Using Close Reading as a Method for Evaluating Visualizations



Figure 1: Visualization used in the study; colormap for visualization; and detail of visualization. Visualization shows biogeochemistry data in the Gulf of Mexico [11], created using Sculpting Visualizations tool, Artifact-Based Rendering [20]

Annie Bares Stephanie Cullen D. Daniel F. Francesca Zeller Jackson Keefe Samsel

ABSTRACT

Visualization research and practice that incorporates the arts make claims to being more effective in connecting with users on a human level. However, these claims are difficult to measure quantitatively. In this paper, we present a follow-on study to use close reading, a humanities method from literary studies, to evaluate visualizations created using artistic processes [Bares 2020]. Close reading is a method in literary studies that we've previously explored as a method for evaluating visualizations. To use close reading as an evaluation method, we guide participants through a series of steps designed to prompt them to interpret the visualization's formal, informational, and contextual features. Here we elaborate on our motivations for using close reading as a method to evaluate visualizations, and enumerate the procedures we used in the study to evaluate a 2D visualization, including modifications made because of the COVID-19 pandemic. Key findings of this study include that close reading is an effective formative method to elicit information related to interpretation and critique; user subject position; and suspicion or skepticism. Information gained through close reading is valuable in the visualization design and iteration processes, both related to designing features and other formal elements more effectively, as well as in considering larger questions of context and framing.

Keywords: Methodologies, Human-Subjects Qualitative Studies

Index Terms: Visualization

1 INTRODUCTION

As research presented in previous BELIV workshops and otherwise has demonstrated, evaluating visualizations, particularly for their qualities referred to as subjective is a challenge, though this problem has proven a rich area of inquiry

[7]. In this paper, we build upon our recent Visualization Viewpoints article, which explained our motivation for using close reading as a method and gave the results of an initial study [3]. We found that close reading (described in more detail in section 2.3) showed promise as a method to evaluate how visualizations spark human connection, so we decided to take this research a step deeper. Turning to humanities research on environmental visualizations that attends to how "our understandings of visualizations depend on the contextual knowledge, vocabularies, and sociocultural positionings we bring to them," our work defines human connection as a trait that aligns a visualization's informative aims with individuals' informational contexts, previous experiences, and associations [15]. At best, visualizations promote a sustained, internal conversation that incites viewers to curiosity, imagination, and a desire to act, explore, or learn more. Close reading, a foundational method of humanities research from literary studies, is a tool that gives us insight into what degree specific visual, contextual, and associative features of a visualization work together to spark or

inhibit human connection.

This work is part of Sculpting Visualizations, a multidisciplinary research collaborative that includes computer and domain scientists, an artist, and a humanities researcher, in the model of what Donna Cox refers to as "renaissance teams" [9]. Working in the visualization tradition of bringing the arts to visualization, our team has created tools and resources to break down technical barriers of entry and use for artists to play a central role in the visualization conceptualization, design, and iteration process. Our most recently released tool, Artifact-Based Rendering (ABR) enables the incorporation of physical artifacts, including handcrafted sculptures, drawings, and other artisticallygenerated media into visualizations [20]. An ethos of artist- and design-centered thinking informs both what the tools enable, as well as the ways that those tools are conceived, in the case of ABR's interface, which operates like a printmaker's workshop. By allowing the visualization community to incorporate hand-, machine-crafted, and natural objects, this tool expands the visual vocabulary for the benefit of the visualization community and fields whose data they visualize.

Annie Bares, Stephanie Zeller, and Francesca Samsel are at the University
of Texas at Austin. Emails: {abares, stellerzeller}@utexas.edu,
fsamsel@tacc.utexas.edu.

Daniel F. Keefe is at the University of Minnesota. Email: dfk@umn.edu

Cullen D. Jackson is Beth Israel Deaconess Medical Center & Harvard Medical School. Email: cdjackso@bidmc.harvard.edu.

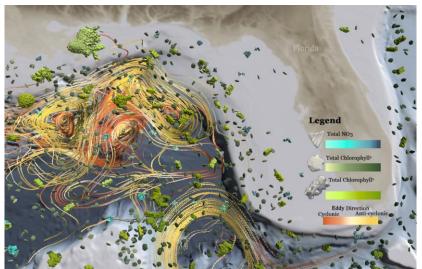


Figure 2: Visualization used in the study of biogeochemistry data in the Gulf of Mexico [11] created using Sculpting Visualizations tool Artifact-Based Rendering [20]

We believe that the visualizations created in alignment with this ethos are more evocative, intuitive, and therefore, effective. However, in the same way that we discovered that visualization can benefit from a richer vocabulary by turning to the arts, we also discovered that the visualization community could benefit from a richer vocabulary to describe and evaluate *why* and *how* visualizations created using artistic and design expertise are more effective. To develop and study this vocabulary, we turned to the humanities. As a discipline, the humanities have for centuries developed methodologies for describing, analyzing, and interpreting works of art and culture. The humanities provide unique theoretical and practical insights into how and why humans create, respond to, and use artistic and cultural texts to examine and mediate complex societal problems.

