Using Visual Design Expertise to Characterize the Effectiveness of 2D Scientific Visualization Methods

Daniel Acevedo, David Laidlaw ({daf,dhl}@cs.brown.edu) **Cullen Jackson** (cjackson@aptima.com) **Fritz Drury** (fdrury@risd.edu)
Computer Science Dept., Brown University **Aptima Inc.** **Illustration Dept., RISD**

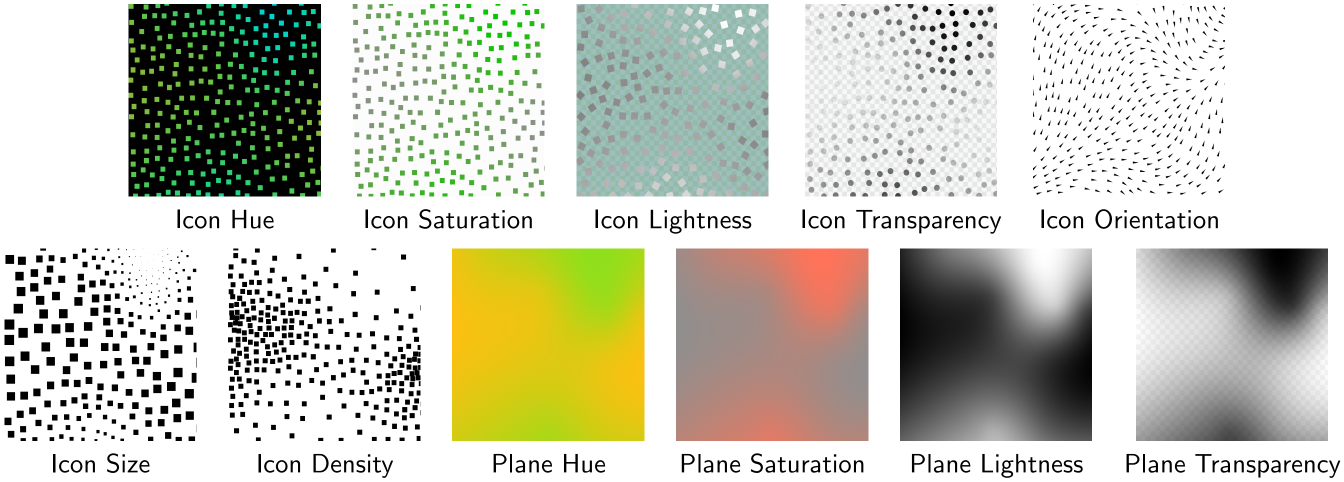


Figure 1: Eleven different visualization methods that represent the same continuous scalar dataset. We are characterizing the effectiveness of each one of these methods, both individually and in combination, to represent scalar datasets in 2D.

We present the results from a pilot study that evaluates the effectiveness of 2D visualization methods in terms of a set of design factors, which are subjectively rated by expert visual designers. In collaboration with educators from the Illustration Department at the Rhode Island School of Design (RISD), we have defined a space of visualization methods using basic visual elements including icon hue, icon size, icon density, and background saturation (see Figure 1).

In this initial pilot study we presented our subjects with single variable visualization methods. The results characterize the effectiveness of individual visual elements according to our design factors. We are beginning to test these results by creating two-variable visualizations and studying how the different visual elements interact.

1 Introduction

Given the increasing capacity of scientists to acquire or calculate multivalued datasets, creating effective visualizations for understanding and correlating these data is imperative. However, modeling the space of possible visualization methods for a given scientific problem has challenged computer scientists, statisticians, and cognitive scientists for many years [1,2,3,4]; it is still an open challenge. Our goal is to provide scientists with visualization methods that convey information by optimizing the design of the images to facilitate perception and comprehension.

We created a framework for evaluating these visualization methods through feedback from expert visual designers and art educators. Our framework mimics the art education process, in which art educators impart artistic and visual design knowledge to their students through critiques of the students' work.We established a set of factors that characterize the effectiveness of a visualization method in displaying scientific data. These factors include constraints implied by the dataset, such as the relative importance of the different data variables or the minimum feature size present in the data. We also include design, artistic, and perceptual factors, such as time required to understand the visualization, or how visually linear is the mapping between data and visual element across the image. We will describe these in detail in section 2.

Evaluating the effectiveness of visualizations is difficult because tests to evaluate them meaningfully are hard to design and execute [5]. We have researched this issue previously in two user studies comparing 2D vector visualization methods. The first one [6] used scientists to evaluate 6 visualization methods, and the second one [7] studied the validity of subjective measures to evaluate the same methods using designers as subjects. Results indicated that the designers rated the visualization methods in a pattern similar to the results of the scientists of the first study. We also found that designer critiques generally took less time and that designers were able to provide methods for improving the visualizations. This result provides key support for using subjective expert design knowledge as the basis for our visualization effectiveness characterization.

