This is the authors' post-print. The definitive IEEE PacificVis 2024 publication will appear in IEEE Transactions on Visualization and Computer Graphics.

Struggles and Strategies in Understanding Information Visualizations

Maryam Rezaie, Melanie Tory, and Sheelagh Carpendale

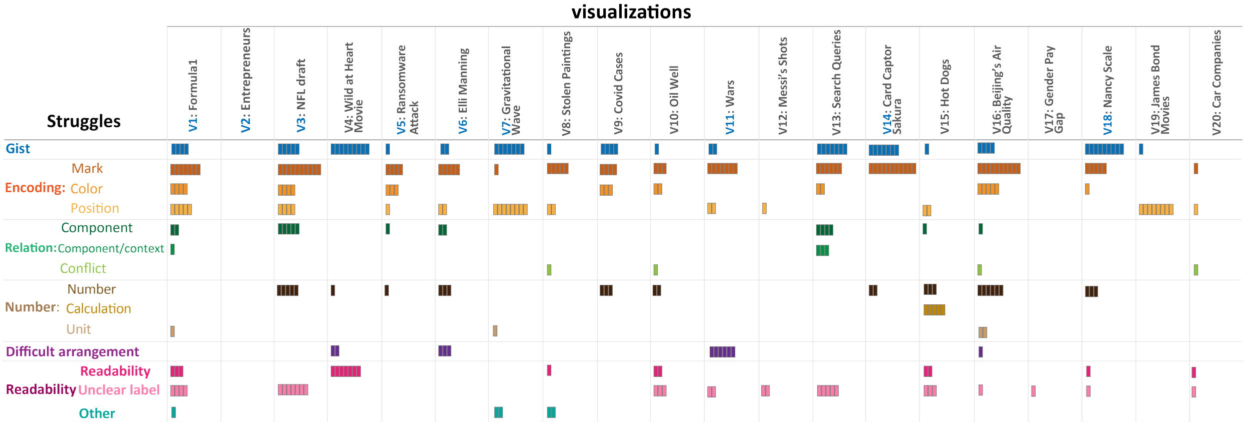

Fig. 1: Struggles identified in the sensemaking of information visualizations in this research. The titles with colored V# indicate that the corresponding visualization is interactive. Length of a bar represents frequency of the struggle. The smallest bar represents one occurrence and the largest represents 11 occurrences.

Abstract—While the visualization community is increasingly aware that people often find visualizations difficult to understand, there is less information about what we need to do to create comprehensible visualizations. To help visualization creators and designers improve their visualizations, we need to better understand what kind of support people are looking for in their sensemaking process. Empirical studies are needed to tease apart the details of what makes the process of understanding difficult for visualization viewers. We conducted a qualitative study with 14 participants, observing them as they described how they were trying to make sense of 20 information visualizations. We identified the challenges participants faced throughout their sensemaking process and the strategies they employed to help themselves in overcoming the challenges. Our findings show how details and nuances within visualizations can impact comprehensibility and offer research suggestions to help us move toward more understandable visualizations.

Index Terms—Information visualization, visualization sensemaking, qualitative study

1 Introduction

While the use and prevalence of visualization continues to expand, the visualization research community is increasingly aware that people often struggle to interpret and use visualizations effectively [39]. The flourish of research into visualization literacy and the challenges faced by people reading unfamiliar visualizations is a testament to the importance of this issue [49, 56]. The visualization research community has put substantial effort into developing visualization training materials [23], literacy tests [38], and guidelines [64] in effort to begin to address these issues. Continuing this direction, we study how people unravel information visualizations (InfoVis) when unassisted.

To realize a future in which visualizations are more readily understandable, we need to first understand in more depth the details of what happens when a person tries but fails to understand a visualization. A recent study by Lee et al. [37] made considerable headway by proposing a sensemaking framework to explain how people interpret unfamiliar visualizations. They observed, at a high-level and without details, that people flounder when they cannot build a mental frame (i.e an understanding of the content and visual encoding) that matches the reality of what they see in the visualization. However, we still do not know what happens while a person is floundering - can they flounder and recover? does floundering always lead to failure? and exactly what are the details in this process? We conducted a qualitative empirical study to expand this understanding. We observed how 14 participants worked towards making sense of a diverse set of 20 InfoVis examples on their own with no additional training, support and given no specific tasks. Through a fine-grained analysis of our extensive data, we observed the details of people's challenges, struggles and strategies with interpreting different aspects of the visualization.

We identified six high-level types of *struggles* as well as five high-level *strategies* that participants employed to overcome their struggles. We look at what leads to people floundering when reading an InfoVis, what happens during floundering and what leads to people giving up without reaching an understanding. By looking at the details of struggles and strategies that people face in their sensemaking process, we identify 6 practical research objectives directions to move this research forward towards human-informed approaches to the design of more understandable visualizations. Our work makes the following contributions:

- Descriptions of the nuances that contribute to people's confusion when they decipher unfamiliar visualizations;
- Insight into how struggles and strategies can impact InfoVis sensemaking;
- Evidence of how these struggles often linked to details of the visualization;

- *Maryam Rezaie is with Simon Fraser University. E-mail: maryam_rezaie@sfu.ca*
- *Melanie Tory is with Northeastern University. E-mail: m.tory@northeastern.edu.*
- *Sheelagh Carpendale is with Simon Fraser University. E-mail: sheelagh@sfu.ca.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

- Descriptions of how unsuccessful strategies can lead to frustrations and abandonment comprehension goals; and
- Six research objectives (RO) to encourage future work towards developing more comprehensible visualizations.

2 Related Work

Our research extends prior research in visualization insight generation and sensemaking, especially for viewers who are unfamiliar with factors such as the data domain, visual encoding, or available interactions.

2.1 Insight generation

Gaining insight into the dataset is one of the main purposes of visualization [14]. Therefore, extensive research has been carried out on the definition of insight [54], its characterization [15, 32, 46, 65], quality of insight [54], how people gain insight [65] and factors that affect insight generation [30, 51]. Also research has been done on evaluating visualizations based on the quantity [30] and/or quality of insights [32]. Visualization systems that automatically generate and recommend insights are proposed in the literature (e.g., [18, 21]).

A recent study [35] argues that additional domain information should be considered in insight-based evaluations. Similarly, different levels of understanding a visualization, grounded in Bloom's taxonomy of educational objectives [7] were suggested for evaluating a visualization [11], particularly in the context of task-based evaluation. Previous work on the insight generation processes have not yet provided evidence of what barriers affect insight generation and what people do to overcome the barriers when viewers are not given specific visualization tasks and when they explore the visualization freely.

2.2 Sensemaking Practices

It is important to look at insight generation as part of a bigger process: sensemaking. Sensemaking can be defined as the activities and behavior people do in order to understand a complex information space [52]. Various models of the human sensemaking process have been proposed as foundational to understanding how people use data and visualizations in the construction of knowledge [28, 50]. These models assume that people continually update their internal frame of reference as a new evidence comes to light. For example, Grolemund and Wickham [28] describe how people create an internal schema, which they then iteratively confirm, update, or reject based on new insights.

Studies have also considered strategies people take to solve problems in the visualization sensemaking process. For instance, Mayr et al. [45] explored different strategies to solve tasks when exploring information visualizations. Similarly, various individual and group strategies within information analysis processes have also been explored [2, 20, 31, 33, 41]. These studies focused on the exploration stage, when participants already knew how to read and interact with the visualizations; participants were given instructions or training materials about what was available in the visualizations.

We focus on the sensemaking practices of people when they encounter a visualization for the first time – without the advantages of training or suggested tasks. Experiences and data practices of less skilled data workers and the public have been investigated in both professional and personal contexts. Many studies focused on spreadsheets as the primary medium within which data workers prepare, organize, and analyze data [4, 17, 24]. Others revealed frequent tool switching and non-linear sequencing [5, 60]. For example, Tory et al.'s [60] study of dashboard users revealed how people frequently dumped data out of dashboards to enable more flexible work practices in more comfortable tools. They characterized breakdowns and strategies, similar to this paper's struggles and strategies, but at a higher level of abstraction, looking at data work practices rather than visualization interpretation.

More pertinent to our work are studies exploring the mechanics of how novices interpret and make sense of data in visualizations. Grammel et al. [27] examined barriers for novice users, but focused on the visualization construction process rather than visualization interpretation. Blascheck et al. [6] explored strategies people employed to discover available interactions in a novel interaction-rich visualization.

