies for interpreting visualizations to start finding common skills/strategies among participants in the subsequent Delphi panels.
7. Based on the definition of data visualization above, what skills or strategies do you want people to use when *creating* a data visualization?	Gathering skills and strategies for creating visualizations to start the process of finding consensus among participants in the subsequent Delphi panels.
8. Is there anything else that you would like to add about interpreting or creating a data visualization?	Allowing for study participants to have a place to provide comments and/or additional feedback about data visualizations.
9. Would you describe yourself as a data scientist?	Need to gather their perspective because I am labeling my sample as data scientists based on data science being an emerging field that encompasses statistics, mathematics, and computer science.
10. Fork from #9, What would you describe yourself as?	Would like to have their perspective on their title if it is not a data scientist.
11. Fork from #10, Would you be interested in sharing a data visualization from your research with the study researcher?	Potential to use those visualizations in the next survey for Delphi Panel 2.

Table A2. Framework for Delphi Panel 2

Question	Rationale
1. What do you perceive is the overall interpretation of the visualization from your perspective?	Have study participants think about the overall story of the visualization based on their interpretation.
2. What do you perceive is the purpose of the visualization from the perspective of the creator (aka. the website)?	Have study participants think about the intended purpose of the visualization.
3. What are the skills/strategies you used when interpreting the visualization on the HealthMap website?	Prepare study participants to start thinking about skills/strategies that they used in order to prepare them for questions 4 and 5.
4. Please select all of the skills/strategies from the list below that you used when interpreting the visualization.	Building consensus among study participants based on the skills/strategies identified from Delphi Panel 1.
5. Please rate the items according to the level of importance when interpreting the visualization.	First step at trying to understand the process of interpreting the visualization and using importance as a proxy for order.
6. Would you like to comment on any of the skills/strategies that were identified from the first survey about interpreting a visualization?	Allowing for study participants to have a place to provide comments or feedback about the skills/strategies identified to work towards building consensus
7. Are there any additional comments that you would like to provide regarding interpreting a visualization?	Allowing for study participants to have additional opportunities to comment in general about interpreting visualizations.
8. Think back and recall a recent visualization that you have created, please describe the visualization in a few sentences.	Prepare study participants to think about the process they engage in when creating a visualization.
9. Please select all of the skills/strategies from the list below that you use when creating a visualization.	Building consensus among study participants based on the skills/strategies identified from Delphi Panel 1.
10. Please rate the items according to the level of importance when creating a visualization.	First step at trying to understand the process of creating the visualization and using importance as a proxy for order.
11. Would you like to comment on any of the skills/strategies that were identified from the first survey about creating a visualization?	Allowing for study participants to have a place to provide comments or feedback about the skills/strategies identified to work towards building consensus.
12. Are there any additional comments that you would like to provide regarding creating a visualization?	Allowing for study participants to have additional opportunities to comment in general about interpreting visualizations.

Table A3. Framework for Delphi Panel 3

Question	Rationale
1. After looking at the overall ratings and your individual rating for each skill/strategy about interpreting a data visualization, would you like to adjust any of your ratings?	Intention of the Delphi method is to bring back the results from the previous panel to study participants to achieve consensus/agreement among the data, so this serves the purpose for skills/strategies about interpreting a visualization.
2. Please move the orange slider for any rating you would like to adjust.	Allows the study participant to either keep their rating the same or adjust it if they feel like their rating is outside of the consensus for that skill/strategy about interpreting a visualization.
3. After looking at the overall ratings and your individual rating for each skill/strategy about creating a data visualization, would you like to adjust any of your ratings?	Intention of the Delphi method is to bring back the results from the previous panel to study participants to achieve consensus/agreement among the data, so this serves the purpose for skills/strategies about creating a visualization.
4. Please move the orange slider for any rating that you would like to adjust.	Allows the study participant to either keep their rating the same or adjust it if they feel like their rating is outside of the consensus for that skill/strategy about creating a visualization.

