

Comparative Analysis of Data Visualization Aesthetics: A User Study Framework

Angeline Moyer

Submitted under the supervision of Daniel Keefe to the University Honors Program at the University of Minnesota-Twin Cities in partial fulfillment of the requirements for the degree of Bachelor of Arts, *summa cum laude* in Computer Science.

April 23rd, 2021

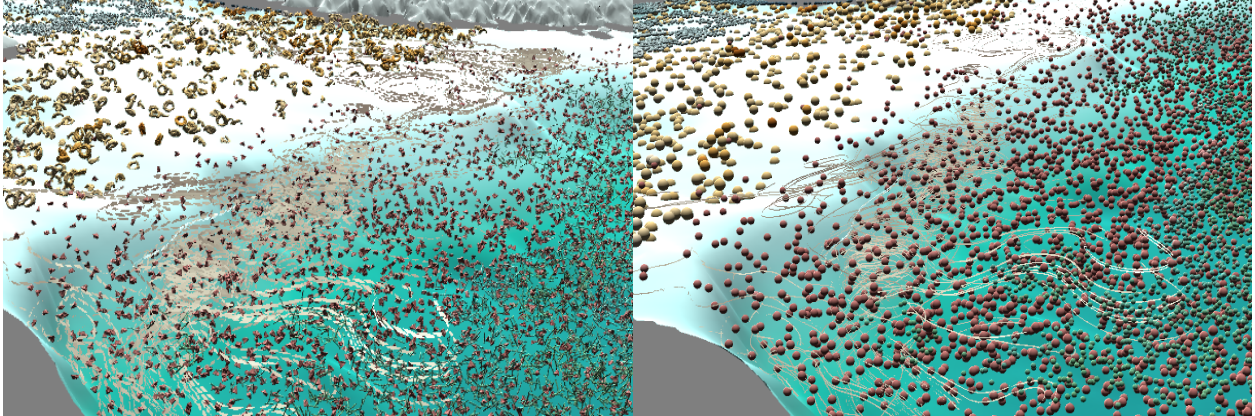


Figure 1: Side-by-Side visual of a SculptingVis scene with two different aesthetics

Abstract

In the world of data visualization, innovative new ways to represent and convey information are constantly being formed, such as in the Sculpting Visualization (SculptingVis) project. But how do we evaluate the human response to a data visualization, and how can we use this data to measure the way the data connects with individuals? To explore this issue, I present a user study framework to understand how user reactions to the artist-made data artifacts in SculptingVis may compare to reactions to traditional data visualizations, and how this affects user interaction and interpretation with the data in Virtual Reality (VR). I have compiled research on data visualization evaluation methods and user study practices in the context of a comparative user study framework that seeks to understand how users respond to the different forms of data. I provide the tools necessary to run an implemented user study and also a prototype of the resulting study. Additionally, I suggest ways this framework can be implemented and built upon, helping to propel further development and exploration of the topic.

Section 1: Introduction and Background

1.1 Introduction and Sculpting Visualization Aesthetics

This work was developed specifically with the Sculpting Visualization project in mind, and provides a framework for comparing user reactions to different data visualizations. Sculpting Visualization (SculptingVis) is a project currently in development, and explores bringing artists into the realm of scientific data visualization and representation [1]. In SculptingVis, artists and researchers can sculpt and create 3D artifacts with physical materials like clay, then scan and upload them to a library of artifacts to be used in scientific visualizations. This is a revolutionarily different approach to scientific data representation and brings in a human touch to the data. The SculptingVis team theorizes that being able to create glyphs to represent data enables individuals to form a deeper connection with the data. In the words of the team, “SculptingVis is predicated on the idea that scientific visualizations are more effective when they are engaging. Humans are drawn to beauty—drawn to engage with it, interact with it, and ponder it for longer periods of time.” [2]. There is also prior research that found artistic data visualizations to be effective in conveying memorable information and encouraging visual exploration [9]. With SculptingVis, a person may select or create a glyph that resembles a characteristic of the data it is representing. That could be a shape that is more “wave-like,” for example, or anything that resonates with the creator. The increased association and connection to the visuals is then theorized to improve the ability to interpret the data [4]. However, presently there is no formal framework for evaluating this theory for Sculpting Vis to traditional data visualizations in a comparative manner. The SculptingVis team has explored how individuals interpret the data aesthetics through close readings [5, 6] and has made many interesting

observations on the associative properties of SculptingVis aesthetics. Yet, these close readings examine the SculptingVis aesthetic as a standalone. The work presented here builds off of existing work for examining the associative qualities of the SculptingVis aesthetic, yet does so in a new comparative format. This puts SculptingVis next to a justified traditional data visualization style and examines the user interactions with the visuals in a side-by-side comparative style. This creates a new way to measure how this artistic data representation performs in direct contrast to a more standard approach.

1.2 Related Work

The area of user studies for analyzing aesthetics in scientific data contains many interesting dives into the topic. There has been some related exploration of aesthetics where specific vector techniques have been evaluated [7]. In fact, this work and its original pilot [8] were much of the inspiration for my work. In the full study [7], visual designers were brought in to critique scientific visualizations. The designers were asked to do three tasks related to the data that sought to expose the effectiveness of the visualization. The study found that quantitative evaluations are important and that nonexperts can often be effective at detecting problems with a visualization. The study framework I propose in this paper follows this and is created with the intention to be performed by inexperienced users. There is also relevant work on criticizing visualizations, where researchers ran a critique similar to those done in the art world, except instead of paintings hanging on the walls, their scientific visualizations were the subject of critique. [9]. A major point in their paper is that visualizations can be evaluated in the same way as art. I see user studies as an extension of this, where the user study is essentially a public critique of a visualization. Some of the relevant work in this area touches upon psychological elements where the human visual system is taken into account. This work on perceptual proxies

ran a user study on the ease of comprehending a data visualization and compared the results to known human perceptual tendencies [10]. They found in their case that “the established hierarchy of perceptual precision for single-value comparison does not extrapolate to judgments across larger sets of data.” This all the more stresses the importance of showing users a larger data visualization in addition to single objects. There are countless other resources on the psychology of interpretation of objects, such as the Greeble [11, 12], which can be extended to interpreting the unique glyphs in SculptingVis. Circling back to the original purpose of SculptingVis, works such as [13] cite climate visualizations as a challenging problem due to the large data sets, capturing intricate linkages between related areas of study, and the need for complex modeling. Sculpting Vis possesses the ability to handle large data sets with complex models, and here I seek to further explore how its aesthetic may help capture such intricate details.