Most relevant is the work of Lee et al. [37], who investigated how novice participants made sense of 3 unfamiliar visualizations. They proposed a framework of sensemaking activities, NOVIS, that consists of five activities: encountering visualization, constructing a frame, exploring visualization, questioning the frame, and floundering on visualization. Of these sensemaking activities, a case study [42] investigates 'constructing a frame', in the context of exploring how visitors of a museum decode visual elements of a complex visualization. The last activity, floundering, which refers to "failing in constructing a frame and not knowing what to do" is reported as one of the observed activities of the sensemaking 'process'. It is reported that some participants experience challenges and try to respond to those but they give up understanding the visualization once they flounder. *Our research builds on this work to directly investigate those challenges by characterizing the types of sensemaking struggles that cause viewers' confusion – including whether they decide to give up on understanding the visualization –and the strategies people applied to overcome those struggles.*

3 Our Qualitative Study

We designed our qualitative study to better understand the minutiae of people's confusion in their visualization sensemaking processes. Our goal was to explore the challenges people faced as they worked with visualizations without being trained or assigned any specific tasks. We observed participants as they tried to understand 20 information visualizations for the first time. Below we describe the visualizations, our participants, the procedure, data collected, and our data analysis.

3.1 InfoVis Examples Used in the Study

Our set of 20 diverse InfoVis samples were showcased and publicly available with open access on distinguished platforms, such as respected visualization galleries and websites: *Information is Beautiful* (7), *Tableau* (6), *Behance* (2), *Visual capitalist* (1), *Fivethirtyeight* (1), *Reuters* (1), *ScienceNews* (1), and *Reddit* (1). Publishers on these sites tend to be professional visualization practitioners sharing completed work. As part of our selection process, we considered diversity across the following criteria:

1) Diversity in the context of the datasets: Previous work has reported that having background knowledge about the context of the dataset affects sensemaking. We included a diverse set of visualization contexts, so that most participants would be familiar with some but not all of the contexts. These contexts included: COVID-19 data (V9), Messi's shot attempts (V12), and major car companies (V20), as well as instances that might be unfamiliar to many participants, such as NFL draft (V3, V6), gravitational waves (V7), and Cardcaptor Sakura (V14).

2) Interactive/static: We wanted to know how people understand visualizations when interaction is not available (static visualizations) and how they use available interactions to understand visualizations. V1, V2, V3, V5, V6, V7, V11, V14, V18 were interactive.

3) Unfamiliar visual encoding: Our examples were selected to extend our understanding of sensemaking with unfamiliar encodings. Except for V2 (line chart) and V5 (bubble chart), all other examples had elements of unfamiliarity, such as unfamiliar representation (e.g. V14), unusual presentation (e.g. V4, V6) or some deviations from standard charts such as icons as visual marks (e.g V15).

4) Visualizations with more than one type of visual mark: While previous works investigated sensemaking of visualizations with one type of visual mark, we included visualizations with more than one type of visual mark. For example, V16 (shown in Figure 2 a) includes points, area, and bars (line) as its visual marks. Similarly, V1, V8, V9, V13, and V14 are featured with both lines and points as visual marks.

Each InfoVis example included a title that described the topic. All interactive features were familiar interactions such as tooltips, filters, and highlighting. Ten visualizations had visible legends or how-to-read parts. The 20 InfoVis examples are listed in the supplementary material. In our post-study interview, none of the participants reported seeing any of these 20 visualizations prior to the study, but all reported familiarity with line (V2) and bubble chart (V5) encodings.

3.2 Participants

We recruited 14 participants (8 female, 5 male, 1 declined declaration; ages 18-35) using word-of-mouth snowballing, posting on social media and through flyers at the university. Participants' fields of study were: computing science and technology (7), humanities and cognitive science (3), and art and design (4). While it is possible that student participants were more willing (than others) to spend time and effort working through their struggles and strategies, this provided more opportunities to observe nuances. All participants were familiar with bar charts, scatter plots and line charts and 5 were familiar with parallel coordinates and cord diagrams. While none of the participants were visualization designers, all had experience creating basic charts and 2 considered themselves proficient with Tableau and D3.js.

3.3 Apparatus and Study Setting

We used a desktop computer with a 32-inch, 4k monitor equipped with Microsoft Windows as well as a standard mouse and keyboard and an HD USB camera. Using Zoom Video technology, our video recordings included participants' faces, their interactions with the visualizations, and the experimenter's and participants' verbalizations. We developed a simple web page that showed thumbnails of the 20 visualizations in a 4 x 5 grid. Participants could see the visualization at the original size by clicking on each visualization thumbnail. They could also zoom in by hitting control+plus buttons on the keyboard. Below each visualization, a button labeled *more information* or *more information+i* was available in case they wanted to see the original visualization where it was published. The *i* indicates the original visualization is interactive and that they could interact with the visualization.

3.4 Procedure

Pre-study (∼ 5 minutes): Participants were welcomed and asked to fill out a consent form and a pre-study questionnaire about their demographics and their overall background knowledge and experience (see supplementary material) regarding their experience with visualization. The web page and the buttons were explained to ensure that they could freely use the buttons to see the visualizations on the original location and to know which visualizations were interactive. Participants were asked to go through all the visualizations and try to understand them while thinking aloud. Using other visualization examples, the experimenter explained the think-aloud protocol.

Study (∼ 50 minutes): As the participants were working on understanding the visualizations, the experimenter paid attention to the participant's interaction with the visualizations, such as moving the mouse around, zooming, reading (if they were moving the mouse on textual elements, or they were reading out loud), and silence. The experimenter asked questions about what seemed confusing and why they did a specific action and took notes during the study. Participants could spend as much as time they wanted with each of the visualizations.

Post-study (∼ 5 minutes): We asked 7 questions in the post-study semi-structured interviews to collect more information about the overall understandability of the visualizations (see supplemental material). In appreciation of their time, participants received a $20 gift card.

3.5 Data analysis

Our resulting data was rich, with lengthy videos (∼700 minutes). We analyzed the videos instead of the verbal transcripts because the videos captured the participants' interactions with the visualizations as well as their verbalized thoughts. We took an inductive data analysis approach [61] to analyze the data in multiple rounds.

Preliminary analysis (Phase 1) As a preliminary step, one coder went through all the videos to become familiar with the details of the data, writing analytic memos of the unexpected and common observations. In this first inclusive pass, all details, visuals, verbals, actions and expressions were considered and commented upon in the memos.

Following this, a second coder reviewed the memos alongside relevant video clip segments. Together, the coders discussed various perspectives for data analysis that could provide more detail about nuances of challenges and struggles in understanding visualizations. Our initial observations suggested that participants faced various issues in understanding visualizations, and they employed various approaches to overcome those issues but some participants gave up on sensemaking before they fully understood the visualization. Noting that a similar observation was briefly reported as "floundering on a visualization" in previous work [37], we decided to investigate whether more light could be shed on these moments of confusion and frustration.

Focused Coding Based on the understanding we developed in phase 1, we focused on two high-level codes – the *struggles* and the *strategies* – faced in deciphering a visualization. A *struggle* is the expression of needing some more information to understand some part of the visualization during in the sensemaking process. For example, P14 asked *"what does this line mean"* about V11. A *strategy* as a unit of analysis is one or more *actions* that participants took to make an attempt to understand a visualization – for example, moving the mouse randomly on the visualization or reading tooltips. Iteratively, following the guidelines of Inductive analysis of the data [61], the first coder went through the data again and elaborated the coding schema into fine-grained ideas within each high-level category. For example, when P11 asked, *"I don't understand what red and blue mean here"*, the basic assigned code was *struggle: understanding color*. When they tried to read available text to resolve this confusion, it was coded as a *strategy: reading short sentences*. The first coder discussed initial observations and codes with the other researchers to clarify the codes and to identify possible categories for grouping fine-grained codes.

Synthesis and Analysis In this iteration of data analysis, in a comparative process, the first coder went through the fine-grained codes and categories, and discussed and compared the codes with two other researchers in multiple sessions to confirm the codes and resolve disagreements following guidelines for reaching agreement [12]. Consequently, to enhance clarity and prevent redundancy, some codes were grouped together. For example, labels and sentences were merged to be considered as short textual elements. Conversely, certain codes were overly general when compared to other codes. For example, visual search encompassed general ideas which were broken down into subcategories. Finally, the 3 researchers (in an iterative process) carefully examined how the observed outcomes aligned with findings from the literature. The distilled insights and avenues identified for further research are elaborated upon in Section 5.

4 FINDINGS

Generally our participants approached the study with good humor and a puzzle-solving mindset. However, they did go through multiple struggles and tried various strategies during their process of understanding visualizations. In this discussion, for participant comments, we use PxVy to refer to a quote that participant x said while understanding visualization y. In discussing participant struggles and strategies we use the word *marks* to refer to visual marks, such as points and lines, and *visual variables* to refer to the ways designers controlled the appearance of marks such as colors, length, position, etc. We also use the word *component* to refer to the parts a visualization could be decomposed into, as perceived by some participants, also known as visual chunks [9,40]. Figure 2 shows examples of these types of components.