The current pilot study builds on those two studies and is the initial step towards our final goal, which is to create a mathematical model of the knowledge collected from design experts and use that model to find an optimal solution for a data visualization problem.

2 Definitions

In this section, we will define the two main components of our study: our set of visualization methods and the design factors we defined to characterize them.

In general, a visualization method takes a scientific dataset and produces a visualization display. A method corresponds to a layered combination of our visual elements (see Figure 1), where the different data variables being represented are mapped to one or more of the available elements. To express these mappings we created Evolvis, a language for describing multi-layered scientific visualizations of multivariate 2D datasets. We have used this language to generate all the images used in our studies, from the single-variable visualizations shown above to complex examples of multi-layered multi-variable displays.

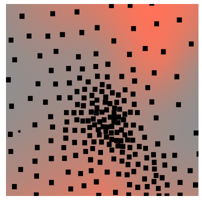
The second component, the design factors, are quantitative measures of the effectiveness of a visualization method. The goal of our visualizations is exploratory: scientists need an accurate representation of their data but have no simple specific tasks in mind, other than exploring how the different variables interact. In this sense, the factors we define here provide information about the quality of the data presented and the capability of a visualization method to work in combination with other methods. Said factors are:

- *data resolution:* the number of different levels of a data variable that can be distinguished by a viewer;
- *feature resolution:* the minimum spatial feature size that can be reliably represented with the method, expressed as a percentage of the image width;
- *linearity:* the perceptual linearity of the mapping from data value to visual property; this factor is measured by asking subjects to indicate the locations where they see the values of 0, 0.25, 0.5, 0.75, and 1.0 along the image for a linear dataset visualization;
- *visual bandwidth:* the percentage of of a method that can be covered when combined with other methods but still remain readable;
- *dominance:* the forcefulness or *punchiness* of the data mapping. This indicates how much a method would dominate the composition when combined with other methods, measured as a value from 0 to 10;
- *time to read:* the time it takes an average user to comprehend the data, measured in seconds.

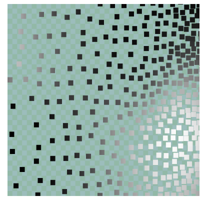
Bertin [1] developed a similar classification of his "retinal properties" (size, value, texture, color, orientation, and shape) according to their level of organization (whether they could be used to represent quantitative, qualitative, or ordered information) and the number of steps they could take (our data resolution factor). Our factors introduce new measures, like linearity,

	Highest	Average	Lowest
Data Resolution	Icon Orientation 3 ***(18)***	Icon Hue 2 ***(10)***	Icon Transparency 2 ***(1)***
1/(Feature Resolution)	Plane Hue 1 ***(50)***	Icon Orientation 2 ***(18)***	Plane Saturation 3 ***(1)***
Visual Bandwidth	Plane Lightness 1 ***(99%)***	Plane Hue 3 ***(50%)***	Icon Density 3 ***(0%)***
Dominance	Icon Transparency 2 ***(8.3)***	Icon Saturation 2 ***(5)***	Icon Lightness 2 ***(2)***
1/(Time to Read)	Plane Hue 1 ***(1.5)***	Icon Density 3 ***(0.45)***	Plane Saturation 3 ***(0.04)***

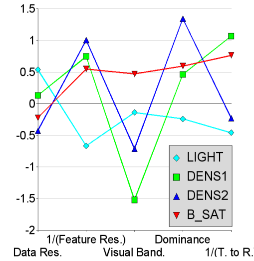
Summary of results for single-variable



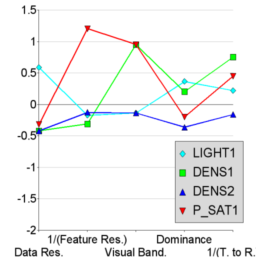
Saturation 1/Density 1



Lightness 1/Density 2



Single-variable z-scores



Two-variable z-scores

Figure 2: For our pilot study, we showed subjects 3 different parameterizations for each of the 11 visual elements (see Figure 1). On the summary table, the number following the element's name indicates which parameterization obtained those results. Next, we show two examples of two-variable visualization methods and evaluation results (z-scores) for each design factor, both individual and in combination. Note how values change when another dataset is present. The perceptual conflict between visual elements is obvious in the 'Lightness/Density' example.

and also capture some composition characteristics, like visual bandwidth and dominance. Our data resolution and feature resolution factors capture the fact that we are targeting quantitative datasets.

3 Methods

Building on these visualization methods and design factors, we have developed an approach for acquiring knowledge about our space of visualization methods. We have expert illustration educators critique and rate simple visualizations where only one visual element changed. The goal is to determine the relative effectiveness of each of these visual cues for representing single-valued data (see Figure 3). At this stage we also evaluate our set of design factors based on the comments from expert designers.