1.3 The “Typical” Data Visualization

In order to evaluate the effects of any new style of data visualization, a baseline must first be established. To establish a baseline for SculptingVis, I looked through and categorized relevant examples to draft a “typical” data visualization aesthetic. Since the current application of SculptingVis is data in climate science, I turned to some established climate science conference papers through the American Geophysical Union (AGU) conference to see how this data is typically represented. Looking through the different papers, I recorded what symbols, colors, and methods they used in their visualizations to convey data. This includes line formations, shapes, and colors. Of the articles I reviewed, many visualizations included features such as: dotted and dashed lines [14, 15, 16], red-green-blue gradients [17, 18, 19], and basic symbols to represent data such as triangles, squares, diamonds and circles [20, 14, 15]. Specific

to 3D visualizations, I recorded features such as: layering and bright colors [21], rounded data points [22], and primary shapes like spheres and tetrahedrons [22, 23, 24]. A noticeable pattern in 3D visualizations in particular was the use of spheres differentiated by color. Based on these findings, I believe that a typical visualization would have contrasting bright colors, and contain objects from the typical primary shapes set, potentially going up to multifaceted polygons that retain some form of symmetry. Based on these conclusions, I created a set of new glyphs detailed in section 2.1.

One potential shortcoming of this method to address is that some of the prior work reviewed is two dimensional, while SculptingVis operates in a three dimensional environment. Despite this, my objective is to record the overall aesthetic of the works, and translate that into a 3D environment when necessary. I am also not considering the specific arrangement or nuances of the data representations, I am merely recording the general shapes, colors, and line patterns found. Shapes such as a 2D circle can easily be translated to a 3D sphere. Two dimensional is also the most prominent visual style used in the field. According to a survey of researchers on what types of climate science visualizations they regularly used [25], the majority of visualizations in the field are 2D, with 90% of researchers reporting use of time charts and 77% use of bar charts. Only 37% of individuals reported using some form of 2.5D or 3D visualization, which were most commonly height maps and not VR. One may find slight differences if looking exclusively at three dimensional data, however since examples of this are limited in the field and I am seeking to represent what is typical of the field, I believe the typical characteristics identified in two dimensional visuals are reasonable to include.

Section 2: Creating Visual Comparisons

To conduct a study on comparing data visualization, a toolkit for creating visual comparisons must exist. In the scope of this project, I addressed this by creating new glyphs, camera flythrough recordings, and identical state files that can be used for side-by-side data visualizations.

2.1 Creating new Glyphs

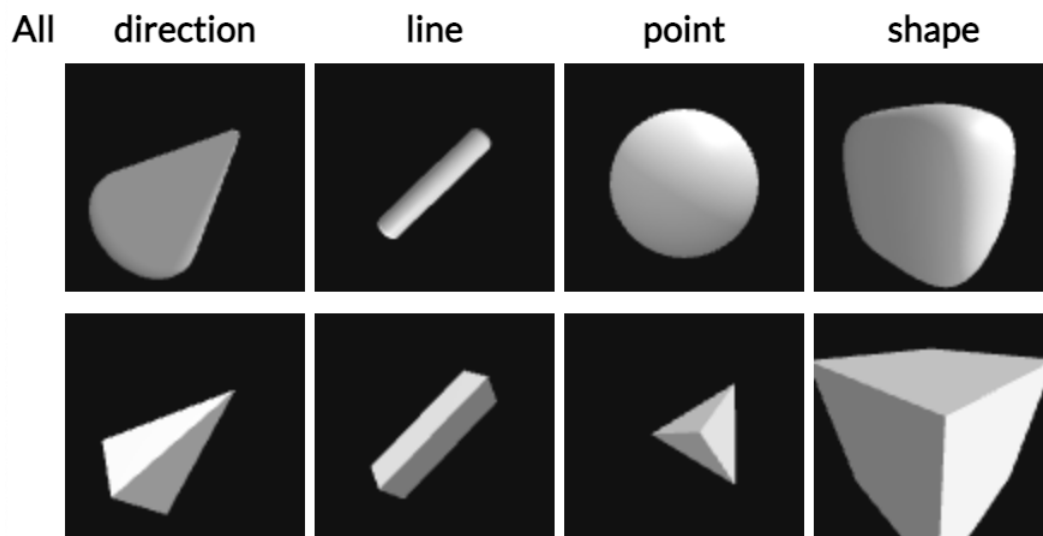


Figure 2: Artificial glyphs as they appear in the SculptingVis library. Top row is curvilinear geometric glyphs and bottom row is linear geometric.

The first step in creating comparison visualizations is creating a new set of glyphs. These follow the typical aesthetic covered in section 1.3. I created two new types of SculptingVis glyphs: curvilinear geometric and linear geometric. Determining these involved discussing with the SculptingVis team which shapes would accomplish my goal and also be useful to the team in the long run. Initially, I planned to create only linear shapes because of their contrast to the

smooth organic feel of the existing glyphs. After some discussion and consideration of prior work however, I expanded this to include the curvilinear shapes. This gives a greater range of expression while still maintaining a computer-generated aesthetic. The process of creating these was fairly simple. I created the 3D objects using Blender and exported them as .OBJ files. No specific textures were used and I selected a number of vertices and faces high enough for each object so that their faces retain a smooth curve when enlarged. Those files were then uploaded to the SculptingVis library to be used to create visualizations. The glyphs can currently be found in the public library [26] under the class name `z_Artificial`, and can be seen pictured in Figure 2. This “typical” aesthetic contrasts the handmade feel of SculptingVis well. Where artifacts from SculptingVis contain artistic notches and irregular shapes, the “typical” artifacts are computer-generated, smooth, and symmetrical.

The existence of these new glyphs enables the creation of identical state files, where everything is identical in a state except for the glyphs used. This ability is an essential component of running a comparative visual study, and the range of glyphs ensures there is a “typical” counterpart to most of the existing artist-made SculptingVis glyphs.