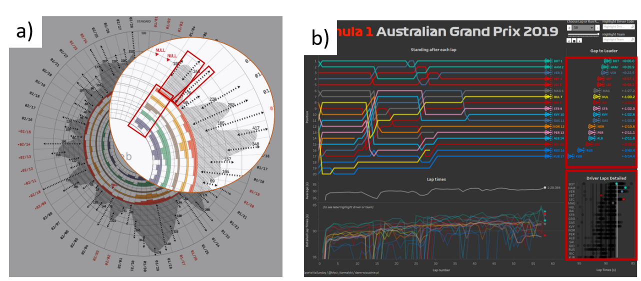

Fig. 2: Examples of perceived components of the visualizations: a) a group of bars, an area, and directed dotted lines are 3 examples of various components that p9 perceived in V16; b) P4 also perceived views in a linked-views type of visualization (V1) as separate components

4.1 Struggles

We discuss struggles under five headings: encoding (145), gist (69), readability (46), numbers (39), components and their relationships (27), and spatial arrangements (13). The bracketed number indicates the occurrences of this type of struggle. The frequency of these struggles and how they are associated with which visualizations is shown in Figure 1.

4.1.1 Struggles: Visual Encoding – Marks and Variables

Struggling to understand the data encoding when combining all variations of visual marks and variables, had 145 struggle instances with the variables position (40) and color (24) being the most frequent. Understanding the visual encoding was particularly challenging when there was no visible legend. However, having a legend did not guarantee that participants could understand the visual encoding. Participants struggled to understand what data entity was represented by a visual mark, saying something such as, "what does each line show" or "what is each dot".

Visual variable: Position. Participants struggled to understand the structure of marks when the axis and/or labels were missing (e.g., V7) or placed in a way that was not clear with what the label is associated (e.g., V1). Many participants simply asked what the x-axis and/or y-axis are, which showed their struggle with understanding the meaning of axis, especially when the elements clearly followed a horizontal or vertical pattern. Interestingly, with some visualizations in which there were no visible axes but the marks seemed to be organized horizontally (e.g. V1, V7, V19), many participants assumed that the horizontal position encoded a data attribute. For example, 4 participants assumed the x-axis was a timeline in V7. P12V7 said: *"x-axis is a timeline ... There is a flow [of events] but I don't know whether it started from left to the right or right to the left?"*. In visualizations where there was a visible axis, participants struggled only when the label of the axis was missing or its contextual meaning was not clear.

Visual variable: Color. Participants struggled to understand the meaning of colors when they were not explained in a legend. Even with a color legend, participants sometimes struggled to make sense of colors when the labels were not sufficiently descriptive or when the use of the color was too far apart to support comparison. For V1 (Formula 1 Racing) where color was used to show teams of cars, some participants struggled to understand the colors. P3 stated: *"I don't understand why certain colors are chosen for certain cars"* and guessed either they chose the colors for the cars randomly or the designers wanted to show some kind of grouping. Later by interaction, P3 realized that colors show teams of drivers. P4V1 also mentioned *"These are two different things so why do they have the same color?"* In addition, redundantly encoding a data attribute with color and position caused some participants to struggle in finding meaning. This result is rather surprising because visualization design principles typically suggest that redundant encoding is helpful rather than harmful, as previous studies suggest that multiple representations are beneficial for learning new ideas [1] P3V5 said: *"I don't know why they used colors when they are already organized spatially"*. Then, P3 hypothesized that they used color in case one would change the position of the colored circles for some reason, then they would still know to what group the circle belongs. Furthermore, in V9 where color was used to categorize data points, P13 struggled to understand some uncolored marks.

4.1.2 Struggles: Gist of the Visualization

At the beginning of their sensemaking process, generally participants wanted to know the basic idea or topic in order to have something to relate their developing understanding to. Struggling to uncover the gist (topic, focus and purpose) of the visualization - was a common struggle (69 instances). It occurred at least once for all visualizations except V2, V17, and V20 and for most of the visualizations many participants struggled to discover the gist (such as 10 participants for each of V4 and V18) They expressed this kind of struggle with statements such as *"looking at this, I have no idea what this is saying at all" (P3V7).* Since understanding the visualization's topic can provide a basis for their sensemaking, participants usually found this important enough to attempt to solve it first. For some visualizations, P13 took the notion of gist a step further and tried to find the *message* the visualization was trying to convey. They said, *"Ok, there are [covid] cases, their contacts, and their relationships, but what should I conclude from this chart?" (P13V9).* Similarly, they mentioned, *"...So there are some violence, there are some sex scenes, and some violent scenes in the movie. This doesn't have any information about anything. I don't know why would anyone show a movie like this." (P13V4)*

4.1.3 Struggles: Readability Issues

In total we saw 46 instances of readability issues - 29 of which were about unclear labels. Readability issues varied considerably both from participant to participant as well as from visualization to visualization. These issues included: clutter and overlapping visual marks; colors being perceived as indistinguishable; bright colors on a dark background being painful to look at; difficult fonts; difficulty due to sizes of icons; and challenging placement of labels - raised questions about what the label belonged to. Some participants took extra time and effort to connect legend items with visualization components, especially when there were several visual variables. For example, P5 could not find the solid line that the legend of V16 was describing – causing them to go back and forth between the legend item and the chart for comparison.

Some characteristics such as clutter, irritating colors, and small fonts can affect people's comprehension and overall experience influencing their willingness to spend more time with the visualization. As noted in previous work [63], these issues are expected to be problematic in the sensemaking process. There were some instances of not noticing a textual or visual component entirely. For example, P13 noticed the title of V1 upon their 2nd attempt to understand it and said, *"...the Formula one, oh, it has a title, I have a problem with reading titles!" (P13V1).*

Unclear Labels When labels are not clear or not descriptive enough, that can increase confusion. This struggle often follows another struggle, when participants try to use labels to overcome their lack of understanding. In some cases, the label itself was not understandable (e.g. 'POLYUNSATURATED' in V10). However, in many cases the label itself was understandable, but participants could not relate it to the context. For example, in V13, which was about the social credit system - a Chinese national reputation system and people's approaches toward it, different categories of approaches are shown by color. On the legend, the labels of the colors were *neutral*, *positive*, *negative*, etc., without further explanation, which made it hard for people to relate them to the context of visualization. *"I don't know what* approaches *is, is it the approach of the website that is negative? or is it the query?" (P13V13).* Since labels are associated with some visual marks or variables, unclear labels were often associated with struggles related to visual encoding.

Other Issues Sometimes participants invented meanings for factors that did not embed data. For instance, V16's bars associated with pollutants were placed in a concentric circular arrangement, and participants assumed that the size of the circles represented data. This generated incorrect understandings *"..and the PM2.5 is the most dangerous one [pollutant]; because it [the ring on which the associated bars were placed] is the biggest circle" P8V16.*

4.1.4 Struggles: Numbers

We observed that participants notice numbers and often choose to understand them before other items in the visualization. In total there were 39 struggles with numbers – most being generally about number but 5 were specifically about how a calculation was done and 4 were about units of the numbers. Many times participants simply asked *"what is this number"* and they wanted to know how to think about the number in regards to the context of visualization. This struggle relates to the participants trying to establish a relevance in terms of the visualization. These struggles (30 of them) may relate to trying to build a semantic model. For most of the visualizations the number was explained in the textual elements explicitly. For example, in V14, *50 chapters of adorable cuteness* was written on top of the visualization, far from where the numbers 1-50 were placed; for those who missed reading this, it was hard to understand the numbers.

Calculation method: With numbers such as score, quality index, etc. that implied "calculation" instead of measurement, some participants strove to understand whether or how these numbers were calculated in relation to existing factors in the visualization. For example: *"AQI is based on which of the parameters (pollutants)? is it the average of all of them?" (P3V1)*.
Unit of measurement: Many participants cared about and struggled to understand the unit of measurement of a number. To find that, they scanned the visualization.

4.1.5 Struggles: Components and Their Relationships

As part of sensemaking, people collect information components and try to piece them together. This is a process of finding the relationships between the pieces of information. We observed that participants sometimes struggled to relate different types of information pieces. Most frequently these occur: between 2 or more components of the visualization (16); between a component and its context (4); when there is a conflict between the understandings from 2 or more components (4); and when the participants think there is a missing component (3).
Relation between a component and its context: When participants tried to understand a component of the vis, sometimes they struggled to find the relation between the newly found item and the visualization topic or the knowledge they have built so far in their sensemaking process. For example, when P3 discovered that the interactive visualization V7 includes sound elements, they asked *"what does the sound mean in terms of the gravitational waves?" (P3V7)*. Similarly, P3 stated: *"I don't understand why it is talking about the 1984 and Black Mirror" (P3V10)* when they couldn't understand how one legend item related to the context of the visualization.
Relation between two components/visual marks: Some participants could understand components (depicted in Figure 2) separately, but could not find the relation between them. For example, P4 said, *"It seems these are two completely different charts" (P4V1)*. Similarly P10 said, *"I'm trying to understand what is the relation between this query and these blogs" (P10V13)*. This shows that they understood the *query* as a text block and *blogs* represented as a column of urls; however, they could not figure out how the components were related. Five participants did not understood how the *references* component of V13 was related to the context or other components. Understanding how two visual marks were related was also challenging. For example, *"I'm struggling to understand the relation between this area and these bars on top of it" (P9V16)*.
Conflict in the relation between two understandings: There were cases where understanding a visual mark conflicted with a participant's knowledge of the context. For example, the bend in the lines in V6 spanned 20 years on the y-axis. P1 guessed this meant the theft took place over 20 years. However, this was the case for all the bends of the lines on the visualization, which seemed unreasonable as it meant all thefts in this visualization took 20 years exactly. As another example: *"I cannot understand what the solid lines mean, legend says air quality standard guide; if it's the standard level, why do we have a lot of them? [...] standard is usually one number, is it about a specific day? like, warmer days are different. I don't know" (P5V16)*. In this case, there were different standard levels because there were different pollutants in the chart. It is likely that P5 could not find the relation between these solid lines and the pollutants, which are shown by colors.