In this paper, we present a follow-on study using the close reading process with three different audiences across areas of expertise in visualization and domain sciences, as well as generalists. Due to the COVID-19 pandemic, we had to make significant changes to the evaluation method by conducting these evaluations at a distance via video and audio calls, a shift which we describe in sections 3.2 and 4.1. We also conducted more rigorous quantitative data analysis than in our initial studies and present new results here. Conducting close readings on an individual basis helped us to refine the value that we see in using this method for formative evaluation. Close reading produces rich, complex data that includes specific ideas from users to guide future iterations of visualization design and illustrates how crucial it is to understand how individual user experience and background informs how that user interprets the visualization. While our previous study demonstrated the importance of context to visualization, without knowing much about participants' backgrounds, we could only hypothesize about the importance of individual user background and subject position, to their interpretation of the visualization. User subject position describes how various aspects of a user's identity situate and inform their perspectives and interpretations of visualizations. By conducting this study with scientists, visualization experts, and generalists, we are able to better test our previous hypothesis that, in addition to informational context, user context is central to visualization interpretation. As our results demonstrate, close reading provides an evaluation method that recognizes the relational qualities of data and data visualization that humanities and

information science researchers have pointed to in proclaiming that "data never stands alone" and that "data is always already a cultural product" [16] [2].

2 BACKGROUND

2.1 Art, Humanities, and Visualization

Our work partakes in the rich tradition of interdisciplinary collaboration in visualization. Artists are experts in expressing complex ideas visually, and as such, the visualization community has recognized the value of creating workflows, tools, and resources to incorporate artistic expertise. As we have noted previously, the relationship between the humanities and visualization has primarily been characterized by the digital humanities' tendency to apply methods from computing to humanities content. Distant reading, a foundational digital humanities practice, seeks to draw generalizations about literature by conducting computational analyses on large corpuses of texts [19]. In keeping with the larger ethos of Sculpting Visualizations, we invert this principle by applying a humanities methodology, close reading, to computationally-generated data and data visualizations. Our work draws on a wider recognition of the crucial value of the humanities' interpretive lens to the positivism of the sciences in approaching societal issues like climate change [17]. Humanities, information studies, and Science, Technology, and Society scholars have brought critical methodologies, including close reading, aesthetic analysis, and historical contextualization to studies of Big Data, visualization, and the production of scientific knowledge [2] [10] [15] [16] [27]. Humanities scholarship has begun to treat visualizations like cultural objects, in recognition of the ways in which visualizations "yield entangled epistemologies dependent on culturally contingent responses to color, temporality, and genre conventions" [16]. Close reading offers a method for unpacking these entanglements and understanding how specific formal features combine with content and cultural context to produce meaning for individual users.

2.2 Comparison to other Approaches

In our previous study, we found that close reading provided distinct results in evaluating visualizations qualitatively based on aesthetic qualities like "beauty" [23] and the use of artistic critique as an

evaluation method [18]. Close reading, particularly as conducted in this study through verbal, recorded interviews differs from these

Table 1: CATEGORIES USED FOR CODING RESULTS OF CLOSE READING STUDY

	Category	Subcategory	Example of Results
Making Meaning	Critique Participants pass judgement on visualization or features of visualization	Positive Critique: Participants express favorable judgement	"The visualization is very engaging and interesting and it makes you want to spend more time with it and look at it more."
		Negative Critique: Participants express unfavorable judgement	"There's far too much information."
		Iterative Critique: Participants make suggestions for improvement	"When [the glyphs] get dark, it gets hard [to distinguish]You could cross hatch one and put a texture in addition to the hue."
	Exploratory Participants used interpretive process to develop understanding of the visualization	Gaining Understanding: Participants observed in the act of coming to a better understanding of the visualization through interpretation	"It seems clear that the bright blue colors indicate where there's high levels of NO3 that are conducive to producing the the seaweed. The sargassum appear to be in quiet central areas where you don't have very much eddy flow."
		Talking about Understanding: Participants explicitly describe how features of the visualization or its context contributed to their understanding of it.	"Lwent in looking at it as a whole pictureand then kind of broke it into parts. So the cyclonic/anti-cyclonic is where it kind of helps me looking at it like, "Oh, it's off the coast of Florida." So maybe we're looking at a hurricane"
	Meta-Visualization Commentary Participants referenced how visualization would or could be read based on a variety of factors outside of the visualization itself		"I tend to work with data that are uniform grids, so I'm either looking at color, shading, or contour lines. I'm not used to looking at discreet elements."
		Referencing Audience for Visualization: Participants explicitly reference how potential audience for visualization relates to interpretation	"If I were a member of the public and I saw numbers on it, it might be kind of a turnoff This is a little too detailed for what I need to know."
		Referencing Purpose of Visualization: Participants explicitly reference how the potential purpose of the visualization relates to interpretation	"What's the overall message supposed to be?"
		Referencing Subject Position: Participants referenced their own subject position during the interpretive process, in most cases, citing it as a reason for an interpretive point that they're making	"I think I went to the map because as a climate scientist and ocean modeler I'm always going to try to orient myself, "like where am I?" "
Descriptive	Features of Visualization Evoke Need for Different Vocabulary Participants explicitly note that visual features don't fit current visaulization vocabulary or they result to metaphors to describe visual features		"The glyphs are all different sizes and shapes, but I could have a limited vocabulary. That's why I don't know exactly how can I even express this. In my visualization field, I would say it's a set of polygonsWhat could be a synonym for that in layman's terms? Maybe blobs?"
	Straightforward Description Participants provide description of the visualization as a part of interpretation		"I'm seeing a visualization of the Gulf of Mexico, Florida is very visible. And I am seeing sort of streams of what I assume to be ocean flow."

methods in focusing more on eliciting subjective, open-ended information about the visualization than objective or discrete criteria that measures a visualization against standards of aesthetics or relies on an expert perspective, as with the use of critique as an evaluation method. Instead, we use close reading as a method that elicits more holistic analysis that complex, multi-faceted visualizations demand, in the same way that close reading is used in the humanities to elicit user interpretation of complex, multifaceted texts. Like in the humanities, we do not use close reading as a tool to determine whether a visualization is "good" or "bad," but instead in service of a more robust mode of formative evaluation of a visualization's features that need human evaluation. However, unlike in the humanities, where close reading is a mode of analysis that a scholar typically uses in service of larger arguments, not in service of feedback on the text subject to close reading, we use close reading as a rigorous method of data collection that provides us with information about how users interpret key visual features.