2.2 Creating Visual Flythroughs

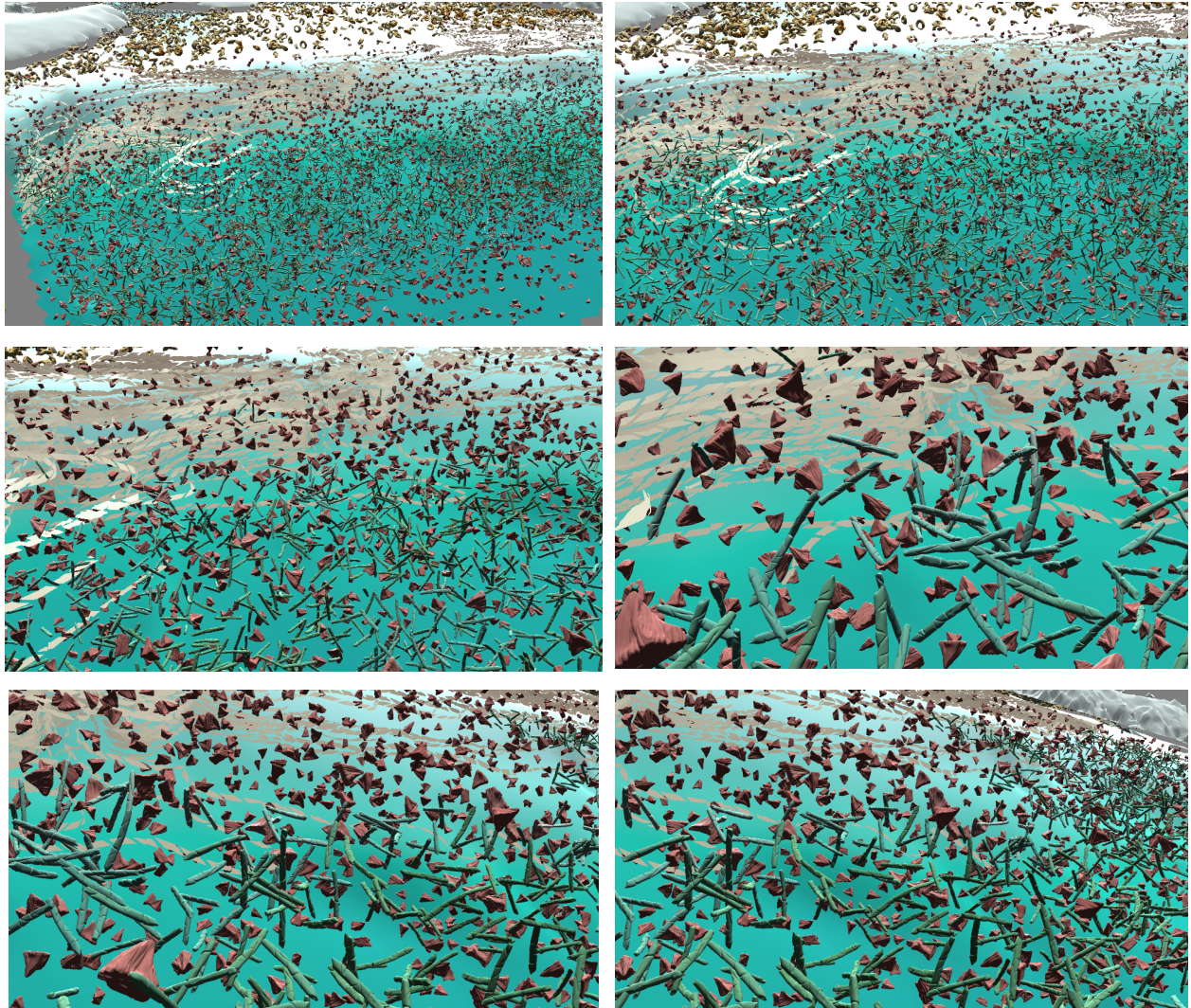


Figure 3: Stills of a flythrough, read top to bottom left to right.

As a key component of conducting a user study on comparing visualizations, I created flythroughs that could be shown to a participant during said study. The key benefit of flythroughs is that they can convey the dimensionality of a 3D environment in a 2D delivery method. This makes them an ideal strategy in conducting a remote user study. Several key objectives were on my mind when I created the flythroughs: Seeing glyphs relative to each other and in the context of an environment, seeing full 360 views of glyphs, and seeing how the two glyph styles appear

differently in identical environments. The flythroughs I recorded sought to capture these objectives. In Figure 3 some sample stills from a flythrough are placed in sequence. One can see that the flythrough begins further away and zooms in on some glyphs, then rotates to show the glyphs at more angles. Starting zoomed out orientates the user to the relative size of the imagery, and rotating once up close with the glyphs helps to grasp the depth of the scene and the three-dimensionality of the individual glyphs. As part of this project, a repository of recordings of these flythroughs was created and shared with the team for use in any future conducted studies. The flythroughs were created using the existing DeltaMover scripts from the SculptingVis project. A very specific process had to be followed in order to replicate the flythroughs in two different scenes. Since there was no existing documentation on how to use this tool, I wrote detailed documentation on how I accomplished the creation and reproduction of the flythroughs to aid in further production of them in the future.

Section 3: Designing the Study

3.1 User Study Goals

To design a study that captures the different reactions people have to data styles, I had to consider many different factors, as well as the goals of the study. The primary purpose of the study is to capture the reactions and ability to analyze and interact with the SculptingVis data aesthetics as compared to a traditional style. The specific objective of this portion of the project was to create an experimental design for a user study which accomplishes the primary goal and utilizes the created visual comparison tools. Breaking that down further, the goal of the study is

to gain comparative insights into the associations users have with the individuals glyphs, and gather data on the comparative ease of interactivity between the two aesthetics.