During their critiques, our subjects provide three different kinds of information for each design factor: numeric ratings, specific suggestions for directions of improvement, and explanations of their ratings. We videotape the sessions, which last approximately 3 hours, and we encourage in-depth explanations of their numerical ratings.

We are also beginning our exploration of how combinations of elements work together. In this case, our subjects critique and rate combinations of visual elements to map the ability of cue combinations to represent complex relationships within multivalued data sets. The two images on the left of Figure 2 show two examples of visualizations of two-variable datasets. This step allows us to understand how the individual visual elements interact when put together in the same visualization display. We will compare the results obtained here with the results of the single-variable study to better characterize the element interactions.

4 Results

Four illustration educators have performed the study. We have obtained a characterization of methods for each of the design factors studied, summarized in the table in Figure 2. We also obtained the z-scores for an easier comparison across factors, effectively normalizing all values to have a 0 mean. Figure 2 shows an example of this classification for four methods (chart labeled 'Single-variable z-scores').

For the conditions with two visual elements combined we know, for example for the cases shown in Figure 2, how the parameterizations of icon lightness and density work in isolation, and we can see the design factor evaluations changed significantly when we combined both. On the other hand, saturation (red line in the charts) was not affected significantly by its combination with a different density parameterization (except for its dominance value).

5 Discussion

As we expected, no method dominates all factors for the single-variable case. For the linearity factor (not shown in the charts), icon orientation visually conveyed linearity very accurately, while almost all other methods failed to do so.

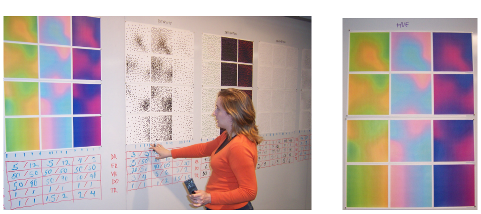


Figure 3: A subject in our pilot study critiques visualizations of 2D datasets with a single scalar variable. Illustration educators are shown a total of 132 visualizations corresponding to 3 parameterizations (columns, on the right) of 11 visual elements and 4 different datasets (rows, on the right). For each parameterization, they evaluate all 6 of our design factors (left, bottom).

Commenting on the appropriateness of our design factors, one of our subjects noted that a choice of visualization method will be affected by what the data actually is, e.g., visualizing temperatures is not the same as looking at wind speed or altitude data. In our case, we want to apply our resulting design knowledge to any type of scalar data, so we are considering the use of a seventh factor called *intuitive association*. This would measure whether there are any associative readings of a method that might interfere with the desired numerical reading and should be avoided.

For the two-variable examples, following Bertin's principles [1], icon density and lightness would have similar levels of organization so they would visually conflict with each other when combined, something that is obvious by observing the example above and which our pilot data confirmed. The non-conflict between saturation and icon density is also predicted by Bertin, due to their different levels of organization. The problem in this case is that the density results change drastically while the saturation ones remain mainly the same. One of our subjects explained that, although we are using the perceptually balanced LAB color space, creating saturation ranges that do not change their lightness is extremely difficult. Our eyes are very well trained to detect lightness changes, so this might be causing the density method to become unclear, like in the lightness/density example.

Note that our characterization of 2D visualization methods acknowledges that the input we get from the designers is directly targeted at the needs of scientists, and is not about artistic qualities, visual appeal, or aesthetics. Our subjects, illustration educators, are experts at evaluating visuals for targeted communication goals; while their results are often appealing and aesthetic, they first have to satisfy those communication goals which, in this case, means presenting scientific data for effective exploration.

6 Conclusion

With our current results, given requirements for all our design factors, we can probably find by hand, in the single-variable case, an optimal or close to optimal solution. But when multiple variables are involved, the optimization process will be much harder. This simple result already supports our research idea that a mathematical model, combined with a constrained optimization process, is necessary to find effective multivalued visualization methods. Our goal is to be able to build such a model from data collected in our designer critiques. A key for the success of this project is to gather as much information as possible from design experts in areas of the space that we can explore in a structured and exhaustive way, such as visualization methods with a single layer of visual elements representing a single data variable.

Also, obtaining information about visual element interactions at this point will facilitate our exploration of more complex areas of the space, by exploiting the knowledge gathered for this simple cases. The current results, although preliminary, provide an idea of the difficulty of the problem and the need for a formal study of this space of visualization methods, which we intend to pursue.

7 References

[1] Bertin.Semiology of Graphics. U. Wisconsin Press, 1983.
[2] Mackinlay.Automating the design of graphical presentations of relational information.ACM TOG,5(2),1986.
[3] Healey et al.Perceptually-based brush strokes for nonphotorealistic visualization.ACM TOG,23(1),2004.
[4] Tory et al.Human factors in visualization research.TVCG, 10(1),2004.
[5] Kosara et al.User studies:Why,how,and when.CG&A,23(4),2003.
[6] Jackson et al.Designer-critiqued comparison of 2D vector visualization methods: A pilot study.SIGGRAPH'03 Sketches.
[7] Laidlaw et al.Comparing 2D vector field visualization methods: A user study.TVCG,11(1),2005.