In this particular study related to a specific visualization finds precedent in research conducted by Hogan et. al. and Nowak et. al. Both close reading and the phenomenological interview center the subjective experiences of users and apply non-traditional qualitative methods to visualization evaluation [14] [25]. However, our application of close reading differs from a phenomenological approach in several important ways. First, the goal of using close reading as an evaluation method is primarily to break down the interpretive process that users go through when engaging with a visualization. Essentially, close reading helps us to understand how a visualization's content (its subject matter) combine with its form (how it presents its subject matter through its aesthetic features) along with contextual information and individual users' particular associations due to their prior knowledge and subject positions. Rather than walking participants through the act of evoking a previous experience with the visualization, as a phenomenological approach does, close reading allows us a window into user interpretation of a visualization by breaking down close reading into a series of steps that act as prompts that the user goes through, worksheet style, in writing or as an interview. And, while we include a question about how the visualization and the process makes the user feel as a part of the reflection prompt, the act of interpretation is at the center of close reading.

2.3 Close Reading

Our work has previously defined close reading as a foundational humanities method that relies on close attention to a text's content (what a text is "saying") and form (how it's being "said") [3] [6]. As the development of close reading demonstrates, texts now include a wide variety of cultural and discursive artifacts and are not limited to literary texts like poems and novels [13]. Like any other disciplinary method, there are different outcomes when practitioners with different levels of expertise employ close reading, though one need not be an expert in literary studies in order to conduct or to follow a close reading. In visualization, "active reading," has been explored as a mode for enhancing user understanding of a visualization through annotation [29]; close reading has been considered as a way for video game designers to test the effectiveness of their games [5].

Our work builds on these uses of reading, but presents close reading as a defined method that can be used specifically for the evaluation of visualizations. Close reading is a rigorous, foundational humanities method that we gave structure to (described in 3.1). We drew on the expertise of our team's humanities researcher to develop a rigorous evaluation method that consisted of a set of scaffolded questions. These questions prompt interpretation of the visualization. We wrote the steps down in enough detail so that a scientist or any other researcher who is not familiar with close reading could conduct the study, which could be conducted on any visualization, artistically-crafted or not.

3 THE STUDY

3.1 Method

The goal of our study was to determine if close reading is an effective method for evaluating visualizations with users across a range of expertise. For our study, the basic operations of close reading are a series of scaffolded steps that begin with (1) a summary of the visualization, (2) observations about features of the visualization, (3) analysis of features by hierarchizing features based on significance (4) providing contextual information (5) synthesizing the previous steps and context (6) reflecting upon the visualization and close reading process. These steps were adapted from the Close Reading Interpretive Toolkit (CRIT), a method created by English professors at the University of Texas at Austin

to teach students how to close read literary texts [8]. We provide a complete explanation of these steps in prior work [3] and provide a worksheet template that other researchers or teachers can adapt for evaluating or teaching visualizations [28].

Key adaptations that we made to the CRIT steps in this study include adding a follow-up question to step 4, after providing users with a paragraph of context asking the user what questions remain after the reading contextual information. In step 5, we added language to the prompt that asked users to tell the story of the visualization to clarify what we were asking them to do in synthesizing analysis of formal features and context. In order to maintain consistency with our previous studies, we provided users with a similar visualization as was provided to previous users, of biogeochemistry data in the Gulf of Mexico from the E3SM coupled climate model for the purpose of exploring scientific data, created with Artifact-Based Rendering [11] [20]. In this study, the visualization included a legend (Figure 2).

3.2 Procedures

While the steps of close reading that users were asked to respond to were largely unchanged from the previous studies, the procedures of this study were different from those of previous studies due in part, to COVID-19 and in part, to enable us to interview non-local researchers. In the two studies conducted previously, members of the research team conducted the evaluation with groups of approximately twenty undergraduate students at once, in-person, leading the classes through close reading of a printed visualization with a hard copy of a worksheet that they filled out independently. Originally, due to the results of those studies, we planned to adapt close reading for 3D visualization. However due to COVID-19, we were unable to conduct the study on groups of students in-person. And, as few people have access to VR headsets in home settings, we also were unable to conduct studies of 3D visualizations.

Instead, we chose to pursue another research path of interest: exploring the suitability of using close reading as a method to evaluate visualizations with users who had varying degrees of knowledge in visualization and domain sciences. We conducted the close readings individually, rather than in a group setting, over video or calls, which were recorded. The most significant change to the procedure, described below, from our previous studies was that most participants responded to the prompts verbally, rather than in writing. On average, the study took 31 minutes to conduct per participant.