3.2 Task Identification

In what I call the task identification stage of developing this method, I considered many types of questions and tasks that users would be asked to complete in a study format. Metrics from these task-centered questions provide data that can be examined to understand how users respond to different visualization styles. In considering what tasks most effectively accomplish this goal, I considered the complexity of the task and if the results could be easily quantified for analysis. Simpler, more quantifiable responses provide ease of analysis, especially if the study is run with many participants. Therefore, in a specific associative task I would keep responses limited to a few words. Additionally, keeping complexity reasonable in any task is an important factor so that users with little experience or knowledge of the subject matter can still complete the task without losing too much of its meaning. I also considered the imagery shown to the participants and how it can be read and interpreted. Specific aspects of this include considering the depth of field of any images, the resolution, and speed for any moving imagery.

I considered tasks including identification, point-and-click, ranking, sorting and grouping. Each of these can be implemented in different ways to achieve different results. In an identification task, users would be asked to identify different qualities of two comparative images. They may be shown a scene with glyphs and asked to identify all of the different symbols they see in the image. This could go a level further and ask more specific identification questions, such as giving the user a specific shape and asking how many variations of that shape the user sees. Identification tasks can also be a bit more scientific, to get at the ease of interpretation of the data that the user may experience. For example, a user may be shown an

image and given some basic information on what it represents, then would be asked to identify where on the image they believe a specific scientific phenomenon is occurring. Depending on the target audience for the study, these questions would have to be understandable enough for a user unfamiliar with the subject matter. A reasonable example of this would be informing the user what a glyph represents, then asking the user to identify an example of its represented quality, such as a tendency to group together.

In a point-and-click task, the user can be shown an image and their response would be recorded as where they physically clicked in the image. An example of this would be showing the user an image and asking them to click on the glyph they notice first in the image. Variations of this question could be asking what glyph reminds the user most of a certain shape, or the glyph their eyes are drawn to the most. Many identification tasks can be implemented with the point-and-click method, such as the scientific identification previously mentioned. However, a further implementation of a scientific point-and-click task could include timing, where the metric response is how long it takes for the user to click on the image, as well as where they clicked.

In a ranking task, the user may be asked to rank different qualities of two comparative images on a 5-point Likert scale [27]. These qualities would target specific aspects of the visuals in an attempt to gather insightful data. An example of this could be asking a user to rate how familiar two comparative images are. These questions could be asked relative to an entire scene of glyphs in their contextual environment or to a standalone glyph.

As mentioned above, most of these task-types can be modified to fit the specific needs of different studies. However, in my case I select some specific tasks aimed at understanding how reactions to different data visualization styles may vary.

3.3 Experimental Procedure and Implementation

The study I am proposing is an experimental procedure designed to compare two different visualization styles to measure their effect on users. The general structure of the study centers around this comparison, and also involves some independent glyph analysis. When I speak of comparative images, I am referring to the exact same data visualization replicated with different glyphs. I suggest beginning with comparison of individual glyphs, particularly the SculptingVis glyphs to orient the participant to its unique style. Then, increase the user interaction with the glyphs with grouping and other glyph-centered tasks. Finally, glyphs are shown in their complex environment and the user performs tasks from those images. The next section will contain a sample runthrough of a study.

The study itself is an arrangement of certain tasks to evaluate visualization strategies. This is merely a methodology for a study, yet was created with a remote format in mind. In other words, I do not assume users are physically present for the study, nor do I assume they have access to a VR headset. These assumptions have the major benefit of accessibility. By ensuring this study can be perceived without a VR headset and without being physically present in a specific location, many more are able to participate when they otherwise would not have been able to. The overall study design that I propose in the following section therefore, is intended for delivery in an online remote format such as a webpage or crowdsourcing platform, and has no specific equipment requirements.

Section 4: Resulting Study Runthrough

In this section, I provide a runthrough of a user study implementation of the study methodology I have described. This is the resulting content of this work and can be applied to many different delivery methods. This version was created with remote delivery in mind and does not assume access to VR equipment. Throughout the study, only one question is visible at a time and all answers are kept relatively short to help elicit more “first-impression” answers.

The study begins with basic questions asking the user’s familiarity with the topic at hand. Topics include VR, data visualization, scientific climate data, and sculpting. The user can select their familiarity level on a 5-point scale ranging from “Completely Unfamiliar” to “Very Experienced.” The low end of the scale is defined as no prior experience or knowledge on the topic, while the high end indicates that the participant has worked in the topic area in some professional capacity. These provide interesting metrics on the users participating in the study, and may show interesting trends with the results. An overview of the time commitments of the study and clarification that no headset is required to complete this study are also included at the start of this stage. This stage can be seen pictured in Figure 4.

Thank you for participating in this study. No special equipment or knowledge in any particular area is required to complete the study. It should take around 40 minutes to complete.

For demographic statistics, please rate your own level of expertise in the following fields:

Data	<input type="text"/>
Visualization: Virtual Reality Scientific Climate Data	<input type="text" value="1 - Completely
Unfamiliar
2 - Somewhat"/>
Visual	<input type="text"/>
Design/Art Sculpting	<input type="text"/>

Figure 4: A sample image of what the study opening page looks like.

Moving into actual study content, the user is first presented with a series of still images of SculptingVis glyphs, and is asked to write a single word that first comes to mind to describe the glyph. The text field accepts only single-word answers for ease of later analysis. Each question appears standalone on the page, as pictured in Figure 5. The images come directly from library screenshots and contain no color, so everything is standardized. These questions show the glyphs out of their normal context and seek to capture the immediate association an individual has with its shape. The second half of this stage then shows a series of still images of the typical glyphs and asks the same question: to write the first word that comes to mind to describe the glyph. In total, around ten glyphs should be shown at this stage, five SculptingVis and five traditional. At this point, the user has not received any explicit context to the shapes they have seen. This is

designed to be a relatively quick portion of the study and users are encouraged to not spend too much time on one question.