These examples illustrate fine-grained details of how people extract information from visualizations and then reconcile that information with their internal frame, either re-assessing the observation or updating the frame in the case of conflicts. Lee et al. [37] also observed people re-assessing their frame, but our data here reveals more explicit details about participants' internal reasoning and struggles in these scenarios.

4.1.6 Struggles: Difficult Arrangements

The arrangement of the marks could cause participants' confusion in visualizations that did not include explicit or implicit explanations for such arrangements, though with much less frequency: 7 instances of line arrangement struggles; 4 instances of confusion caused by circular arrangements; and 2 instances of separation struggles .

Continuity of connected lines: The arrangement of lines in V11 (Figure 3 a)) was confusing. As P3V11 stated, *"I don't know why all lines are stemming from 4...it directs my eyes to it"*. Six participants struggled to find a reason why all lines were connected to number 4, even before they knew enough about the visualization to know what the lines were showing. This relates to Gestalt principle of continuity which states human eyes follow lines from beginning to end [59].
Circular arrangement: The circular arrangement was confusing to many participants especially when the circle represented a timeline. For example, *"why a circular arrangement? are we going to watch the movie again? or the story comes back to the beginning?" (P6V4)*, or *"I think a circle is probably not the best idea for something that's supposed to be explained over time because when I think of time I think of something linear" (P3V4)*.
Component separation: In some visualizations, some of the visual marks did not follow the same arrangement as others. For instance, a group of aligned visual marks is slightly separated from other visual marks in V6 (Figure 3 b)) to show Eli Manning's playoff games. P6 figured this out accidentally by interacting and choosing another view of the chart where the separated elements were labeled. P8 noticed this but could not figure out the reason for this arrangement.

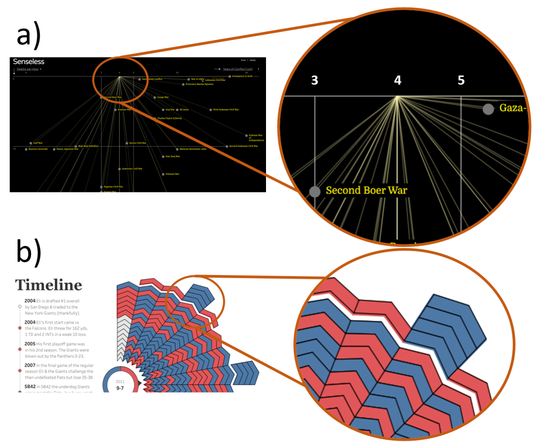

Fig. 3: Confusing spatial arrangements: a) all lines are connected to a point number 4 on the x-axis in V11 and the meaning of this connectivity was not clear; b) some elements were spaced out to distinguish regular and Superbowl games in V6 without clear explanation.

4.2 Strategies

Participants applied various strategies to overcome their struggles.

4.2.1 Strategy: Making a Guess

The most frequent strategy was making a guess (176 instances). Participants sometimes made guesses without gathering more information if the topic could be understood from the title. For example, P1V3 said: *"It is about players and their goals"* after reading the title "NFL Draft" which was not what the visualization was actually about. More frequently, participants collected more information before making an educated guess. Participants varied in what information they chose to collect. Collection involved 3 high-level activities: visual search, picking something to read, and incorporating more information.

- **Visual search:** in this activity, participants visually scanned the visualization for more information and to know what was available to them. P3 mentioned, *"I let my eyes decide. I try to understand what stands out first"*. Based on what was available on the visualization they searched in different ways:
  - Eye movements: they looked at different visual elements of the visualizations without looking for their meaning at that time.
  - Random interaction: they randomly moved the mouse around the interactive visualization. This helped them notice the elements and affected their process of understanding. For example, seeing

that some visual marks are associated with tooltips helps them realize these are the main entities of the visualization.
  - Scroll: scrolling through the page to see what is available. This activity was only observed for V7, because of its extended height.

- **Pick something to read:** in all cases of making a guess, participants relied on reading some textual elements unless the visualization was featured with icons (e.g. V1 has car icons). Some visualizations had paragraph(s), some had only short sentences and labels and some had with both.
  - Short textual elements: participants picked short textual elements such as axis or legend labels and guessed about a possible topic they could be related to. For example, *"Chinese civil war, Vietnamese war, oh, so it's related to the wars "(P12V11)* or *"sex, violence, Lula's smoking; is it about a movie?" (P5V4)*.
  - Long textual elements: participants read long sentences and/or paragraphs until they could make a guess about the topic. P5 read the available paragraph on V7 until they reached the word *black hole*, skipped reading the rest of the paragraph and said: *"...so, this is about black holes"*.
  - Tooltips text: this strategy was always combined with moving the mouse, which randomly activated the tooltips. Participants picked some words on the tooltips to read (which were mostly short textual elements). Different participants picked different text to read given the same tooltip as many tooltips had many textual elements. This strategy was particularly helpful for V3, in which each dot is an NFL player, and their names were written on the tooltips. P12 realized the dots represent people by noticing the names are changing as P12 as they were reading different tooltips.

- **Incorporate information:** Participants incorporated their own knowledge in understanding visualizations. Our observations, in line with previous works [37,55], suggest that 3 types of knowledge are incorporated into the sensemaking process.
  - Context: knowledge about the context helps people to make a guess, especially about the gist of the visualization and understanding the visual variables. Obviously, to incorporate their knowledge about the context, they must have at least read some textual elements, most often, the title. Participants tended to incorporate the most related attributes in their guesses, especially about the color and size. For example, the title ransomware attack (V5) suggests the size of the visual marks shows the impact of the cost of the attack, while the size of the marks actually represents the size of the organization. Similarly, Participants made guesses about colors immediately if the context was familiar and there were only two color values. For example, red and blue person icons were the main visual marks of V9 whose context was COVID cases. P10 incorrectly guessed red means dead and P1 also incorrectly guessed color shows male and female while the color was actually used to show whether people catch the virus locally or not. Similarly, the conceptual meaning of numbers was also guessed by incorporating knowledge. Where the context of the visualization allowed, the average of certain quantitative data attributes in the visualization was the first guess for the meaning of a number.
  - Visualization: previous knowledge about visualizations can help them guess about the visual encodings. For example, to some participants, a line between two visual marks immediately means some type of connection.

  We also observed how people come up with guesses by a combination of these 3 approaches. For example, *"...I can see green, red and yellow and I saw avocado (a label). So I guess it is about how healthy these foods are" (P5V10)*. Similarly, sequences of these approaches were also observed. For example, P1V14 tried understanding the labels first and read the paragraph as the labels were not clear, P1 said *"I have to read this text because I am not able to understand what these (labels) mean" (P1V14)*.