In step (1) the close reading facilitator, Bares, explained the purpose of the evaluation and information about IRB protocols to the interview participant. In step (2) she began recording audio and/or video using video call application features. In step (3) she asked the participant to open an image file of a 2D visualization. In step (4) she used the chat feature of the video call application to send the user the prompt for Step 1 of the adapted CRIT method. She also explained the prompt aloud, as a teacher might. She explained that the question could be answered verbally or in writing by responding in the chat feature. The user was able to ask questions for clarification. In step (5) users responded to the prompt, mostly verbally (11 of 13 users), though 2 users did type out their answers and respond in the chat. We allowed users to determine on their own whether they would write or speak to allow them to determine which mode of expression best suited them, knowing that both writing and speaking mediate thought. We discuss differences and similarities between writing and speaking in section 5.2. Users were able to continue to look at and refer back to the visualization as they answered the prompt. Steps (4) and (5) were repeated 5 additional steps of the adapted CRIT method.

TEXT ANALYSIS: VISUALIZATION CLOSE READING RESULTS

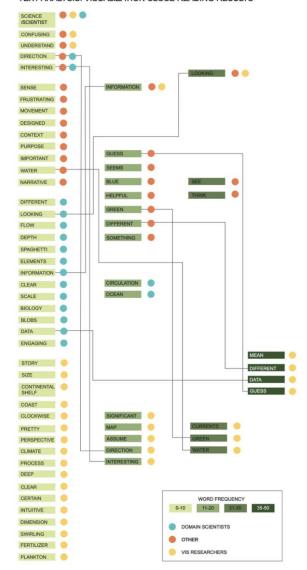


Figure 3: To map word frequency by user group, all participant answers were run through text analysis software. Filler words and words that appeared on the legend or those variables given in the contextual paragraph were removed from this view. In this figure, lines indicate words that were used across participant groups, but at different frequencies.

3.3 Materials & Setting

The study was conducted over Zoom (San Jose, CA) as the video call application, by one member of the research team, Bares. Two participants requested that they conduct the study with audio only. Users were provided with a digital file of an image of the visualization via email. This was a significant change from previous close reading studies that were conducted in-person with

groups of 20+ students at once. The image of the visualization of the biogeochemistry data was very similar to the image that we used in previous studies, but this time, we added a legend to it. The primary goal of the scientific visualization used in this study is to allow scientists to explore their complex, multivariate data intuitively by aligning its visual features with principles of art and design meant to hierarchize information in an intuitive way.

3.4 Participant Profile

In previous versions of the study, in which participants were relatively anonymous, all we knew about them was that they were undergraduate students. However, in this study, a single member of the research team conducted individual interviews with 13 participants. Four of the participants were domain scientists, six of the participants were visualization researchers, and three of the participants didn't fit into those categories (middle school science teacher, science writer, university history professor who conducts digital humanities research), referred to in the grouping named "other" throughout the paper.

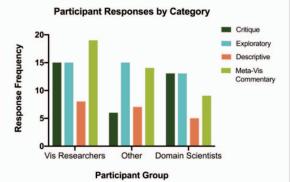
3.5 Data Analysis Method

After the close reading sessions, Bares used Temi (San Francisco, CA), an AI transcription tool to transcribe participants' answers. After ensuring that the transcript matched the recording, Bares then conducted qualitative data analysis (QDA) to analyze participant responses [31]. QDA usually is not conducted on close readings as they are not typically aggregated and studied. In the humanities, close reading is the analysis method, rather than a method of data collection, as in our study. We conducted QDA on the close reading responses to further understand how user subject position influenced how different participants interpreted the same visualization using close reading.

Rather than searching for predetermined keywords or phrases, Bares read through all of the responses to develop coding categories from the data based on common types of responses and patterns. The results of this qualitative analysis appeared to show differences in some categories between how each of the three participant groups (domain scientists, visualization researchers, and others) responded during close reading (see Table 1 for example responses). To more quantitatively investigate the differences seen in the qualitative data analysis between the participant types and the four main coded response categories: critique, exploratory, descriptive, and meta-visualization commentary, we calculated the frequency of each of the response types for each individual within the three groups, which produced a 3x4 contingency table (Figure 4 shows a graphical representation of the frequencies). Given these nominal data, we used the Chisquare statistic [32] on the full table to test for any significant relationships between the three participant groups and the four response categories.

We also used text analysis software to determine frequently used words (Figure 3) and three-word phrases. The most frequently used phrases by visualization researchers were: "I don't know;" "something like that;" "kinds of chlorophyll;" "the continental shelf;" "the original data." The most frequently used phrases by domain scientists were: "Gulf of Mexico;" "In the ocean;" "too much information;" "that looks like;" "engaging and interesting." The only frequently used phrase by the Other group was "I don't know." A caveat to using frequency to measure words, phrases, and responses among aggregated groupings is that there are different numbers of participants in each of the three groups and each participant spoke for different amounts of time.

Figure 4: Colors on bars correspond with categories of responses



and indicate response frequency (vertical axis) based on participant group (horizontal axis).

4 RESULTS

4.1 Description of Results

The raw data that comprise the answers to the prompts presented in the experiment give information about user responses in five broad categories: information about how viewers summarize the visualization presented without context (step 1: summarize); recognizable features of the visualization and designation of which of these features are most significant to each participant (steps 2 and 3: observe and analyze); information about what contextual information participants felt like they needed before being given a paragraph of text about the visualization and what they still felt was missing after receiving the text (step 4); information about what story the visualization tells each user or how they synthesize the visualization and its context (step 5); reflections on the visualization or the process (step 6). As this paper is focused on the evaluation method, rather than simply presenting their answers as results, we analyzed the raw data to present what other kind information close reading as a method produces. We developed five categories with sub-categories, presented in Table 1, for coding data. We divided them into overriding groupings that refer to as "meaning-making" and "descriptive." These groups distinguish between responses that involved simply describing what participants saw in the visualization in terms of features and those responses that indicated interpretation beyond discerning what the visualization contains. Upon further reflection, we noted that the categories that we created roughly aligned with modes of inquiry that visualizations at their best, prompt: curiosity, imagination, and a desire to act, explore, or learn more.