MINNEAPOLIS, MN USA
October 23-28, 2005

VIS 05

IEEE Visualization 2005

BEST POSTER AWARD

Presented To

Daniel Acevedo, Cullen Jackson, David Laidlaw and Fritz Drury

For the Poster

Using Visual Design Expertise to Characterize the Effectiveness of 2D Scientific Visualization Methods

From the IEEE Visualization Committee

Baoquan Chen, *Conference Co-Chair*

Kelly P. Gaither

Kelly P. Gaither, *Conference Co-Chair*

IEEE vgtc

Sponsored by the IEEE Computer Society
Visualization and Graphics Technical Committee (VGTC)

IEEE

IEEE Computer Society

Using Visual Design Expertise to Characterize the Effectiveness of 2D Scientific Visualization Methods

Daniel Acevedo, Cullen Jackson, and David Laidlaw, Computer Science, Brown University, Providence, RI
Fritz Drury, Illustration, Rhode Island School of Design, Providence, RI

VIS 05 MINNEAPOLIS, MN USA

Contributions

- An experimental methodology for capturing quantitative knowledge from visual design experts
- Measurement of a set of design factors to characterize visualization methods for representing scientific data
- Effectiveness quantification of individual visual elements for visualizing single-variable scalar datasets for exploratory visualization
- Hypothesis for using the measured design factors for single-variable visualizations to predict measurements for multivariate cases
- Description of a parameterized space of multilayered multivariate scientific visualization methods and a language to define it

Design Knowledge and Effectiveness Quantification

Six design factors characterize our visualization methods and form the basis for quantifying their effectiveness

Our design factors (see right) are based on concepts that our expert design educators can understand quickly and report on numerically. The feedback obtained is aimed at measuring the amount and quality of information transmitted by the visualizations, not at their esthetic value.

Expert illustration educators have experience in class critiques where they teach design and artistic concepts by evaluating how effectively a message is visually conveyed.

Requirements from scientists come in the form of goals for some or all of these factors. We quantify effectiveness based on how well a method fulfills the set of goals.

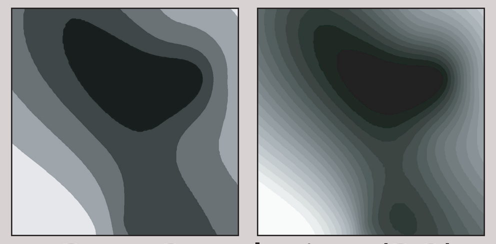
Research Objective

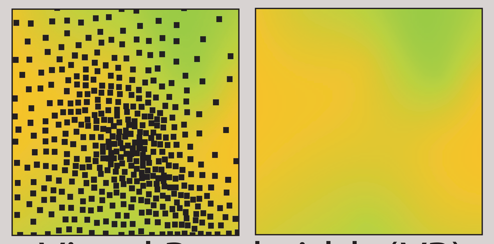
To find effective methods for multivariate multilayered scientific data representation optimized to facilitate perception and comprehension

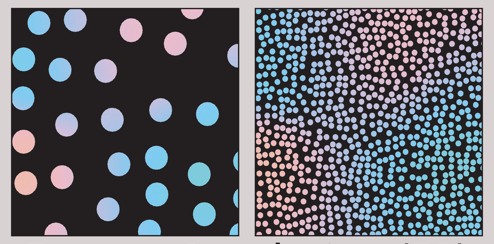
To achieve this objective we are developing a mathematical model of visual design knowledge based upon evaluations of visualization methods performed by expert visual-design educators. Our evaluation sessions, designed after "crit" sessions from art classes, create a familiar context in which expert designers can provide valuable quantitative information about the methods presented.

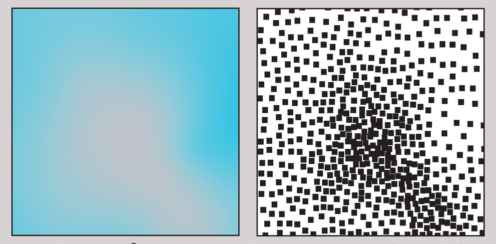
The first step has been to evaluate individual visual elements for representing single-variable datasets. After this, we will combine and evaluate these elements pair wise to represent two-variable datasets. Observing how perceptual interactions affect the evaluations, we will hypothesize the design factor ratings for three-variable visualizations and test a subset of them.