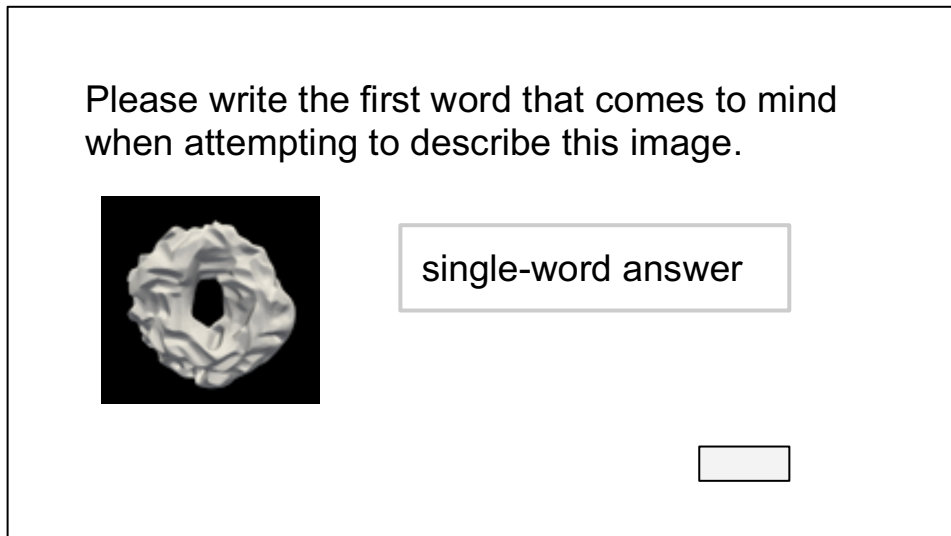


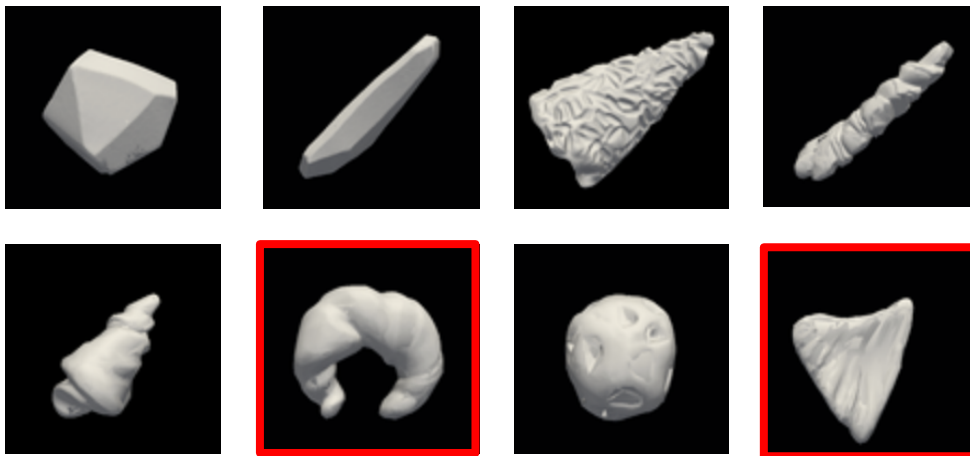
Figure 5: Still images of glyph description question



Figure 5: Still images of glyph description question

The next stage of the study still does not involve contextual information, but shows a set of glyphs and asks the participants to categorize them. The glyphs are shown in the same black-and-white style as previously, except now the user sees multiple glyphs, as shown in Figure 6. The user is asked a series of questions that ask them to identify and group glyphs based on their

characteristics. In practice, I believe the most straightforward way to implement this would be to allow participants to select multiple answers, similar to a check-all-that-apply style of question. Questions in this section range from characteristics pertaining to climate data to informal associative qualities. For example, a data-driven question may ask “Which of these glyphs could represent something in water?” Similarly, an example of an informal associative question would be “Which of these glyphs remind you of food?” Participants can then click as many of the glyphs that they feel this applies to. Another question asks participants to arrange glyphs into families or place them in order. This requires a card style of implementation, where users can click and drag to move glyphs into a particular order, such as in Figure 7. I suggest keeping the glyph set shown consistent between all questions to explore the different ways the glyphs can be interpreted. The same questions and tasks can then be repeated for a traditional glyph set.



Which of the pictured glyphs look familiar? Click all that apply, then click Next

Next

Figure 6: Grouping of several glyphs. In this image, the user selected two glyphs. The quality “familiar” can be replaced with any associative or surmised representational qualities.

1	2	3	4
5	6	7	8

Rank the glyphs by order of familiarity, with 1 being most familiar.

Click and drag to move to the glyphs into the desired spot, then hit Next when finished.

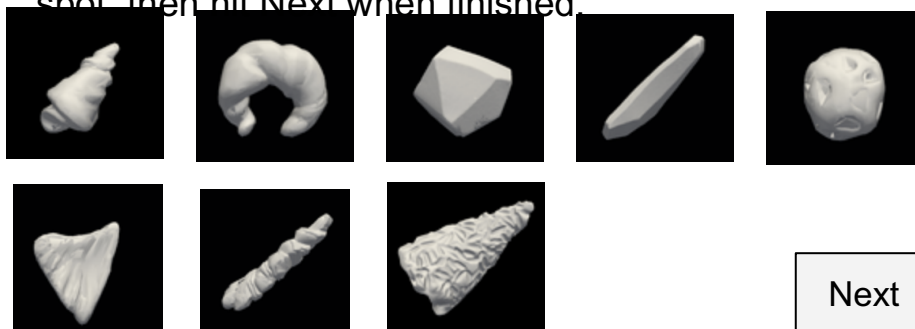


Figure 7: Sorting task of several glyphs

Click on the part of this image that you notice first

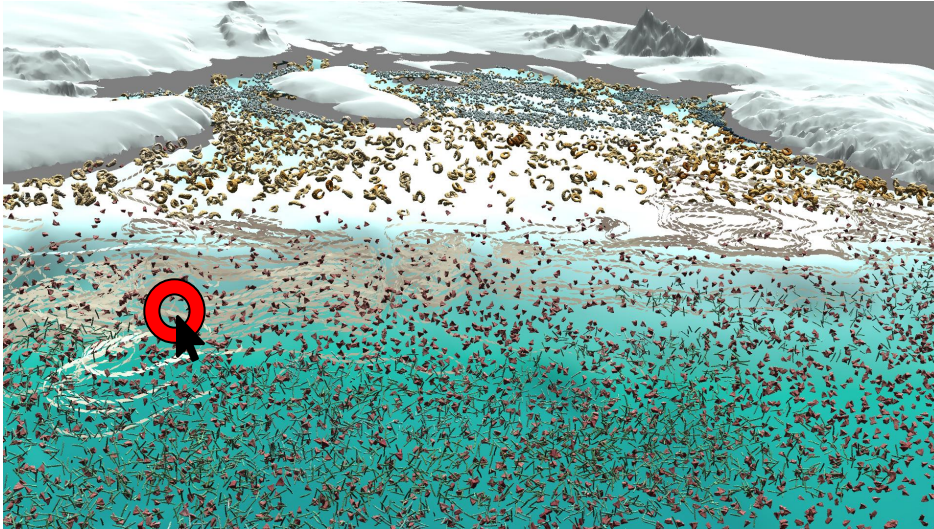


Figure 8: Click to identify task with feedback to indicate to the user where they clicked.