The results are summarized in Figure 4. The Chi-square yielded a test statistic equal to 5.012 with 6 degrees of freedom and a corresponding p-value of 0.5423, so we failed to reject the null hypothesis and did not find any statistically significant quantitative differences between the response groups and the response type categories.

4.2 Summary of Results

Our results demonstrated that close reading effectively delivers useful and interesting information from careful, close readings of the raw data and the coding process. We found close reading to match the exploratory goals of the visualization by prompting equally exploratory, creative information by way of interpretation by participants. From their answers to the prompts, we learned which features of the visualization seemed most important, where confusion emerged, what additional context they needed, and the importance of understanding the purpose of the visualization. Another shared finding of the evaluation methods was the

confusion that showing 3D visualizations as 2D images provokes even for experienced users. This remains a problem for visualization and visualization evaluation as a whole, as technologies to create more sophisticated visualizations outpace widespread adoption or ease of use of these technologies. However, key findings that emerged or were refined in this study that we discuss here include instances where interpretation turned into critique; instances of meta-visualization commentary (questions about purpose of the visualization, metadata, audience for it, and reference to the participant's subject position); how conventions of visualization train viewers; suspicious interpretations of visualization; and how user background influences approach to visualization.

Despite the fact that our quantitative analysis did not conclusively support our hypothesis, we suspect that with a larger sample size, if similar results were returned, we would begin to see more conclusive evidence of a relationship between user background and the kinds of interpretive statements that they made. Our sample size of 13 participants, though sizable for in-depth qualitative studies in the visualization community is small for statistical analysis. Future work will expand our participants, scientists and professionals from the field, from 13 to 25 to enable more robust statistics.

5 DISCUSSION

5.1 Effects of Writing versus Speaking

As a result of changing our procedure from administering close reading to an entire class as a written worksheet to conducting individual interviews, 11 of the 13 participants chose to respond verbally. As in the previous studies, it is clear that many aspects of the visualization, due to their artistically generated nature, are unfamiliar to visualization users. While some users did refer to glyphs in more evocative, metaphorical ways (Figure 5) we found that participants in this study were less apt to come up with metaphorical descriptors. We hypothesize that conducting close reading as an interview, rather than as a written exercise encourages participants to assume that the interviewer knows which aspects of the visualization she is referring to, rather than writing to an imagined reader. We also hypothesize that writing carries with it greater encouragement to interpretation, a plus for close reading. One of the two users who did choose to write his answers to the prompts noted that he did so because "I need a couple of minutes to kind of look at things. It's just the nature of how I look at things and edit. So, I'll look at it and I'll start writing a sentence. And I think, 'Hang on!' I'll look at it a bit more, knowing I'll change that sentence when I go back to it. And then I feel like I've kind of addressed that point." However, speaking answers does capture more immediacy of reactions and allows for less mediated commentary, suggesting that using these methods in tandem, as teachers often do when leading a class in close readings (asking users to write down their answers and then explain them, for example) could be useful.



Figure 5: A detail of visualization used in study (Figure 2) that shows glyphs and streamlines described by participants as "Baby Ruth bars" and "spaghetti."

5.2 Interpretation and Critique

While close reading is primarily considered a mode of interpretation, study participants also partook in moments of critique, passing judgement on the visualization as a whole and specific features. While some moments of critique are very broad ("I see a total mess!" or "'this looks like art!'"), others were very specific ("the lighting is a bit strange"), and some were framed specifically as iterative suggestions ("you could use more distinct color contrast"). Moments of critique derived from users' comments illustrate the strength of close reading in encouraging open-ended, exploratory feedback that might not be gleaned from surveys or task-based assessments. This feedback can help motivate visualization design iteration process.

5.3 Conventions of Visualization

A key insight that close reading provided in this study was the way that conventions of visualization influence user experience with them. We found some evidence that this is the case in the previous studies [3], particularly related to the content of visualizations (for example, assuming that a visualization about science in the Gulf of Mexico would be about a hurricane), and it was confirmed in this study. However, more significantly, in this study we found that conventions of visualization exist not only in content, but in form. That is, users become accustomed to certain standard aesthetic features of visualizations and are thus, trained to read them in specific ways. Close reading as a technique in literary studies draws out moments where a text follows formal conventions (for example, a poem following a standard sonnet rhyme scheme) and those where it departs from them, sparking interest when those moments are particularly expressive, thought-provoking, or complementary to its content.

As scholars of visualization have noted, the same can be said of user experiences with formal conventions of visualizations, which come loaded with aesthetic, epistemological, and affectual conventions [16] [27]. One obvious example of this is users who noted wanting to see a legend and a user who noted in the analysis step, that he determines which features of the visualization are most important based upon the order that they come in on the legend. Close reading proves to be an especially useful method for determining when these conventions are broken and the effects thereof. One participant, for example, noted the confusion that came because the glyphs did not align with "canonical" or "textbook" notions of what a representation of a diatom should look like. Whereas, another participant shared that his wife saw the visualization and said, "She was very drawn to the visualization and said, "Wow you've got something cool on your screen now." As

opposed to what a visualization normally looks like. And she has a very good aesthetic sense, she said, "Oh that looks like art." Likewise, another visualization researcher used interpretation as a jumping off point for reflecting on this very issue: "These glyphs are definitely a newer approach. They are done with the technique that's not too common. It is much more three dimensional here than a lot of what I've seen...Part of it is creating like a language, of the visualization...(Figure 6) the more people become familiar with a certain technique, maybe the more they'll kind of have an intuition for what it means. At the same time, some of the best visualizations are fresh and sort of surprising. And so I think, you know, that's also a tradeoff basically like how to make it fresh and surprising versus how to make it intuitive by requiring a whole lot of explanation or investment by the viewer." Samsel, the visualization designer on the team, identified the three comments above as particularly useful for the design and iteration process in expanding visual vocabulary and understanding.