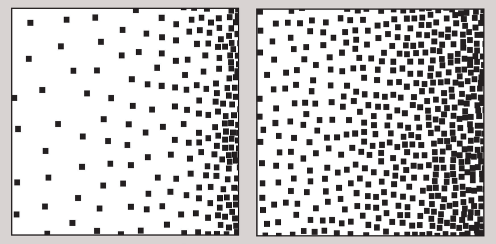
This process will allow us to build our mathematical model for more complex visualization methods, involving multiple icon and color-plane layers and representing multiple variables for exploratory scientific visualization.

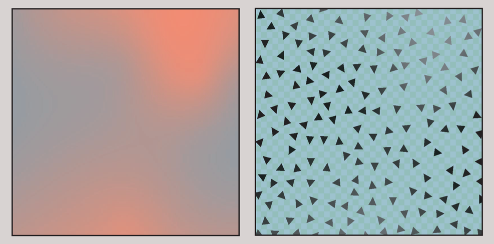

Data Resolution (DR)
levels of data a method can represent


Visual Bandwidth (VB)
% coverage a method can support and remain readable


Feature Resolution (FR)
size (% image width) of the spatial feature a method can represent


Visual Dominance (DO)
how much (0-10) a method dominates a composition over other methods


Visual Linearity (LI)
mark 0%, 25%, 50%, 75%, and 100%


Time to Read (TR)
seconds it takes to see the data

Visualization Language

Collaborating with expert designers, we have created a visualization language for multivariate visualization using basic visual attributes for multiple layers of icons and color-planes

Combinations of the elements we chose, shown below, generate an expressive space of visualization methods. The main elements found in the visualization literature are represented, so our results will be applicable to a large set of problems. The high dimensionality of the space poses a big challenge, but we are approaching the exploration of the space from the ground up, building up our knowledge from simple to more complex visualizations.

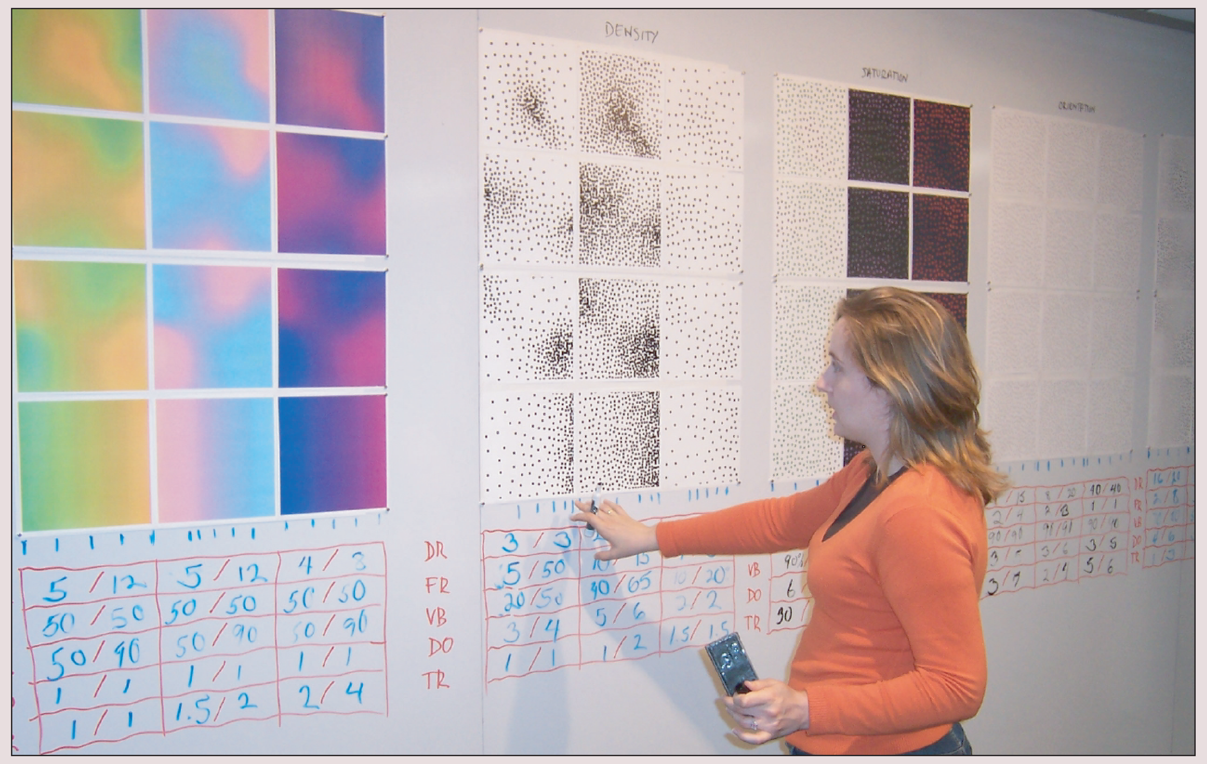
Experimental Setup

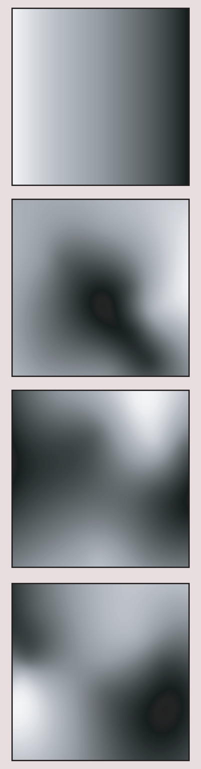
We present all stimuli simultaneously on paper so that subjects can easily compare them; this helps them provide consistent quantitative evaluations