In the final stage of the study, participants now see complete images of the glyphs in their full contextual environment. Participants are shown a variety of images and short video flythrough recordings from two state files, one of each aesthetic, and are asked a series of questions on the images. One question shows the user a still image of a complete visualization, as in Figure 8, and asks them to click on the data point that they notice first. The image gives visual feedback to indicate where the user clicked. The prompt to the user can be more associative in nature, asking the user to click the center of the region where they see the most “familiar” looking glyphs. “Familiar” in this case can be replaced with any associative quality. The same style of question is then repeated with a traditional data visualization.

In similar style to prior questions, the user is next shown a brief video or still image of a complete visualization, but the time spent on the question is recorded. Here, the user is shown an individual image of a glyph and asked to click on the first glyph they see in the visualization that matches the shape of the glyph. A question like this may have a layout as depicted in Figure 9. An alternative to this question asks the user to identify where they believe a specific scientific phenomenon is occurring. This question is kept intentionally simple, and could be something such as identifying the highest concentration of a particular glyph. These questions aim to evaluate how easily an average user can see patterns in the data.

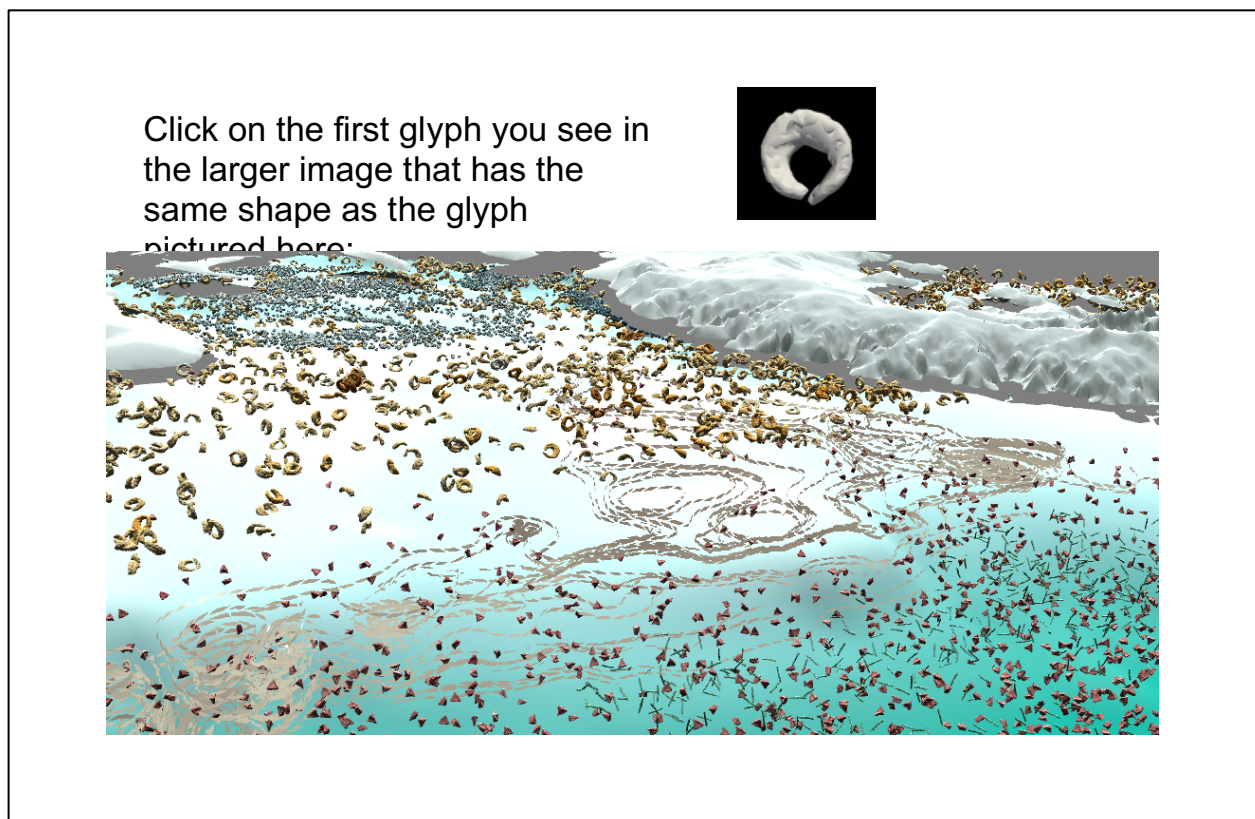


Figure 9: Identifying a specific glyph in a larger visualization

Section 5: Discussion of Work

5.1 Discussion of Study

The goal of this work was to create a reproducible framework which can be used to run a comparative study on SculptingVis versus a typical aesthetic. I believe this goal was achieved, as the resulting tools allow for the creation and running of such a study. The study methodology and variety of tasks create a well rounded look into how an individual perceives and interprets these visual aesthetics. Results of this study could reveal a wide variety of interesting tendencies and associations, and perhaps common patterns among data interpretation. Should this be implemented in a crowdsourcing software, this study would also be the first chance for mass-participation in a comparative study on SculptingVis, and would give valuable insight to human interaction with the project.

In the description of tasks and in the study runthrough, I intentionally left items open-ended, and gave multiple samples of what questions may be asked. In a true choose-your-own-adventure style, I believe that this researched framework can be implemented in many different ways depending on the specific goal, the target audience, the tool, and even the study delivery method. In my case, this work was designed for Sculpting Visualization and I wanted to ensure multiple approaches were covered for whatever factors may exist when this is implemented. Part of the beauty of research in this area is that it can be used as inspiration for further studies and research as well. By keeping tasks and delivery formats customizable, I hope to foster that inspiration in others.

