

## AAPOR, Script 2

### Slide 1:

To start, I'm going to read from Johanna Drucker's great book on visual forms of knowledge production. The book is cited later in the slides.

*The rendering of statistical information into graphical form gives it a simplicity and legibility that hides every aspect of the original interpretive framework on which the statistical data were constructed. The graphical force conceals what the statistician knows very well—that no “data” pre-exist their parameterization.*

### Slide 2:

I'm Aaron Hill, I'm an assistant professor at Parsons School of Design in New York and I also am the director of the Master of Science program in data visualization. I co-taught the course with Michael that you just heard about. I want to recognize my collaborator, Clare Churchouse, who is here today. Our work on this started in the capacity of faculty member and research assistant, but quite quickly Clare became a true collaborator in every sense of the word, and this work is much richer because of her involvement.

Today I am presenting some of our work on new ways of representing uncertainty in data visualization. More specifically, we're looking to the fine arts for better ways to visually express in graphs the uncertainty and ambiguity that exists in the underlying data.

### Slide 3:

Our data consists of photographs of approximately 400 works of fine art (showing a small subset here). They have in common one thing: we believe that in some way, they do a really good job of visually expressing uncertainty.

### Slide 4:

To give you a sense of the kinds of art that are represented in the data, here's a table of the frequency of the medium of each piece. About a third of the works are paintings, and then about an equal number of photographs, drawings, sculpture, prints, and installations. Although we believe that one of the most promising ways of expressing uncertainty is through animation and motion, the focus here is on what is currently one of the main forms of data display: still images. There are however three videos or films represented in the data, and of note here is that our focus is on the visual attributes of this work, which applies across mediums, whether still or moving.

Additionally, the work comes from a variety of museums and collections. A great many of them are held by the Museum of Modern Art in New York and the Tate Modern in London. But they come from other museums and private collections as well. Most of the work is modern and contemporary.

### Slide 5:

We needed a framework for encoding and understanding the qualities of the visual elements of these works of art, so we could systematically compare, aggregate, summarize, and search for meaning in these images. There have been many attempts at creating a universal language or grammar of graphics. We decided on Jacques Bertin's model in his 1967 book *The Semiology of Graphics* (cited later in the slides).

Bertin defines seven visual variables, which I'll get into in more detail in a moment. They are: placement (called position by Bertin), value, texture, color, orientation, shape, and size.

The header of this slide shows one way you can use the visual variable of placement to indicate uncertainty. There are points around the outline of the font for the word "uncertainty." But you'll see that it is incomplete and the points are moving. Multiple random samples of the points were taken and so we only see a subset of the full set at any given time. Points move to their next position as the sample changes. This is one way placement can show the incompleteness and intuitively express complex concepts such as sampling, coverage, and distributions.

**Slide 6:**

Bertin also defines three elemental marks: points, lines, and areas. We see here treatments on those three marks, by six of the visual variables for various ways you can graphically express the data. It's important for me to note here that this system was not developed explicitly to show uncertainty, rather, it's much broader than that. It's also important to note that this was created as a system for graphics, not for art.

**Slide 7:**

But with our data, we did use this system to help us define the visual attributes of the works of art. Our first finding is that this system did really well for this purpose. We found that marks on the various canvases and media could be classified as a point, line, or area, without being too reductive. We also found that the seven visual variables were fairly exhaustive and excellent at defining the visual attributes of modern and contemporary art, also without being too reductive. This framework was enormously helpful for encoding the attributes of the 400 pieces of art in our data.

**Slide 8:**

So we comprehensively applied Bertin's model of graphics on this art. With the 400 images, we classified each image on a 0-5 scale for each attribute, with 5 representing a strong presence of that attribute, to 0 for the absence of an attribute. We also added a few other attributes we felt important, such as whether it contained text, is it kinetic, does it have a spatial quality, etc. It's important to note here that the marks and visual variables are not mutually exclusive. In most of these pieces, there is a combination of marks and visual techniques.

*6 min*

**Slide 9 (a blank slide):**

Now I'm going to show you images of about 20 of the works of art, chosen because they represent a diverse set of mediums, techniques, and use of visual variables. After, I'll summarize that full set of data and summarize the conclusions we drew. The full citations for the works of art are included later in the slides and we'll provide a URL to access the slides.

**Slide 10:**

The images I'm showing are organized by primary visual variable. The first visual variable is placement, which is the chosen location of a mark in any given coordinate space. We'll look at some images where we feel placement elicits a sense of uncertainty.

**Slide 11:**

This drawing uses perforation to make the marks. The holes in the paper are subtly aggregated in one area, with hesitancy created especially by judicious perforations at the far right. It certainly looks like--but is not explicitly stated to be--a scatterplot. The vertical axis is not a straight line, and the full distribution of points is not centered.

**Slide 12:**

This drawing, which is graphite of cut-and-pasted Japanese paper atop other ordinary paper, shows several rectangles. But the borders of the areas are not clear. Where does one begin and the other end? And the semi-transparent nature of the paper means that the overlap creates various color values. Added to this are strategically placed arrows, numbers and notations close to the edges pointing in different directions.

**Slide 13:**

This is a photograph, and we're looking at an upside-down reflection off of a puddle of water, which had collected after a heavy rainfall. The rocks you see are actually in the puddle. But even knowing that, it's hard to grasp exactly what's going on here. It's a great use of inverted perspective to call into question what is real.

**Slide 14:**

This photograph with acrylic paint additions is surreal and has many interpretations and meanings. The use of two very and obviously different things and scales makes us wonder how this happened and what will happen next.

**Slide 15:**

The next images have in common the use of shape in a way we believe expresses uncertainty.