Experienced educators can focus on evaluating a single method, but they explain their decisions more easily when they can compare across several examples. In our setup we present:

- 11 visual elements
- 3 different mapping ranges (transformation from data variable to visual variable) per element. See columns in the image below.
- 4 datasets (shown on the right) per mapping

We videotape the sessions, usually 3 to 4 hours long, and we encourage our subjects to explain their scoring criteria as they go. We hope to incorporate their comments as tie-breakers when conflicts arise in our model.

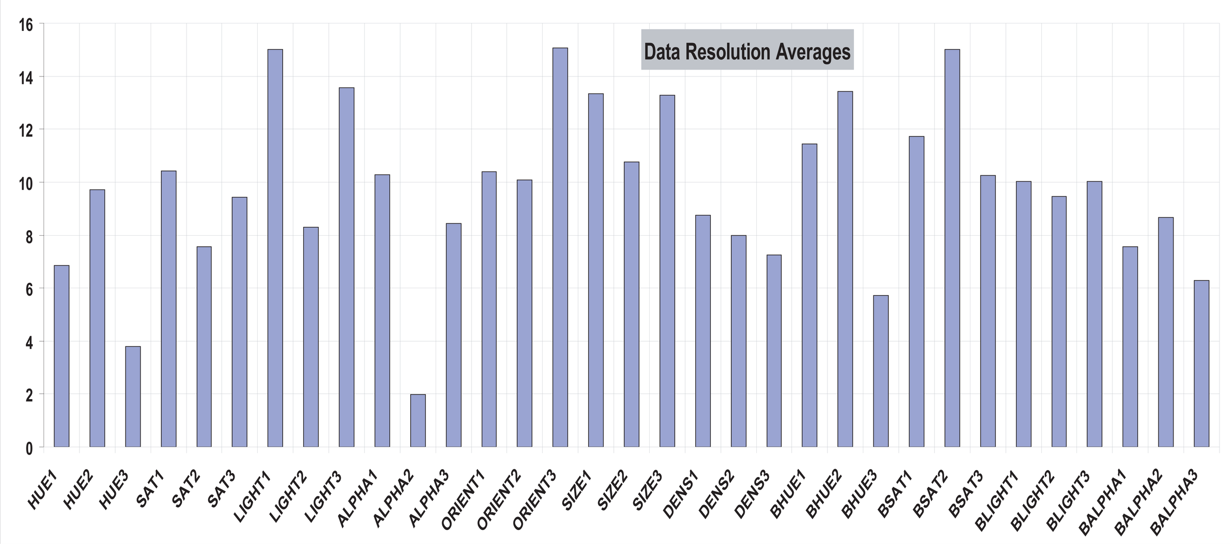


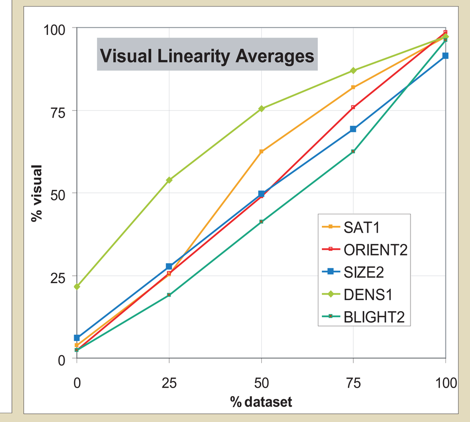


Results

We obtained a characterization of each of our eleven visual elements with respect to our six design factors along with valuable insights from our expert design educators

Five professors from the Illustration Dept. at RISD have evaluated our single-variable visualization examples. The graph on the left shows the results for the data resolution (DR) factor (# different levels of data a method is able to represent.) The results for all 33 methods are shown (11 elements x 3 mappings per element.) We have analogous characterizations for all the design factors. The graph on the right shows the results for the visual linearity factor for some of the methods.