5.2 Limitations

In the study itself, I believe the biggest challenge is the organization of the three study sections: individual glyphs, grouped glyphs, followed by full-scene imagery. The prototype study runthrough I presented in this paper follows that ordering. However, limiting the study to be run in only this order risks introducing a bias to the study. By introducing glyphs by themselves at the start of the study, users may form a certain level of familiarity with the glyphs that gives an advantage in the later part of the study, when they see the glyphs in a full environment. This is undesirable because it does not accurately simulate the experience someone would have if looking at a complete visualization for the first time. One can make the argument that in practice it is likely researchers would encounter individual glyphs before seeing a complete visualization, however I do not believe this fully negates the issue, as the study will not necessarily be run on a group of researchers familiar with SculptingVis. A viable solution to this issue is to randomize the order of the study sections when running it with large groups. This would mean randomly distributing the order of the three sections when running the study. A balanced randomization between three objects should not be difficult, and can certainly be done in implementation.

I also made some assumptions about target audience and delivery format in my method that create limitations for further work. Due to the study being developed during a period of remote learning, I attempted to create something that would not require access to VR equipment or physical presence in a space. This has its benefits in accessibility, but also has the disadvantage of needing to rely on recorded flythroughs to grasp the depth and dimensionality of a data visualization scene. Obviously, the experience would be much richer with complete access to VR technology spaces. This also means an in-person implementation would look different from what I proposed here. In an implementation of the study run in real life, none of the

comparative still images would be necessary. The complete visualization would simply be made viewable to the participant in VR. These differences provide a whole new set of considerations when implementing, and offer up many opportunities for further work in the area, discussed more in section 5.3.

5.3 Impact of Work

The main impact of this work is that it adds to the toolkit of visualization evaluation methods in its unique lens of direct aesthetic-based comparisons. Particularly in the vis community, folks are often most interested in evaluating how their work is received, yet studies like this one which directly compare a hand-crafted aesthetic to a more established traditional geometric aesthetic are far less common. This experimental design adds to the toolkit of the broader visualization community, but also adds something completely new to the toolkit of the SculptingVis team that will aid them in further evaluation of Sculpting Visualization. All of the work that I have done is made directly available to the team for use and further development.

5.4 Further Directions

Much of this work involved gathering imagery—specifically flythroughs—that could be shown to a participant in a remote format. I found that the exploration of capturing flythroughs opens up new questions on evaluating these. For example, what constitutes a “good” flythrough? Are some flythroughs better than others, depending on your target use or audience? This is a fascinating unintentional spin-off research question that I encountered. While I do not directly explore this topic, I do explain my criteria for a suitable flythrough for my purposes.

As alluded to throughout this paper, there are also many different directions in which someone could expand this work. My specific use case was to directly compare SculptingVis aesthetics to a traditional aesthetic, though much of this methodology could be expanded to explore different related topics, such as the effectiveness of a single aesthetic. Some may also opt to survey a group with specific expertise on the subject of the visualization, and therefore may wish to expand on the complexity of the scientific tasks.

In early phases of this project, a discussed implementation was creating a switch functionality that allows the user to click a button and flip between two visualizations. In this case, that would be swapping instantly between a visual with traditional glyphs and SculptingVis glyphs, as pictured in Figure 1. Implementing this provides benefits beyond simply running a study, but could also be used to showcase multiple state files at once in a demo. A potential way to create this would be to cache an existing data state and provide a menu item to swap your active state with the cached one. This would work well for user studies as one could swap between contrasting visuals, and it would be useful in the everyday situations where a researcher may want to compare two different datasets. A more specific way to go about creating this feature would be to build it out to the specific use case of swapping glyph sets for the same visualization. Meaning, a menu option exists to swap out the current glyph set with a different predefined one for the same data set and visual scene. While this satisfies the exact purpose for a user study, its application is somewhat limited.

Another potential direction that would be useful for both a remote and in-person format is the ability to automate the replication and repetition of a flythrough path between state files. One might consider implementing specific camera pathing with pre-set stops that display a desired section of the visualization that can be shown to a study participant, or to an interested party

during a demo of the tool. A contribution of this project includes written documentation in shared materials about the process for the flythroughs I created, which can be considered if the team opts to automate some of these functions.

Section 6: Conclusions

There were two specific goals of this project, both of which were achieved. The first goal was to expand upon existing work and provide comparative tools for evaluating visualizations. This was completed in the form of new “typical” glyphs, flythrough recordings and process, and the ability to create identical state files. The second goal was to draft an experimental design for evaluating the SculptingVis aesthetic in direct comparison to a typical aesthetic. The tasks and framework describe this, and a resulting study mockup that follows the framework was also created. The study sought to uncover intuitive associations, levels of engagement, and interpretations individuals have with the SculptingVis aesthetic, which is a topic that has been explored in prior close readings on SculptingVis. However, the study also achieves a new direct comparison approach, which frames the aesthetics in a direct comparative manner and has not been done before on the team. Overall I have presented research and a new framework for comparing and evaluating visualization aesthetics. This can now be picked up and implemented by both the SculptingVis team and others interested in comparative analysis of aesthetics.