4.2.2 Strategy: Verification

Participants wanted to verify their understanding at various levels. This ranged from verifying their guesses, to verifying insights gained based on building up an understanding of different parts of the visualization. One interesting strategy that 2 participants employed to understand lines in the visualization was to understand the visual marks connected by the lines and make a guess about possible types of connection between these two components. For example, in V13, where blogs were connected to their website categories, P5 said: *"I guess lines are connecting websites to categories of websites but I am going to do some sanity check...Let's pick something easy to understand, social network is connected to Facebook, yes, I am right."* Out of 174 cases of making a guess, participants *verified* their guesses 49 times as they proceeded with their sensemaking process. Verification was mentioned a few times as "confirming" or "testing" in previous work [37]. Our observations suggest that a verified guess affects the rest of the sensemaking process; if the guess is right, the participants had some solid foundation to continue their sensemaking process (including the explorations). If the guess was proved wrong they may feel frustrated and few participants tried to make another guess. Verification strategies include:

- **Matching numbers with textual elements:** to verify the guess made about meaning of a number, they tried to find a textual element that describes the number. For example, P6v14 saw the numbers and guessed they are chapter numbers, and went back to a textual element to confirm it. P13V16 *"I think it [a number] shows the day and month, I think, from January to February I guess. [noticing the textual labels in the middle of the chart] yeah, Jan to Feb 2013."*
- **Matching visual elements with the numbers by counting:** participants counted the number of visual elements to match it to the numerical range they saw on the visualization. For example, p14V6 counted the number of groups of bars to see whether they correspond to the range 2004-2016 in the paragraph. Similarly, P11 tried to verify whether the circles (marks) on the graph correspond to "50 events" written within the paragraph by counting the circles.
- **Read tooltips:** scanning related information presented in tooltips was another verification strategy. P13 guessed V3 was about a sports team such as a basketball team, then went through around 20 tooltips before seeing the word football confirmed their guess.
- **Size and value comparison:** This strategy especially happened for V5 and V20 (bubble charts where each circle is associated with a number). Participants guessed that the size of each circle corresponds to the number written on the circle or its tooltip. However, P13V20 picked two circles of approximately the same size and then compared their (different) values, disproving the assumption.
- **Color and categories comparison:** This strategy was only observed in V5 in which bubbles were organized spatially by year and most bubbles of each year were colored the same, which suggests that color represents year. However, P6 and P3 picked two bubbles from the same year (based on their spatial region) and realized they have different colors, disproving their assumption.
- **Incorporating background knowledge and reasoning:** Many participants tried to verify their understanding by reasoning, which needed some level of knowledge about the context. For example, the bend on the lines in V8 spanned 20 years on the y-axis. P1 guessed this meant the theft took place over 20 years. However, this was the case for all the bends of the lines on the visualization, which seemed unreasonable to P1, as it meant all thefts in this visualization took exactly 20 years. Similarly, in cases where participants guessed that the x-axis represents a time-line, some participants tried to verify that by recalling the time of familiar movies; For example, P2V19 *"I know some of these James Bond movies so the x-axis shows years"*.

4.2.3 Strategy: Using Legends

Ten of our visualizations had legends and guides intended to help participants understand the visual encoding. Some legends were noticeable, meaning that most participants found and used them at the beginning of their sensemaking. Others were found later, in the middle of sensemaking. Also, V3's guide was only available through interaction. Legends were mostly used to tackle understanding the visual encoding.

- **Find the legend item on the chart:** Participants look at the legend item first, read the label, and try to look at the marks to find what corresponds to legend item. Although this strategy seems easy to apply, sometimes it is challenging. For example, different types of dotted lines were featured in V8 and V16, and P13 and P10 could not find them on the chart easily.

- **Find the attribute on the legend:** Some participants started with the chart, and then tried to find information about marks via the legend.
- **Read the legend labels only**: Many participants used legends solely to read the labels, particularly as a strategy for understanding the gist of visualization, and equivalent to reading short textual elements. When the legend labels were unclear and thus not helpful, participants felt frustrated and left the visualization.

4.2.4 Strategy: Random Interaction

With interactive visualizations, we observed several random interactions. Here participants did not look for a specific answer about the data set; they interacted with a seemingly random selection of marks to see the outcome rather than looking for a specific outcome. These random interactions helped them to know what was available to them.

- **Moving the mouse pointer around on the visualization to see how the elements will react:** This was specifically the case with V14 in which there are many visual elements such as packed circles and lines connected to them. Participants hovered over random marks and followed the lines connected to them.
- **Choosing a random drop-down list entry:** Participants selected random entries to see how the chart would change. This strategy was particularly effective for P3 in understanding color-coded lines in V1. When they selected one of the team names from the drop-down list they observed that only the lines corresponding to that specific team became highlighted, all uniformly colored. Since each line is a driver in V1, this observation helped P3 understand that color represents drivers on the same team.
- **Hovering over random marks and reading the tooltips:** Starting their sensemaking process with this strategy helped participants form an idea about the gist of visualization and to understand visual marks.
- **Moving random marks:** Surprisingly, P3V1, P3V6 and P13V1 attempted to move the visual marks. While this interaction was hard to code from the screen recordings, and as a result there might be more instances of this interaction, in these three cases, participants explained their action as they were attempting to move the marks.

4.2.5 Strategy: Planned Interaction

Planned interactions are those where participants had a specific planned outcome.

- **Interact with component to see how another component changes:** To overcome struggle 4.1.5, P3V1 (several times), P13V1 and P3V14 tried to select a component in the visualization to see how another one was affected. For example, P3V1 interacted with the car icons on the top left view to see how the line in the middle view was affected. P3 said: *"clicking on different cars doesn't impact what the lap time is, so I'm not sure if this [line] references a specific driver or not"*
- **Moving the pointer horizontally or vertically:** In this interesting strategy that was applied specifically in V3, participants moved the pointer horizontally on some subsequent marks and investigated the tooltips to see what changed from one mark to another. In this case, other than some names, the year was also increasing so the participants immediately realized that the x-axis was a timeline.

5 Discussion

While considerable attention has been paid to why visualizations are not instantly understandable, many of the details of exactly what happens still needed filling out. For example, a previous study [37] reported that before failing to understand a visualization, people *floundered* on the visualization – then either managing to understand or giving up. In this study we have opened up this stage revealing the struggles and (not always successful) strategies our participants used when trying to

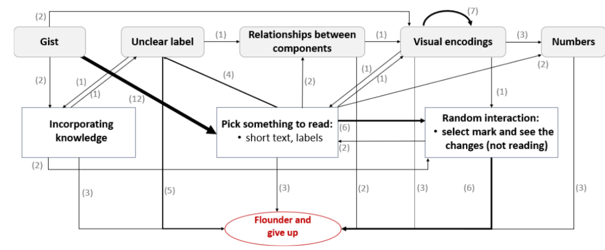

Fig. 4: Transitions between struggles and strategies in instances of floundering and giving up. Line thickness and numbers represent the frequency of these particular transitions within this data subset. Only a subset of the strategies occurred in these specific instances.

interpret new visualizations. Previously, exactly what happens during floundering was not understood. Our research drills into floundering to understand what actually happens and how we might address these challenges. In this section, we reflect on what we discovered and suggest research opportunities (RO) that delineate possible steps to take towards more creating understandable visualizations.

To investigate reasons for failing to understand a visualization, we went through all the videos of the instances of floundering, defined as failing to make a correct frame and giving up on understanding the visualizations [37]. To understand the details of what led to floundering, we extracted the sequence of struggles and applied strategies for these instances. Figure 4 is a diagram of patterns within floundering showing the struggles people went through and the specific strategies they tried as well as the transitions between them. Note that these strategies did not always solve their struggles and they either moved on to another struggle and strategy or gave up. The thickness of the lines shows number of times that transition happened. In all the instances of failure in understanding visualization, participants often tried to read short text and labels which were not helpful in their struggles. Also, they tried randomly interacting with the visual elements in interactive visualizations and they either did not read the text on tooltips or what they read on tooltips was not clear. Furthermore, they attempted to incorporate their own knowledge but their knowledge was not enough for the problem to be solved. Note that this diagram summarises only the *paths* of those who floundered in a visualization.

Figure 4 also shows how participants moved on from trying something out to trying other things or moved on to new struggles. We also looked at sequences of struggles and strategies and we found that people go through a maximum of 4 consecutive struggles before giving up sensemaking. This suggests that participants could tolerate some amount of frustration but there were limits (see supplementary material). Ordered by the number of participants who floundered in the visualizations: V3 (9), V18(8), V7(6), V14(5), V1(4), V13(3), and comparing it to Figure 1 and the characteristics of these visualizations one might have an idea of what struggles are the most important contributors to confusion in understanding the visualization.

5.1 Frustration and Reward Factors in Sensemaking

All participants in our study experienced multiple struggles and applied various strategies. A strategy that is helpful for one participant might not be helpful for another. We found that participants who floundered went through as many as 6 struggles and unhelpful strategies, but only if these pairs were interspersed with some helpful factors. They only tolerated at most 4 consecutive struggles before giving up. In contrast, participants who understood the visualization (at least partially) often got there by applying helpful strategies to overcome their struggles.

The sequences of struggles and unhelpful strategies indicate that viewers have limited tolerance of frustration. However, often participants persisted in the sensemaking when they encountered enjoyable and rewarding factors. The frustration in sensemaking came from *failing* to understand something. The feeling of frustration escalated when they repeatedly tried strategies that were not successful. P3 stated, *"I think it's kind of frustrating to, like, not understand what's happen-*

ing and then also not have anywhere explain better what's happening." Readability issues also add to this frustration, thus also impeding sensemaking. However, participants did feel rewarded when their strategies lead to more understanding, and were much more likely to continue until they understood the visualization, unless struggles and failed strategies arise again. One possibility is that when challenges in a visualization match the skills of participants it can lead to enjoyment [53]. We also observed that being surprised by unexpected discoveries kept viewers engaged in the sensemaking process, even if they discovered they were wrong about something (e.g. P11V16, was very surprised when they discovered that what they thought was the best air quality was actually the worst).