Figure 6: A detail of visualization used in study (Figure 2) that shows glyphs and streamlines that participants described as evocative of the need for "a new language" for visualization.

5.4 Suspicious Reading

One result in this study that was markedly different from others is that through interpretation, some users expressed suspicion towards the visualization. In literary studies, close reading has become associated with "the hermeneutics of suspicion," with the idea that attending to a text very closely leads to readings of it that are skeptical of the text's claims or aim to uncover something untoward or ideological about it [12]. Three users expressed outright suspicion of the visualization or its data. One experienced visualization scientist pointed to the fact that due to the distribution of information in the visualization he " is more skeptical of a visualization when I don't know the original data that it came from. So I just want to correlate with the original data, because I personally do data sampling. So this is interesting to me because if I show you a sample of the data, then I might be misleading you by saying, this is the only region where nitrate is found. I'm not saying that that's true in this case, in general, [this is] why I'm pausing a couple of seconds before I say anything, because when I do random sampling of a region, then I can be randomly showing you some regions of the data." He suggested that he would like to have information about the metadata to better interpret the visualization.

Another visualization researcher made an off-hand joke during the context step about the fact that because of the visualization's renewable energy-related subject matter, he expected to see a note that it is "sponsored by ExxonMobil." To be clear, neither the science in the visualization nor the visualization research is sponsored by ExxonMobil; see acknowledgements. However, that issues of funding and corporate influence are top of mind for an experienced visualization researcher points to the fact that interpretations of visualizations are inseparable from wider

political and ethical contexts of science, an argument that scholars across disciplines, from climate communications to literary studies have made [4] [16] [26].

And the other participant, who has conducted extensive research in visualization for the humanities, expressed that based on his previous experience with scientific visualization related to his brother, who is a chemical engineer, he has observed that, "It's almost like [in some scientific visualizations] there's a deliberate design to be opaque to the general viewer or to the general reader...[the visualization used in the study] is a very complex topic, obviously made even more complex by quite a difficult to grasp diagram." Here the participant's previous experiences with scientific visualization inform his frustration with the visualization at hand as being difficult and opaque in his interpretation of it.

These moments of suspicion or skepticism may seem unhelpful or easily dismissed given the nature of the visualization, which is not meant to be persuasive or in support of a particular policy or position. However, scholarship on visualization in humanities and Science and Technology Studies reminds us that it is a form loaded with larger associations about Big Data [15] [16]. Visualizations can remind users of possibilities for manipulating science in service of larger political aims, a problem that speaks to a lack of clarity or trust on the part of the public about the scientific process [26].

Similar to the comments about glyph design, Samsel found comments where users expressed skepticism about the data to be particularly useful for informing future visualization design research. Close reading uncovered these visualization issues more clearly than other evaluation methods, like perceptual studies that do not typically take into account context and user background. Close reading does not just ask the user to recite what is on the page, but instead highlights how visualizations are relational sites of meaning-making. As such, close reading as an evaluation tool can help makers of visualizations understand how viewer context like the need to know the metadata or ethical questions related to funding sources, can be just as important as formal design features, like glyph color or shape.

5.5 User Subject Position

One of the clearest findings of this study, across all of our findings, was the extent to which individual participant subject position influenced their interpretation of the visualization. We do not see the fact that close reading elicits information that arises from an individual's specific background as a weakness of the method. Instead, we consider it a strength in that it allows us to study how user background influences their interpretation of visualizations, a factor which is often purposefully excluded from other evaluation methods, despite the fact that this separation of user background and experience with a visualization does not reflect real world conditions of interpretation. We did attempt to protect against confounding effects of different experiences with close reading as a method in the evaluation by excluding individuals with extensive training in close reading, except for the participant who is a history professor.

The idea that an individual's background, associations, and experiences influence their interpretation of data and information is not a new one and has been studied across fields including science communication, science and technology studies, and museum studies [4] [24]. However, this study points to the effectiveness of close reading as a method to explicitly study the effects of user subject position in interpreting visualizations, as a majority (9) of participants made explicit reference to their subject position, when making interpretative claims. In our previous studies, we only had a vague sense of our participants' backgrounds as undergraduate students. However, in this study we sourced

participants from previous collaborator networks. Some were intimately familiar with the research program, others were not. Some were well versed in reading and/or making scientific visualizations, while others were not. However, in each case, we knew more about each participants' individual professional background and training. In addition to the moments when participants explicitly referred to their subject position or how people in other subject positions might interpret the visualization, we were also able to think about their response to the visualization in reference to that knowledge about their backgrounds. This knowledge of participants proves essential to how we think about tailoring visualizations to users and audiences.