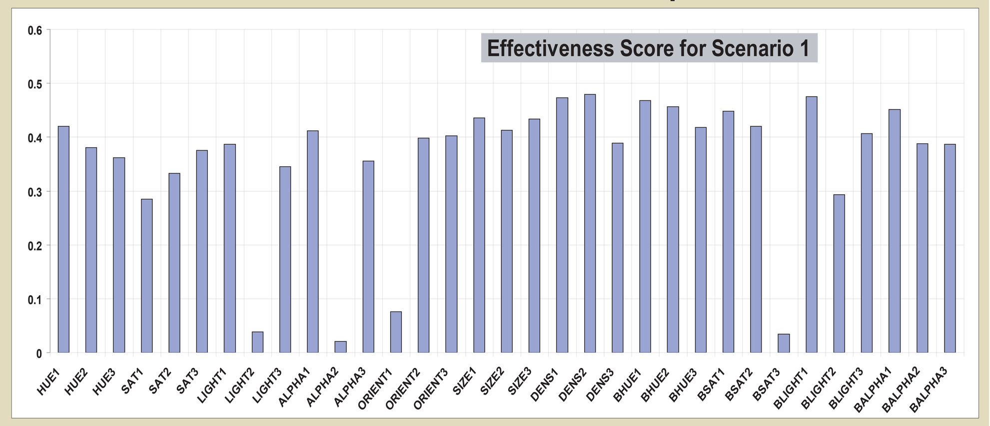


One advantage of using design experts to do the evaluations is that they can pinpoint reasons why a method does not work. They commented about the neutrality of the shapes used, the effects of negative space, the distinctive features a method like orientation creates that makes it readable, or technique of squinting their eyes to perceive the overall composition without focusing on single icons. We used these comments to refine our design factors.

With this characterization we devised five different scenarios, shown on the table below, setting goals for each design factor. These scenarios represent requirements expressed by scientists exploring the data. We measured the effectiveness of each method based on how it fulfills the desired set of goals. The graph below shows the results for one of these scenarios. Note that only three factors are constrained. For the other three factors, linearity can be fixed by adjusting the data-to-visuals mapping, while visual bandwidth and dominance are not required for single-variable examples.

Sample Design Goal Scenarios

	Data Resolution	Feature Resolution	Time to Read
Scenario 1	max	max	min
Scenario 2	DR<3	max	min
Scenario 3	4<DR<6	2<FR<5	min
Scenario 4	--	--	min
Scenario 5	4<DR<6	5<FR<10	--



What's next?

We are exploring a very high dimensional space using a very sparse sampling of visualization methods

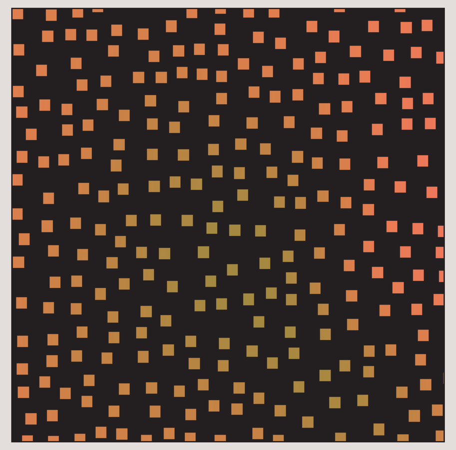
Based on these initial results we hypothesize that we will be able to predict and test design factor measurements for multi-valued visualization methods using combinations of our visual elements. Dominance and visual bandwidth, key design factors when layering visualizations, will become much more prominent in these more complex methods.

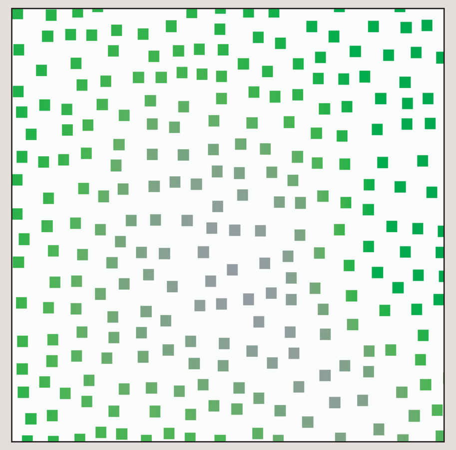
To evaluate this hypothesis we are beginning to bring our experts back to evaluate combinations of two visual elements representing two different scalar variables.

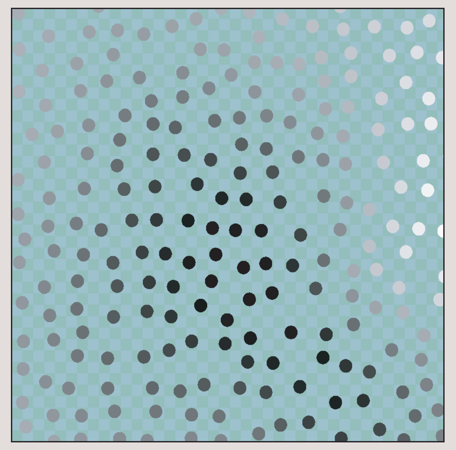
Using the design factors we will compare the evaluations with the results of the single-variable study. These combinations create very complex perceptual interactions (one of our subjects described trying to measure some of the factors as "visual heavy lifting"), and a direct analysis of the data must be coupled with a protocol analysis of the videotaped sessions.