Table A4. Descriptive statistics for ratings of interpreting skills/strategies from Delphi Panel 3

Skill/Strategy	*N*	Mean	*SD*	Min.	*Q1*	Median	*Q3*	Max.	*IQR*
Understanding the layout of the visualization.	12	4.08	0.64	3.00	4.00	4.00	4.25	5.00	0.25
Reading axes.	11	4.82	0.37	4.00	5.00	5.00	5.00	5.00	0.00
Reading captions/text.	11	4.64	0.48	4.00	4.00	5.00	5.00	5.00	1.00
Understanding the definition/meaning of variables displayed.	11	4.27	0.86	2.00	4.00	4.00	5.00	5.00	1.00
Drawing comparisons among variables.	9	4.22	0.63	3.00	4.00	4.00	5.00	5.00	1.00
Constructing meaning from the visualization/gaining insight.	7	3.86	0.64	3.00	3.50	4.00	4.00	5.00	0.50
Identifying a purpose of the visualization.	11	4.00	0.85	3.00	3.00	4.00	5.00	5.00	2.00
Exploring data by interacting with the visualization.	8	4.00	1.12	2.00	3.00	4.50	5.00	5.00	2.00
Critical thinking skills	7	3.71	1.03	2.00	3.00	4.00	4.50	5.00	1.50
Comprehension of statistical methods.	3	2.67	0.94	2.00	2.00	2.00	3.00	4.00	1.00
Interpreting the visualization within a larger context within the field.	3	3.33	0.47	3.00	3.00	3.00	3.50	4.00	0.50
Understanding the methods of data cleaning and data staging.	2	3.50	1.50	2.00	2.75	3.50	4.25	5.00	1.50
Additional skill/strategy: #Intuition and experience.	1	4.00	0.00	4.00	4.00	4.00	4.00	4.00	0.00
Comprehension of machine learning techniques.	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Making extrapolations of the data.	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table A5. Descriptive statistics for ratings of creating skills/strategies from Delphi Panel 3

Skill/Strategy	*N*	Mean	*SD*	Min.	*Q1*	Median	*Q3*	Max.	*IQR*
Defining the purpose of the visualization.	11	4.55	0.66	3.00	4.00	5.00	5.00	5.00	1.00
Aesthetic sense.	10	4.00	0.63	3.00	4.00	4.00	4.00	5.00	0.00
Designing visualizations with clear and efficient meanin.g	10	4.30	0.46	4.00	4.00	4.00	4.75	5.00	0.75
Highlighting main points/patterns (relationships/trends).	10	3.80	0.75	3.00	3.00	4.00	4.00	5.00	1.00
Facilitating comparisons among graphs in a visualization.	9	4.22	0.63	3.00	4.00	4.00	5.00	5.00	1.00
Thinking about the research questions from the study/experiment.	8	4.00	0.71	3.00	3.75	4.00	4.25	5.00	0.50
Using color to highlight multiple variables.	7	4.29	0.70	3.00	4.00	4.00	5.00	5.00	1.00
Critical thinking skills.	6	4.50	0.50	4.00	4.00	4.50	5.00	5.00	1.00
Labeling all aspects of the visualization (axes, legends, etc.).	5	4.80	0.40	4.00	5.00	5.00	5.00	5.00	0.00
Scaling axes appropriately.	5	4.20	0.75	3.00	4.00	4.00	5.00	5.00	1.00
Using story telling techniques.	5	4.80	0.40	4.00	5.00	5.00	5.00	5.00	0.00
Thinking about the intended audience of the visualization.	12	4.25	0.92	3.00	3.00	5.00	5.00	5.00	2.00
Visualizing multiple variables at once.	7	4.14	0.83	3.00	3.50	4.00	5.00	5.00	1.50
Knowledge of foundational statistics concepts.	7	4.14	0.99	3.00	3.00	5.00	5.00	5.00	2.00
Considering the context of the data.	7	3.86	1.13	2.00	3.00	4.00	5.00	5.00	2.00
Engaging in an iterative process when creating the visualization.	4	3.50	0.50	3.00	3.00	3.50	4.00	4.00	1.00
Quantifying variability.	3	4.33	0.47	4.00	4.00	4.00	4.50	5.00	0.50
Transparency of information about the data.	3	4.33	0.47	4.00	4.00	4.00	4.50	5.00	0.50
Using symbols to highlight multiple variables.	3	3.67	1.25	2.00	3.00	4.00	4.50	5.00	1.50
Multidimensional geometry.	1	3.00	0.00	3.00	3.00	3.00	3.00	3.00	0.00
Constructing dynamic/interactive visualizations.	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00