In some cases, users referred to their subject position as a way to signal their understanding or confusion about features of the visualization, as in the case of a climate scientist: "With my limited biology knowledge, the glyphs don't really ring a bell for any specific species. We think of diatoms and couple of the forums and things in these regions and the glyphs tend to look a lot like the canonical view of those and these look kind of like arbitrary glyphs. So it's hard for me to tell what the differences are." In two other cases, users pointed to their backgrounds as a position for expertise to give suggestions for iterations ("Because I have a background as a graphic designer, I probably suggest some maybe stronger colors like a red."). Another participant noted how his background influences his approach to understanding the visualization and organizes what its key features are for him: "I went to the map because as basically a climate scientist and ocean modeler I'm always going to try to orient myself, like where am I?"

In a particularly striking example, one user was able to interpret key aspects of the visualization with no contextual information much easier than the rest of the participants because of his background. However, the piece of his background that informed this knowledge was not only or primarily his knowledge of visualization or science, but rather because he "spent a few years back some vacation time at a beach house on the panhandle of Florida. So I know they have algae blooms, or a red tide problem. So I immediately put nitrates together with chlorophyll and ocean currents." And another participant, when asked in Step 5 of close reading to put together context and the information in the visualization to tell its story noted that, "The context of the graphic is really everything because whether it's a bad news story or kind of a good news story or a kind of a neutral story. In this case, I'd say it's a good news story because it seems to be harnessing what is ecological damage in the form of nitrate runoff and unconstrained growth into something that like biofuels, which is kind of good news and certainly worth looking at. This suggests a new way of doing things. That's more using the powers of biology, rather than just sinking lots of oil drills into the sea. The old way of doing things." Here, the participant's prior knowledge of resource extraction in the Gulf of Mexico, as well as his stance on it, influenced his interpretation of the visualization, the science, and its potential implications.

Close reading points us to the realization that despite the intentions of visualization experts and teams, data visualizations are not simply facts relayed but rely on relational exchanges between user and creator associations, backgrounds, understandings of conventions of visualization, and relationship to the data being visualized. Creators and users, due to their subject positions, knowledge, and socially constructed identities, always bring other information and associations to their interpretations of these images to create meaning in ways that may exceed the "correct" interpretation of the data or visualization, but that are nonetheless influential on their understanding of it.

6 FUTURE WORK & CONCLUSION

This study opens up several future research possibilities related to close reading. The first includes testing the efficacy of using close reading on different genres of visualization. Other possibilities include using close reading to evaluate visualizations with different content and contexts of visualizations: for example, an artistically generated 3D visualization with analytical graphs, poetry, or photographs as context. Future research will specifically focus on using close reading as a method for evaluating artistic and contextual elements of visualizations to better understand how richer contexts and visual vocabularies make visualizations more effective in connecting with users on a human level. As a point of comparison in evaluating close reading as a method for visualization evaluation, we may also compare information that is elicited from asking a participant open-ended interview questions to information elicited from engaging that user in the close reading process.

As noted, we encountered issues in evaluating 3D visualizations as 2D images. As such, we would like to conduct close reading as an evaluation method using video representations of a visualization in 3D. And when possible, we plan to conduct a study testing how close reading can be used as an evaluation method for 3D visualizations experienced in a 3D environment. As we found distinctions between writing and speaking aloud when conducting close reading, as well as conducting it in a group versus individual setting, testing these differences more explicitly also provides areas for future work.

In conclusion, this study demonstrated that close reading as a method is easily adaptable to different conditions, though that these different conditions do inflect results. Close reading is useful because it allows us to evaluate more discrete aspects of a visualization. These include testing baseline understanding of a visualization through user summary, description, and synthesis. Close reading captures this information about the whole of the visualization in the synthesis step of close reading, as well as specific features in results from the observations and analysis steps.

However, because close reading activates users' interpretive responses to a visualization, it also simultaneously allows for more open-ended, perhaps, unexpected evaluation results based on user experience with the visualization. As noted in the discussion, this study demonstrated how close reading is particularly useful for understanding what contexts, associations, and information they bring to the visualization based on their background. This is useful information for further iteration at all steps of the visualization creation and design process and can inform larger questions about the nature of visualization in relationship to data, artistic approaches to visualization, and primary motivations for visualization as a field. Essentially, close reading is a method that allows us to get at the relational nature of visualizations; that is, they are not only a one-way source of information transfer, but instead, a visual site that individuals come to with their own associations, and preferences that inform their interpretation of visualizations.

ACKNOWLEDGEMENTS

This work was supported in part by the National Science Foundation (IIS-1704604 IIS1704904). Brain microstructure applications were supported in part by Dr. Christophe Lenglet at the University of Minnesota and by the National Institutes of Health (P41 EB015894, P30 NS076408). MPAS-O simulations were conducted by Mathew E. Maltrud and Riley X. Brady as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy (DOE), Office of

Science, BER with analyses conducted by Phillip Wolfram, MEM, and RXB under ARPA-E Funding Opportunity No. DE-FOA-0001726, MARINER Award 17/CJ000/09/01, Pacific Northwest National Laboratory, prime recipient.