Note that in the case of encountering visualizations in-the-wild (e.g. in news media), people often have no obligation to learn or work with a given visualization, so they may have less incentive to overcome struggles when they occur. On the other hand, exploring a visualization for which one has a deep interest in the context of the underlying data is rewarding and may increase their willingness to try. From our observations, participants countably persevere through more struggles for visualizations towards which they showed enthusiasm.

This enthusiasm can stem not only from their interest in the data but also from the visual presentation that offers an appealing look and feel. This aesthetic aspect of the visualization serves as a form of reward, motivating individuals to persist in their sensemaking process. For instance, many participants said they would like to spend more time with V14 and explore it more, despite its complexity, because the look and feel of the visualization appealed to them. P1 and P3 mentioned they would spend more time with V7, because they liked the first impression of the visualization and how it looks new to them. This implies that the visual appeal of a visualization affects motivation, aligning with previous research in user interface design that emphasizes how visual aesthetics can grasp users' motivation and situate them in interactions [48]. Note that the appeal appeared to be inherent to the visualization itself rather than its content; these participants did not mention anything about their interest in the topic of the data. But even in these cases, participants' tolerance towards confusion and frustration was limited.

RO1 How can we integrate rewarding elements into complex multivariate visualizations, and how does the presence of these elements affect the sensemaking process?

5.2 Background knowledge is a double-edged sword

Previous research mentioned that knowing how to read graphs (a.k.a. graphicacy [55]) and background knowledge about the content of visualization affect graph comprehension [25, 37, 42, 55, 63]; however, it is important to mention that earlier works in graph comprehension mostly reported on its positive effect on reading and interpreting data from relatively simple graphs such as bar and line graphs rather than the challenge of decoding unfamiliar graphs.

Previous works observed how knowledge about the content of an unfamiliar visualization can affect sensemaking [37, 42]. Lee et al. [37] observed how participants recall domain knowledge in this process. Additionally, Dasgupta et al. [19] found that domain experts prefer familiar visualizations over unfamiliar ones even though background knowledge does not necessarily lead to better sensemaking performance.

Effects of background knowledge have been mostly reported as positive, as participants incorporate their prior knowledge to confirm the knowledge gained from the graph, find errors, keep track of the information in the graph, and decode the visual encoding [37, 42]. Also, it has been extensively reported that knowledge about the content helps with the correct interpretation of and prediction from the data [13]. We extended these observations by providing details and examples of how knowledge about the context and graphs can help people with making educated guesses when they face things they do not understand, verifying their guesses, reasoning about their understandings, and piecing information together in visualization sensemaking.

However, previous knowledge about content can also affect sensemaking negatively as it can introduce biases in data interpretation such as relationships between data attributes that are not depicted by the graph explicitly [25]. Additionally, incorrect decoding of color due to prior knowledge, as observed by Ma et al. [42], was a phenomenon we encountered as well, notably in the case of visualization V9. We noticed similar issues with other visual variables such as position (mainly x-axis). In the case of V7 and V19, viewers mistakenly interpreted the x-axis as representing a timeline because time was related to the context of the data: black hole collisions and James Bond movies respectively. Size of the marks was also mapped with the most related attribute to what it was representing. For example, size of the ransomware attack in V5 and the revenue of each car company in V20. These examples illustrate how knowledge of the context can be harmful, when people use that knowledge to make incorrect assumptions that are not verified.

In summary, a certain amount of background knowledge is needed for the sensemaking of a visualization, yet we cannot control the knowledge a viewer brings into the sensemaking process. Therefore, a carefully designed visualization will include necessary contextual information for understanding. But how should such context be presented? In the visualizations within this study, context was typically written in textual elements. In previous studies, other than textual elements context has been provided via media like PowerPoint slides [13]. However, it is unclear whether users encountering visualizations will be willing to engage in tutorials. In addition, our initial evidence indicates that many participants may be unwilling to invest time in reading lengthy paragraphs to gain the necessary context. P9V14 said *"I'm not interested enough in the visualization to read everything. there are lot of text going on here"*. Therefore, more focused research is needed:

RO2 Explore design strategies aimed at presenting context, while taking into consideration individuals' differing degrees of familiarity with the subject matter, and evaluate the efficacy of these strategies.

5.3 Visual encoding: beyond traditional legends

Participants took different approaches in using legends to understand visual encodings and to get what is available in the visualization. Mostly, they started from the chart, noticed a visual variable, and looked it up in the legend or how-to-read part(s). However, there were also many instances in the opposite direction, i.e. noticing a legend item and trying to find it on the chart. We observed many instances of not noticing the legend until the middle of the sensemaking process or not noticing it at all (e.g. V3 whose how-to-read part was discovered only by P6 and P9). Also, some legends were not used in the way they were intended to be used. For example, V14 has multiple how-to-read parts. However, only some of them were read and used by participants. Furthermore, we had many instances of legends, which were said to be hard to use; because it was hard to go back and forth between the chart and the legend - rather than that the legend was unclear. This was interestingly the case for V8 in which there are multiple types of dotted lines to represent different museums.

There are previous works that aimed at moving beyond traditional legends (i.e. boxes of keys and labels placed at the edges of the visualizations). Among them are alternatives such as direct data labeling, legend on titles and tooltips, gaze-based legend adaptation for digital maps [26], derivable legend [66] and visual encoding explanation patterns in data comics [3]. However, there are few studies to assess these types of legends. Edsal et al. [22] compare 3 types of interactions with legends. Modern visualizations, with advanced visual encodings might also need modern legend techniques. In other words, legends themselves can also be considered as data visualization [66] and further explored in research:

RO3 Explore the possibilities of re-thinking legends and 'how-to-read' sections considering factors such as notice-ability, accessibility and proximity.

5.4 Textual elements and numbers

We observed a spectrum of reading the textual elements, ranging from reading randomly selected short textual elements to reading all paragraphs/long sentences. From those sensemaking processes that failed in understanding visualizations to those in which the visual encoding was successfully discovered, we observed many instances of this high-level path of activities: reading -> random interaction. Interestingly, the difference between those who failed and those who succeeded was that the former only read the labels which were often unclear, and within their random interaction, they did not put effort into reading the textual elements that were changing because of the interaction (e.g. a sentence in V14 and tooltip on V3). However, the more successful viewers tended to read more, such as the long sentences and paragraphs as well as the changing textual elements. In addition, when the textual elements contained numbers, we observed attempts to make mental connections between them and visual elements. For example, '50 events' is written in a paragraph of V7 and P11 tried to link it to the bigger circles of V7. Similarly, '32 picks' and '50 chapters' were not clearly linked to each column of dots and the 50 inner circles on V3 and v14 respectively.

The importance of textual elements in the way visualizations are read [36], recognized, and recalled [8] has been discussed. Our observations also confirm the importance of clear titles and labels with regard to the context of visualization. Our observations suggest that the connection between all kinds of numerical and textual elements and the chart should also be clear. Recently, a data-authoring tool proposed highlighting techniques to link the text and primitive charts [57].

RO4 How can we take advantage of interaction techniques such as linking and highlighting to clarify the connections between numbers, textual elements, and the corresponding visual marks?

5.5 Explanations of inter-component relationships

We observed many instances where participants struggled to understand inter-component relationships in multi-component visualizations. In the case of single-view charts, this was a problem of finding the relationship between data variables. In V13 for example, participants wanted to know how the 'reference' component related to other parts of the visualization. Ma et al. [42] also reported that the connection between two of their visualization components was not always clear. In linked-views visualizations such as V1, people wanted to know how these views are connected together. Many of our participants tried to interact with one component of the visualization by selecting a visual mark or choosing from drop-down list to infer the relationships between them from the changes made on other components. Surprisingly, visualization creators rarely included explanations for these inter-component relationships.

To facilitate finding connections between 'data items' of different visual structures such as views in linked-views visualization, previous works such as visLink [16], connectedCharts [62], and composite visualization [34] linked views together using lines and curves. In contrast, we observed that participants wanted to know the 'type' of relationships between these data items across different views and between data attributes of the same data point. For example, P13 in V1, wanted to know how the view on the lower right of the visualization related to other views, while they could simply see the connection between their data items (shown by point marks) because of adjacency and alignment. That view was statistically summarizing the other views was not clear to P13. In Lark [58], the relationship between each derived view and the 'dataset' is shown using an icon; however, the relationships between the derived views themselves were not explicit.

RO5 Explore when in the sensemaking process, where on the visualization, and how the inter-component relationships could be explained to visualization viewers.

5.6 How can we take advantage of random interactions?

One of the most frequently observed interactions was hovering over and selecting randomly picked visual marks and reading tooltips. This happened in all interactive visualizations, even those with clear textual explanations (e.g. V14). This interaction mostly happened at the beginning of the sensemaking process. For static visualizations, participants often asked whether the visualization was interactive after they felt frustrated from not understanding the static visualization. This shows the potential for taking advantage of interactions for more understandable visualizations. P1 requested a specific kind of interaction: *"If I knew about only one visual mark then I could read the visualization"*.