References

- [1] Sculpting Visualization, *Sculpting Visualization*. [Online]. Available: <https://www.sculpting-vis.org/>. [Accessed: 02-May-2021].
- [2] Francesca Samsel, Seth Johnson, Annie Bares, and Daniel F. Keefe. “Scientific visualization: Enriching vocabulary via the human hand.” *Proceedings of the IEEE VIS Arts Program(VISAP) 2019*.
- [3] Kozik, P., Tateosian, L. G., Healey, C. G., & Enns, J. T. (2019). Impressionism-inspired data visualizations are both functional and liked. *Psychology of Aesthetics, Creativity, and the Arts*, 13(3), 266-276. doi: 10.1037/aca0000175
- [4] Francesca Samsel, Annie Bares, Seth Johnson, PJ Wolfram Jr, and Daniel Keefe. “Sculpting Visualization: Expanding the vocabulary of scientific visualization, facilitating the construction of engaging communication of complex science.” In *AGU Fall Meeting Abstracts*, volume2019, pages ED51D–0828, 2019.
- [5] A. Bares, "Close reading for visualization evaluation", *IEEE Computer Graphics and Applications*, vol. 40, no. 04, pp. 84-95, Jul/Aug. 2020.
- [6] A. Bares, S. Zeller, C. D. Jackson, D. F. Keefe and F. Samsel, "Using Close Reading as a Method for Evaluating Visualizations," 2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV), 2020, pp. 29-37, doi: 10.1109/BELIV51497.2020.00011.
- [7] D. H. Laidlaw et al., “Comparing 2D Vector Field Visualization Methods: A User Study,” *IEEE Trans. Visual. Comput. Graphics*, vol. 11, no. 01, pp. 59–70, Jan. 2005, doi: 10.1109/tvcg.2005.4.

- [8] C. D. Jackson, D. Acevedo, D. Laidlaw, F. Drury, E. Vote, and D. Keefe, “Designer-critiqued comparison of 2D vector visualization methods,” presented at the the *SIGGRAPH 2003* conference, 2003, doi: 10.1145/965400.965505.
- [9] R. Kosara, F. Drury, L. E. Holmquist, and D. H. Laidlaw, “Visualization Criticism,” *IEEE Comput. Grap. Appl.*, vol. 28, no. 3, pp. 13–15, May 2008, doi: 10.1109/mcg.2008.63.
- [10] L. Yuan, S. Haroz, and S. Franconeri, “Perceptual proxies for extracting averages in data visualizations,” *Psychon Bull Rev*, vol. 26, no. 2, pp. 669–676, Sep. 2018, doi: 10.3758/s13423-018-1525-7.
- [11] “Greeble (psychology),” Wikipedia, 20-Apr-2021. [Online]. Available: [https://en.wikipedia.org/wiki/Greeble_\(psychology\)#cite_note-3](https://en.wikipedia.org/wiki/Greeble_(psychology)#cite_note-3). [Accessed: 02-May-2021].
- [12] I. GAUTHIER and M. J. TARR, “Becoming a ‘Greeble’ Expert: Exploring Mechanisms for Face Recognition,” *Vision Research*, vol. 37, no. 12, pp. 1673–1682, Jun. 1997, doi: 10.1016/s0042-6989(96)00286-6.
- [13] J. Johansson, T.-S. S. Neset, and B.-O. Linnér, “Evaluating Climate Visualization: An Information Visualization Approach,” presented at the 2010 14th International Conference Information Visualisation (IV), Jul. 2010, doi: 10.1109/iv.2010.32.
- [14] C. Sutton, S. Kumar, M.-K. Lee, and E. Davis, “Human imprint of water withdrawals in the wet environment: A case study of declining groundwater in Georgia, USA,” *Journal of Hydrology: Regional Studies*, vol. 35, p. 100813, Jun. 2021, doi: 10.1016/j.ejrh.2021.100813.

- [15] V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters, “Mathematical and physical ideas for climate science,” *Rev. Geophys.*, vol. 52, no. 4, pp. 809–859, Dec. 2014, doi: 10.1002/2013rg000446.
- [16] A. E. Kelly and M. L. Goulden, “Rapid shifts in plant distribution with recent climate change,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 33, pp. 11823–11826, Aug. 2008, doi: 10.1073/pnas.0802891105.
- [17] H. M. Worsham, “An Examination of an Enhanced Remote Sensing Method for Agent Attribution of Forest Disturbance.” Cold Spring Harbor Laboratory, Nov. 23, 2020, doi: 10.1101/2020.11.23.394221.
- [18] Craig Anderson, Aaron J Bever, Michael MacWilliams, and Matt Nobriga. “Evaluation of the effect of delta outflow on the water temperature of the low salinity zone and potential management implications for delta smelt.” In *AGU Fall Meeting 2020*. AGU, 2020.
- [19] H. Yang et al., “Poleward Shift of the Major Ocean Gyres Detected in a Warming Climate,” *Geophys. Res. Lett.*, vol. 47, no. 5, Mar. 2020, doi: 10.1029/2019gl085868.
- [20] Ahmad Shakeel, Ryan R MacIver, Claire Chassagne, and Andrew J Manning. “Rheological And Structural Analysis of Mud from Port of Hamburg, Germany.” In *AGU Fall Meeting 2020*. AGU, 2020
- [21] Laurent Ailleres, Mark Jessell, Eric deKemp, Guillaume Caumon, Florian Wellmann, Lach-lan Grose, Robin Armit, Mark Lindsay, Jeremie Giraud, Guillaume Pirot, et al. “Loop: an Integrated and Interoperable Platform Enabling 3D Stochastic Geological and Geophysical 3D Modelling.” In *AGU Fall Meeting 2020*. AGU, 2020

- [22] Jaime Aguilar Guerrero, et al. "From simulated data to AR: 3D Visualization Techniques for Meaningful Representation of Atmospheric Gravity Wave Data and AR Applications." *In AGU Fall Meeting 2020*. AGU, 2020
- [23] Yuling Wu, Ajinkya Kulkarni, Brian Ellingson, Manil Maskey, and Sara J Graves. "Exploring storm morphology, dynamics and lightning characteristics during the lifecycle of a multicell cluster in a field campaign using a 3D visualization tool." *In AGU Fall Meeting 2020*. AGU, 2020
- [24] Mahda M Bagher, Pejman Sajjadi, Julia C Carr, Lorne N Leonard, Peter C La Femina, and Alexander Klippel. "Visualizing earthquake locations in an immersive learning environment." *In AGU Fall Meeting 2020*. AGU, 2020
- [25] Thomas Nocke, Till Sterzel, Michael Böttinger, and Markus Wrobel. Visualization of climate and climate change data: An overview.
- [26] SculptingVis Library. [Online]. Available: <http://sculptingvis.tacc.utexas.edu/library/>. [Accessed: 03-May-2021].
- [27] Likert, R. "A technique for the measurement of attitudes." *Archives of Psychology*, 22 140, 55, 1932