The CRIT method reproduced on the guided close reading worksheet was developed in the Department of English at The University of Texas at Austin by Professors Phillip Barrish, Evan Carton, Coleman Hutchison, and Frank Whigham, and Ph.D. students Sydney Bufkin, Jessica Goudeau, and Jennifer Sapio. CRIT is a product of a Course Transformation Grant generously funded by the Office of the Executive Vice President and Provost. CRIT is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

REFERENCES

- D. Acevedo, C. Jackson, F. Drury, and D. Laidlaw, "Using visual design experts in critique-based evaluation of 2D vector visualization methods," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 4, pp. 877– 884, Jul./Aug. 2008.
- [2] A. Acker, "Data cultures, culture as data special issue of cultural analytics," *Cultural Analytics*, pp. 1-8, 2019.
- [3] A. Bares, "Close reading for visualization evaluation," *IEEE Computer Graphics and Applications*, vol. 40, no. 04, pp. 84-95, Jul/Aug. 2020.
- [4] J. Besley, "Scientists' views about communication training," *Journal of Research in Science Teaching*, vol. 52, no. 2, pp. 199-220, 2015.
- [5] J. Bizzocchi and T. Tanenbaum, "Well read: Applying close reading techniques to gameplay experiences," 2011.
- [6] B. Brummett, Techniques of Close Reading. Newbury Park, CA, USA: SAGE, 2019.
- [7] S. Carpendale, "Evaluating information visualizations," in A. Kerren, J.T. Stasko, J.D. Fekete, C. North, Eds., *Information Visualization*. *Lecture Notes in Computer Science*, vol. 4950. Berlin, Germany: Springer, 2008.
- [8] "Close reading interpretive tool," University of Texas at Austin Department of English, [Online]. Available: https://laits.utexas.edu/crit/home
- [9] D. Cox, "Using the supercomputer to visualize higher dimensions: An artist's contribution to scientific visualization," *Leonardo*, vol. 21, no. 3, pp. 233–242, 1988.
- [10] J. Drucker, Graphesis: Visual Forms of Knowledge Production. Cambridge, MA, USA: Harvard Univ. Press, 2014.
- [11] S. Dutta, R. Brady, M. Maltrud, M. Einar, P. Wolfram, and R. Bujack, "Leveraging lagrangian analysis for discriminating nutrient origins," in *Proc. Workshop Vis. Environ. Sci.*, 2019, pp. 1–8.
- [12]R. Felski, The Limits of Critique, Chicago, IL: University of Chicago Press, 2017.
- [13]J. Gallop, "The historicization of literary studies and the fate of close reading," *Profession*, pp. 181–186, 2007.
- [14]T. Hogan, "The elicitation interview technique: capturing people's experiences of data representations," *IEE Tras. Vis. Compt. Graphics*, vol. 22, no. 12, pp. 2579-2593, Dec. 2016.
- [15]H. Houser, "Climate visualizations as cultural objects," in S. Siperstein, S. Hall, S. LeMenager, Eds. *Teaching Climate Change in the Humanities*. Evanston, IL, USA: Routledge, 2016.
- [16]H. Houser, Infowhelm: Environmental Arts and Literature in an Age of Data. New York, NY: Columbia University Press, 2020.
- [17]M. Hulme, "Meet the humanities," Nature Climate Change, vol. 1, pp. 177–179, 2011.
- [18]B. Jackson, et al., "Toward mixed method evaluations of scientific visualizations and design process as an evaluation tool," in *Proc.* BELIV Workshop: Beyond Time Errors-Novel Eval. Methods Vis., 2012, pp. 1–6.
- [19]S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann, "On close and distant reading in digital humanities: A survey and future challenges," EuroVis (STARs), 2015.
- [20]S. Johnson, et al., "Artifact-based rendering: Harnessing natural and traditional visual media for more expressive and engaging 3D visualizations," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 1, pp. 492–502, Jan.2020.

- [21]D. Keefe, et al., "Scientific sketching for collaborative VR visualization design," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 4, pp. 835– 847, Jul./Aug. 2008.
- [22] R. Kosara, F. Drury, L. Holmquist, and D. Laidlaw, "Visualization criticism," *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 3, pp. 13– 15, May/Jun.2008.
- [23] P. Kozik, L.G. Tateosian, C. G. Healey, and J. T. Enns, "Impressionism-Inspired data visualizations are both functional and liked," *Psychol. Aesthetics, Creativity, Arts*, vol. 13, no. 3, pp. 266– 276, 2019.
- [24] L. Neuburger. "Augmented reality: providing a different dimension for museum visitors," in T. Jung, M.C. tom Dieck, Eds. Augmented Reality and Virtual Reality Empowering Human, Place and Business. Springer International, 2020.
- [25] S. Nowak, "A micro-phenomenological lens for evaluating narrative visualization," *IEEE Evaluation and Beyond-Methodological* Approaches for Visualization (BELIV), pp. 11-18, Oct. 2018.
- [26] N. Oreskes, Why Trust Science, Princeton, NJ: Princeton University Press, 2019.
- [27] B. Schneider, "Burning worlds of cartography: a critical approach to climate cosmograms of the Anthropocene," *Geogr. Environ.*, vol. 3, no. 2, 2016, Art. no. e00027.
- [28] Sculpting Visualization Collaborative, "Close reading as a method for evaluating visualizations," https://www.sculpting-vis.org/index.php/evaluation/.
- [29] J. Walny, et al., "Active reading of visualizations," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 770–780, Jan.2018.
- [30]Y. Wang, et al., "An emotional response to the value of visualization," IEEE Comput. Graph. Appl., vol. 39, no. 5, pp. 8–17, Sep./Oct.2019.
- [31]S. B. Merriam et. al., "Qualitative data analysis," in S.B. Merriam and E. Tisdell, Eds., Qualitative Research: A Guide to Design and Implementation, 4th Edition. San Francisco, CA: Jossey-Bass, 2016, pp. 195-236.
- [32] M. L. McHugh. "The chi-square test of independence," *Biochemia Med (Zagreb)*, vol. 23, no. 2, pp. 143–149, 2013.