Random interactions may be related to an episode of *trial and error* (T&E), defined as "exploring the interface's available functions" [47]. T&E is reported to be preferred over use of help [47] and performs at least as well as help [43]. This is in line with our observations of how people randomly interact with the visualizations; even when help, as legends and how-to-read parts are available. Different strategies for taking advantage of T&E have been proposed [29, 44] for supporting people in learning complex software, but understanding complex software and web interfaces and visualizations requires distinct strategies. Software interfaces involve task-focused interactions with menus and buttons. Visualization sensemaking often leads to direct interactions with visual elements in addition to menus and buttons which might introduce challenges in adapting strategies from complex interface learning strategies.

RO6 Given that people likely interact randomly, what could we do in the interface such that they get more information out of those interactions?

5.7 Limitations

The diversity of InfoVis examples that we chose is a strength that contributed to the richness of our results. In this research we focused on InfoVis understandability. However, future research could expand upon our work by considering other types of visualizations such as spatial visualizations. With our 20 examples, we were reaching saturation in that we were seeing repeated struggles and strategies. Thus we consider it unlikely that adding additional InfoVis examples would reveal many additional struggle and strategy patterns; however, new patterns might arise with different types of visualizations.

Furthermore, while our student participants represented a community of information seekers from a wide range of disciplines who were familiar with visualization reading and creation to some extent, it is possible that our student participants were more prone to enjoying puzzles and might have been more willing to persevere than would be common. However, this allowed us to see more details about struggles and strategies. For this community, we were reaching saturation and seeing the same things recur. This could change with more diverse communities. We recognize the debate around over-sampling students [10], and that there are communities that have been underrepresented. Future work could include visualization viewers from a greater diversity of demographics, backgrounds, and expertise in reading and authoring information visualizations. In addition, we presented the visualizations in the same order to every participant. Though participants displayed no signs of fatigue in their verbal expressions and interactions, there's a possibility that the sensemaking of the last few visualizations might have been impacted.

6 CONCLUSION

Through our study we have seen considerable evidence about the interplay between presence of struggles, the uncertainty of what might be useful strategies, and the absence of readily available support for sense-making strategies. We have described the nuances that contribute to people's confusion causing them to struggle, often with small details, when they try to decipher unfamiliar information visualizations. Our findings have shown how the interplay between struggles and strategies impact visualization sense-making and how unsuccessful strategies can lead to frustrations and abandonment of visualization comprehension goals. As we noted, there are still many challenging research directions: How can we integrate rewarding elements into complex multi-variate visualizations? Can we discover design strategies that will make presenting context within a visualization possible? Will exploring the design space for novel solutions make it possible for viewers to learn visual mappings as needed? How can we show connections between numbers, textual elements, and corresponding visual marks? Can we explore, within a visualization, the visualization of the inter-component relationships? Or can we make use of the now common response of random interaction to better inform viewers? We invite the community to engage in answering these questions to improve visualization comprehensibility.

Acknowledgments

We thank our colleagues and reviewers for their thoughtful comments and our participants for their time and invaluable input. This research was funded in part by NSERC Discovery Grant Interactive Visualization RGPIN-2019-07192, and Canada Research Chair in Data Visualization CRC-2019-00368.

References

[1] S. Ainsworth. Deft: A conceptual framework for considering learning with multiple representations. *Learning and instruction*, 16(3):183–198, 2006. doi: 10.1016/j.learninstruc.2006.03.001 4

[2] A. Alsaiari, J. Aurisano, and A. E. Johnson. Evaluating strategies of exploratory visual data analysis in multi device environments. In *EuroVis (Short Papers)*, pp. 91–95, 2020. doi: 10.2312/evs.20201054 2

[3] B. Bach, Z. Wang, M. Farinella, D. Murray-Rust, and N. Henry Riche. Design patterns for data comics. In *Proceedings of the 2018 chi conference on human factors in computing systems*, pp. 1–12, 2018. doi: 10.1145/3173574.3173612 8

[4] L. Bartram, M. Correll, and M. Tory. Untidy data: The unreasonable effectiveness of tables. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):686–696, 2022. doi: 10.1109/TVCG.2021.3114830 2

[5] A. Bigelow, S. Drucker, D. Fisher, and M. Meyer. Reflections on how designers design with data. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2598153.2598175 2

[6] T. Blascheck, L. M. Vermeulen, J. Vermeulen, C. Perin, W. Willett, T. Ertl, and S. Carpendale. Exploration strategies for discovery of interactivity in visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(2):1407–1420, 2019. doi: 10.1109/TVCG.2018.2802520 2

[7] B. S. Bloom. Taxonomy of educational objectives: The classification of educational goals. *Cognitive domain*, 1956. doi: 10.1007/978-1-4419-1428-6_141 2

[8] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond memorability: Visualization recognition and recall. *IEEE transactions on visualization and computer graphics*, 22(1):519–528, 2015. doi: 10.1109/TVCG.2015.2467732 9

[9] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete. A principled way of assessing visualization literacy. *IEEE transactions on visualization and computer graphics*, 20(12):1963–1972, 2014. doi: 10.1109/TVCG.2014.2346984 3

[10] A. Burns, C. Lee, R. Chawla, E. Peck, and N. Mahyar. Who do we mean when we talk about visualization novices? In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2023. doi: 10.1145/3544548.3581524 9

[11] A. Burns, C. Xiong, S. Franconeri, A. Cairo, and N. Mahyar. How to evaluate data visualizations across different levels of understanding. In *2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*, pp. 19–28, 2020. doi: 10.1109/BELIV51497.2020.00010 2

[12] J. L. Campbell, C. Quincy, J. Osserman, and O. K. Pedersen. Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological methods & research*, 42(3):294–320, 2013. doi: 10.1177/0049124113500475 3

[13] M. Canham and M. Hegarty. Effects of knowledge and display design on comprehension of complex graphics. *Learning and instruction*, 20(2):155–166, 2010. doi: 10.1016/j.learninstruc.2009.02.014 8

[14] M. Card. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999. 2

[15] E. K. Choe, B. Lee, and m. schraefel. Characterizing visualization insights from quantified selfers' personal data presentations. *IEEE Computer Graphics and Applications*, 35(4):28–37, 2015. doi: 10.1109/MCG.2015.51 2

[16] C. Collins and S. Carpendale. Vislink: Revealing relationships amongst visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1192–1199, 2007. doi: 10.1109/TVCG.2007.70521 9

[17] G. Convertino and A. Echenique. Self-service data preparation and analysis by business users: New needs, skills, and tools. In *Proc. CHI Conf. Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pp. 1075–1083. ACM, 2017. doi: 10.1145/3027063.3053359 2

[18] Z. Cui, S. K. Badam, M. A. Yalçin, and N. Elmqvist. Datasite: Proactive visual data exploration with computation of insight-based recommendations. *Information Visualization*, 18(2):251–267, 2019. doi: 10.1177/1473871618806555 2

[19] A. Dasgupta, S. Burrows, K. Han, and P. J. Rasch. Empirical analysis of the subjective impressions and objective measures of domain scientists' visual analytic judgments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, p. 1193–1204. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025882 8

[20] K. Davidson, L. Lisle, K. Whitley, D. A. Bowman, and C. North. Exploring the evolution of sensemaking strategies in immersive space to think. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2022. doi: 10.1109/TVCG.2022.3207357 2

[21] c. Demiralp, P. J. Haas, S. Parthasarathy, and T. Pedapati. Foresight: Recommending visual insights. 10(12), 2017. doi: 10.14778/3137765.3137813 2

[22] R. M. Edsall, M.-J. Kraak, A. M. MacEachren, and D. J. Peuquet. Assessing the effectiveness of temporal legends in environmental visualization. In *Gis/lis*, vol. 97, pp. 28–30. Citeseer, 1997. 8

[23] M. Elgendi. Scientists need data visualization training. *Nature biotechnology*, 35(10):990–991, 2017. doi: 10.1038/nbt.3986 1

[24] S. L. Erete, E. Ryou, G. Smith, K. M. Fassett, and S. Duda. Storytelling with data: Examining the use of data by non-profit organizations. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 2016. doi: 10.1145/2818048.2820068 2

[25] E. G. Freedman and P. Shah. Toward a model of knowledge-based graph comprehension. In *International conference on theory and application of diagrams*, pp. 18–30. Springer, 2002. doi: 10.1007/3-540-46037-3_3 8

[26] F. Göbel, P. Kiefer, I. Giannopoulos, A. T. Duchowski, and M. Raubal. Improving map reading with gaze-adaptive legends. In *Proceedings of the 2018 acm symposium on eye tracking research & applications*, pp. 1–9, 2018. 8