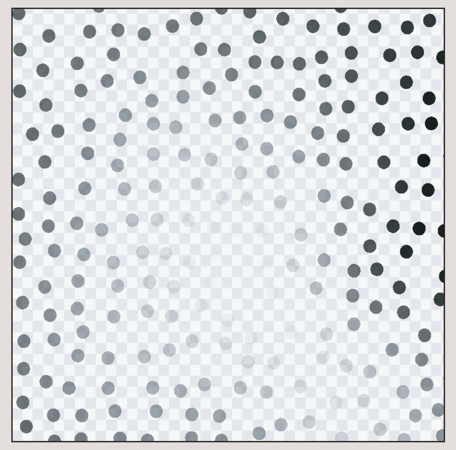
REFERENCES

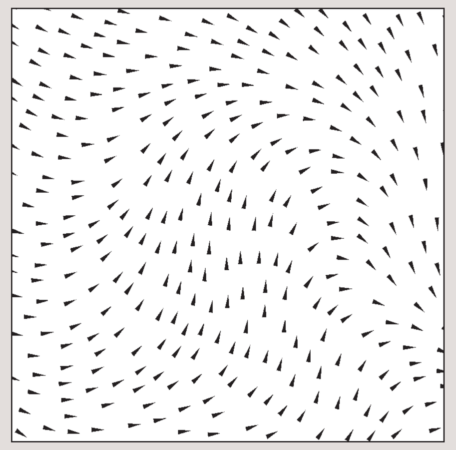
* J. Bertin. Semiology of Graphics. U. Wisconsin Press, 1983.
* J. Mackinlay. Automating the design of graphical presentations of relational information. ACM TOG, 5(2),1986.
* R. Kosara et al. User studies: Why,how,and when. CG&A, 23(4), 2003.
* C. Jackson et al. Designer-critiqued comparison of 2D vector visualization methods: A pilot study. SIGGRAPH'03 Sketches.
* C. Healey et al. Perceptually-based brush strokes for nonphotorealistic visualization. ACM TOG, 23(1), 2004.
* M. Tory et al. Human factors in visualization research. TVCG, 10(1), 2004.
* D. Laidlaw et al. Comparing 2D vector field visualization methods: A user study. TVCG, 11(1), 2005.

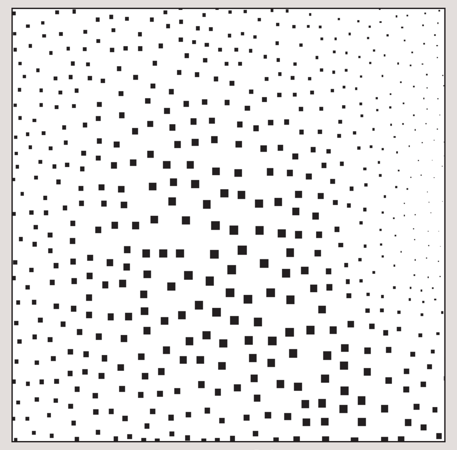

Icon Hue

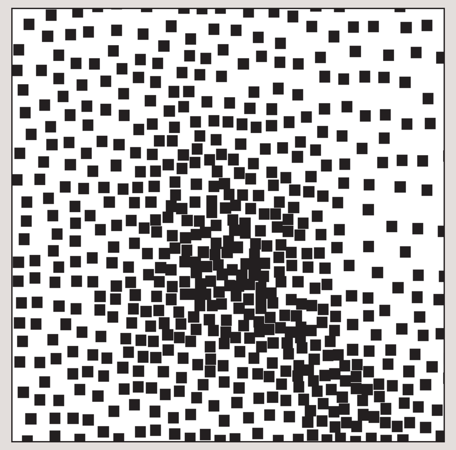

Icon Saturation


Icon Lightness


Icon Transparency

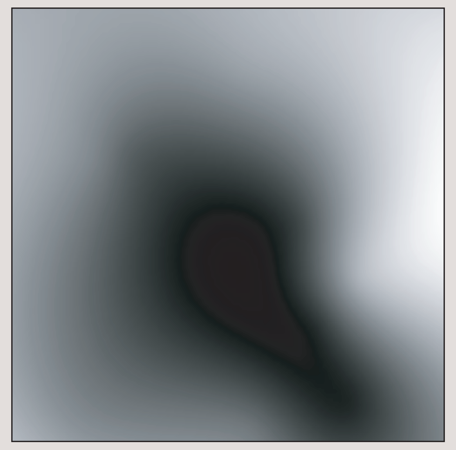

Icon Orientation


Icon Size


Icon Density


Background Hue


Background Saturation


Background Lightness


Background Transparency