**Slide 16:**

This installation, which is wood and paint, reads as an incomplete square. The lines, which are wood, are not quite forming a right angle. And it makes one wonder is it open or closed? And if it is a square, where is the rest of it? The absence and incompleteness creates a sense of visual openness and possibility.

**Slides 17:**

This Dali painting or "hand painted dream photograph," bends physical objects that don't ordinarily bend and behave in this way. It looks real, as though it were a photograph. But because the laws of physics seem to have been broken, we know that it isn't real.

**Slide 18:**

This installation of wood, metal, plastic, ceramic, paper, textile, and wire we chosen to show for the shadows that are projected onto the walls, by some interior light. We can recognize the shapes of the shadows and what they represent, but we also know that they are only representational, and not the real thing itself.

**Slide 19:**

This is a drawing that uses gunpowder with two sheets of paper. It was created by exploding the gunpowder sandwiched between the two sheets of paper, and we see now the unfolded result. The

chemical transformation changed the original form and shape, and we see the residue. We have evidence that what we're seeing isn't what we started with.

*10 min (4 min)*

**Slide 20:**

The next images use texture and blur.

**Slide 21:**

This print shows overlapping, scratched, and textural lines. In the aggregate, it also creates a sense of mass in some areas, but the boundaries are unclear.

**Slide 22:**

This painting of stenciled letters, repeated over and over, becomes more smudged and darker toward the bottom. The handmade physical quality becomes more and more exaggerated, and it becomes harder to distinguish the letters and words.

**Slide 23:**

This oil painting is made of layered paint, squeegeed to literally smear the paint colors and shapes into each other. The effect is a strong sense of ambiguity.

**Slide 24:**

This is a digital print with acrylic on a canvas. The actual piece is gigantic, and up close, you can only see pixels. But from afar, you can begin to make out that this photograph is some street corner, in some city. But it would be nearly impossible to say for sure where it is.

**Slide 25:**

The next images make use of value, which is the various degrees between white and full saturation.

**Slide 26:**

This photograph appears to have an effect created by overexposure. We see not only a source of light, but where it has been. Each time the light pauses, we see a point. And we see the path it took between points.

**Slide 27:**

This painting with oil and glue has parallel lines that fade out in the bottom of the painting. We can still see the full lines, but it becomes harder to see as the blue fades.

**Slide 28:**

This drawing is pastel on paper. The area that is defined by the deep, dark, rich color clearly pushes our attention to the center. But the edges are faded and the delineation is somewhat ambiguous.

**Slide 29:**

In this drawing with charcoal and chalk, you can barely see an apparition in the window. It's clear enough that we can see it, but faded enough that we are reminded that it is probably

not of this physical world.

**Slide 30:**

And now an image where color expresses uncertainty.

**Slide 31:**

This print has strange color combinations whose purpose is not exactly clear. The print in general is ambiguous, but most would agree that it shows some kind of architecture or infrastructure. But the color doesn't seem to communicate anything that would be widely recognized.

**Slide 32:**

Now an image where size conveys uncertainty.

**Slide 33:**

This painting uses size in a very literal way with other variables to create a sense of uncertainty. The units of time are listed in order from smallest to largest. But the typeface size changes, moving from largest to smallest.

**Slide 34:**

And finally, a couple of images where orientation expresses uncertainty.

**Slide 35:**

This watercolor painting shows elements at a very steep angle to the plane, giving the appearance of slipping down off the surface. But the trees are painted parallel to the edges of the canvas.

**Slide 36:**

This installation is a room, a physical space. We can recognize that it's a room, but the space is unusual and disorienting. We can see that this is not an ordinary room, and perhaps not even a real room.

**Slide 37:**

You may be thinking, we'll that's all nice, but how would I do that in R or SAS or Stata or Python? And while we may not be able to exactly, and may not even want to exactly, the hope is that art can become a reference for more subtle visual cues and treatments that could alter our demarcations, delineations, and containers. So that they read as less certain, but not to the point that they read as unbelievable.

*15 min*

**Slide 38:**

As I mentioned earlier, we carefully encoded the attributes of each of the images. Here I'm showing the averages of the visual variable encodings, which were measured on a scale from 0 to 5, with 5 being the strongest expression of a given attribute. We found as I previously mentioned that the artworks used multiple variables. We found that the two most prevalent visual variables in the work are placement and shape. Texture, value, and orientation were

also strongly represented in the data. Size and color were the least frequently used visual variables in expressing uncertainty.

**Slide 39:**

A Google image search of “graphs of uncertainty,” show that error bars are by far the predominant way to express uncertainty in data. We’re not suggesting that should change, as error bars are widely recognized and almost as widely understood.

*16 minutes*

**Slide 40:**

But there’s also more to uncertainty. Another Drucker quote: *The profound challenge we face is to accept the ambiguity of knowledge, the fundamentally interpreted condition on which data is constructed.* And maybe that is where art can be most helpful in expressing uncertainty. Maybe not for the precisely measured known sources of uncertainty in error, but for subtle prompts and reminders that all data represent an abstraction and not a hard truth.

**Slide 41:**

The next steps for this work: we are creating an interface to a searchable database of these images of works of art. It’s not quite ready for prime time, but you can see here a preview of this tool, which allows the user to find and select images based on various visual variables and attributes. Our hope is that this tool will be helpful to designers and analysts in seeking sources of inspiration for new ways to express uncertainty in graphs.

**Slide 42:**

If you’d like to be notified of the launch of this tool, our contact information is listed here. Let us know, and we’ll be sure to notify you when it’s released. There’s also a URL at the bottom that leads you to a PDF of these slides in case you’d like to access them. It includes references to the books I mentioned, as well as citations for the works of art included in this presentation, with links to the sources. Thank you again to my collaborator, Clare Churchouse. And thank you for your time.

*17-18 minutes*