[27] L. Grammel, M. Tory, and M.-A. Storey. How information visualization novices construct visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943–952, 2010. doi: 10.1109/TVCG.2010.164 2

[28] G. Grolemund and H. Wickham. A cognitive interpretation of data analysis. *International Statistical Review*, 82(2):184–204, 2014. doi: 10.1111/insr.12028 2

[29] T. Grossman and G. Fitzmaurice. Toolclips: an investigation of contextual video assistance for functionality understanding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1515–1524, 2010. doi: 10.1145/1753326.1753552 9

[30] H. Guo, S. R. Gomez, C. Ziemkiewicz, and D. H. Laidlaw. A case study using visualization interaction logs and insight metrics to understand how analysts arrive at insights. *IEEE transactions on visualization and computer graphics*, 22(1):51–60, 2015. doi: 10.1109/TVCG.2015.2467613 2

[31] J. D. Haider, P. Seidler, M. Pohl, N. Kodagoda, R. Adderley, and B. W. Wong. How analysts think: sense-making strategies in the analysis of temporal evolution and criminal network structures and activities. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, pp. 193–197. SAGE Publications Sage CA: Los Angeles, CA, 2017. doi: 10.1177/1541931213601532 2

[32] C. He, L. Micallef, L. He, G. Peddinti, T. Aittokallio, and G. Jacucci. Characterizing the quality of insight by interactions: A case study. *IEEE Transactions on Visualization and Computer Graphics*, 27(8):3410–3424, 2020. doi: 10.1109/TVCG.2020.2977634 2

[33] P. Isenberg, A. Tang, and S. Carpendale. An exploratory study of visual information analysis. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1217–1226, 2008. doi: 10.1145/1357054.1357245 2

[34] W. Javed and N. Elmqvist. Exploring the design space of composite visualization. In *2012 ieee pacific visualization symposium*, pp. 1–8. IEEE, 2012. doi: 10.1109/PacificVis.2012.6183556 9

[35] B. Karer, H. Hagen, and D. J. Lehmann. Insight beyond numbers: the impact of qualitative factors on visual data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1011–1021, 2020. doi: 10.1109/TVCG.2020.3030376 2

[36] H.-K. Kong, Z. Liu, and K. Karahalios. Frames and slants in titles of visualizations on controversial topics. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–12, 2018. doi: 10.1145/3173574.3174012 9

[37] S. Lee, S. Kim, Y. Hung, H. Lam, Y. Kang, and J. S. Yi. How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking. *IEEE Transactions on Visualiza-*

tion and Computer Graphics, 22(1):499–508, 2016. doi: 10.1109/TVCG.2015.2467195 1, 2, 3, 5, 6, 7, 8

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

[39] L. Y.-H. Lo, A. Gupta, K. Shigyo, A. Wu, E. Bertini, and H. Qu. Misinformed by visualization: What do we learn from misinformative visualizations? *Computer Graphics Forum*, 41, 2022. doi: 10.1111/cgf.14559 1

[40] R. Lowe. “reading” scientific diagrams: Characterising components of skilled performance. *Research in Science Education*, 18(1):112–122, 1988. doi: 10.1007/BF02356586 3

[41] W. Luo. User selection strategies of interactive data visualization format. *Journal of Computer Information Systems*, pp. 1–13, 2022. doi: 10.1080/08874417.2021.2023338 2

[42] J. Ma, K.-L. Ma, and J. Frazier. Decoding a complex visualization in a science museum–an empirical study. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):472–481, 2019. doi: 10.1109/TVCG.2019.2934401 2, 8, 9

[43] S. Mahmud, J. Alvina, P. K. Chilana, A. Bunt, and J. McGrenere. Learning through exploration: how children, adults, and older adults interact with a new feature-rich application. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020. doi: 10.1145/3313831.3376414 9

[44] D. Masson, J. Vermeulen, G. Fitzmaurice, and J. Matejka. Supercharging trial-and-error for learning complex software applications. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2022. doi: 10.1145/3491102.3501895 9

[45] E. Mayr, M. Smuc, and H. Risku. Many roads lead to rome: Mapping users’ problem solving strategies. In *Proceedings of the 3rd BELIV’10 Workshop: BEyond time and errors: novel evaLuation methods for Information Visualization*, pp. 8–15, 2010. doi: 10.1145/2110192.2110194 2

[46] C. North. Toward measuring visualization insight. *IEEE Computer Graphics and Applications*, 26(3):6–9, 2006. doi: 10.1109/MCG.2006.70 2

[47] D. G. Novick, O. D. Andrade, and N. Bean. The micro-structure of use of help. In *Proceedings of the 27th ACM International Conference on Design of Communication*, SIGDOC ’09, p. 97–104. Association for Computing Machinery, New York, NY, USA, 2009. doi: 10.1145/1621995.1622014 9

[48] H. L. O’Brien and E. G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology*, 59(6):938–955, 2008. doi: 10.14288/1.0107445 8

[49] L. M. Perkhofer, P. Hofer, C. Walchshofer, T. Plank, and H.-C. Jetter. Interactive visualization of big data in the field of accounting: A survey of current practice and potential barriers for adoption. *Journal of Applied Accounting Research*, 2019. doi: 10.1108/JAAR-10-2017-0114 1

[50] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, vol. 5, pp. 2–4. McLean, VA, USA, 2005. 2

[51] S. Rosenthal and T. R. Chung. The role of expertise on insight generation from visualization sequences. In *2022 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pp. 1–10, 2022. doi: 10.1109/VL/HCC53370.2022.9833097 2

[52] D. M. Russell, G. Furnas, M. Stefik, S. K. Card, and P. Pirolli. Sensemaking. In *CHI ’08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’08, p. 3981–3984. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1358628.1358972 2

[53] B. Saket, C. Scheidegger, and S. Kobourov. Towards understanding enjoyment and flow in information visualization. *arXiv preprint arXiv:1503.00582*, 2015. doi: 10.48550/arXiv.1503.00582 8

[54] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE transactions on visualization and computer graphics*, 11(4):443–456, 2005. doi: 10.1109/TVCG.2005.53 2

[55] P. Shah and E. G. Freedman. Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in cognitive science*, 3(3):560–578, 2011. doi: 10.1111/j.1756-8765.2009.01066.x 6, 8

[56] C. Stoiber, C. Walchshofer, M. Pohl, B. Potzmann, F. Grassinger, H. Stitz, M. Streit, and W. Aigner. Comparative evaluations of visualization onboarding methods. *Visual Informatics*, 6(4):34–50, 2022. doi: 10.1016/j.visinf.2022.07.001 1

[57] N. Sultanum, F. Chevalier, Z. Bylinskii, and Z. Liu. Leveraging text-chart links to support authoring of data-driven articles with vizflow. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021. doi: 10.1145/3411764.3445354 9

[58] M. Tobiasz, P. Isenberg, and S. Carpendale. Lark: Coordinating co-located collaboration with information visualization. *IEEE transactions on visualization and computer graphics*, 15(6):1065–1072, 2009. doi: 10.1109/TVCG.2009.162 9

[59] D. Todorovic. Gestalt principles. *Scholarpedia*, 3(12):5345, 2008. 5

[60] M. Tory, L. Bartram, B. Fiore-Gartland, and A. Crisan. Finding their data voice: Practices and challenges of dashboard users. *IEEE Computer Graphics and Applications*, 2021. doi: 10.1109/mcg.2021.3136545 2

[61] D. F. Vears and L. Gillam. Inductive content analysis: A guide for beginning qualitative researchers. *Focus on Health Professional Education: A Multi-disciplinary Journal*, 23(1):111–127, 2022. doi: 10.11157/fohpe.v23i1.544 3

[62] C. Viau and M. J. McGuffin. Connectedcharts: explicit visualization of relationships between data graphics. In *Computer Graphics Forum*, vol. 31, pp. 1285–1294. Wiley Online Library, 2012. doi: 10.1111/j.1467-8659.2012.03121.x 9

[63] J. Wang, X. Cai, J. Su, Y. Liao, and Y. Wu. What makes a scatterplot hard to comprehend: data size and pattern salience matter. *Journal of Visualization*, pp. 1–17, 2021. doi: 10.1007/s12650-021-00778-8 4, 8

[64] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the working conference on Advanced visual interfaces*, pp. 110–119, 2000. doi: 10.1145/345513.345271 1

[65] J. S. Yi, Y.-a. Kang, J. T. Stasko, and J. A. Jacko. Understanding and characterizing insights: how do people gain insights using information visualization? In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, pp. 1–6, 2008. doi: 10.1145/1377966.1377971 2

[66] E. Zgraggen, R. Zeleznik, and S. M. Drucker. Panoramicdata: Data analysis through pen & touch. *IEEE transactions on visualization and computer graphics*, 20(12):2112–2121, 2014. doi: 10.1109/TVCG.2014.2346293 8