Information Visualization
CMSC 838B – Spring 2003

Visual Design

Benjamin B. Bederson
University of Maryland
www.cs.umd.edu/~bederson

This presentation adapted from John Stasko

Semiotics

- The study of symbols and how they convey meaning
- Classic book:
  - J. Bertin, 1983, *The Semiology of Graphics*

This presentation taken largely from John Stasko

Related Disciplines

- Psychophysics
  - Applying methods of physics to measuring human perceptual systems
    - How fast must light flicker until we perceive it as constant?
    - What change in brightness can we perceive?
- Cognitive psychology
  - Understanding how people think, here, how it relates to perception

One (simple) Model

- Two stage process
  - Parallel extraction of low-level properties of scene
  - Sequential goal-directed processing

Stage 1

Early, parallel detection of color, texture, shape, spatial attributes

Stage 2

Serial processing of object identification (using memory) and spatial layout, action

[Ware 2000]

Stage 1 - Low-level, Parallel

- Neurons in eye & brain responsible for different kinds of information
  - Orientation, color, texture, movement, etc.
- Arrays of neurons work in parallel
- Occurs “automatically”
- Rapid
- Information is transitory, briefly held in iconic store
- Bottom-up data-driven model of processing
- Often called “pre-attentive” processing

Stage 2 - Sequential, Goal-Directed

- Slow serial processing
- Involves working and long-term memory
- More emphasis on arbitrary aspects of symbols
- Top-down processing

Preattentive Processing

- How does human visual system analyze images?
  - Some things are done preattentively, without the need for focused attention
  - Generally less than 200-250 msecs (eye movements take 200 msecs)

How Many 3's?

128176875613897654698450698560498282676 2
980985845822450985645894509845098035858
909103020990595959577256467505067890456 7
884578980982167765487636490856091294968 6

How Many 3's?

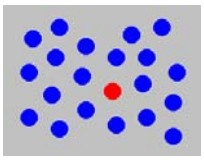
128176875613897654698450698560498282676 2
980985845822450985645894509845098035858
909103020990595959577256467505067890456 7
884578980982167765487636490856091294968 6

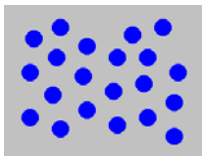
What Kinds of Tasks?

- Target detection
  - Is something there?
- Boundary detection
  - Can the elements be grouped?
- Counting
  - How many elements of a certain type are present?

Hue

Determine if a red circle is present

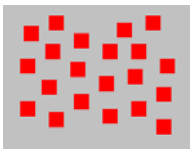


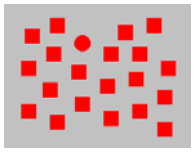


Can be done rapidly (preattentively) by people
Surrounding objects called "distractors"

Shape

Determine if a red circle is present

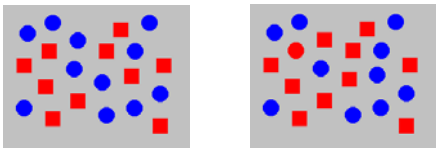




Can be done preattentively by people

Hue and Shape

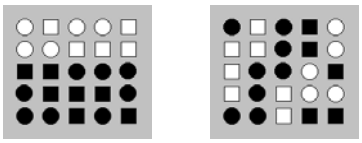
Determine if a red circle is present



- Cannot be done preattentively
- Must perform a sequential search
- Conjuction of features (shape and hue) causes it

Fill and Shape

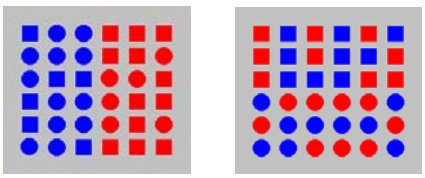
Is there a boundary?



- Left can be done preattentively since each group contains one unique feature
- Right cannot (there is a boundary!) since the two features are mixed (fill and shape)

Hue versus Shape

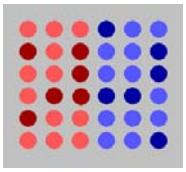
Is there a boundary?

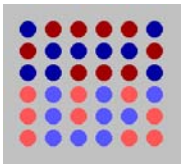


Left: Boundary detected preattentively based on hue regardless of shape
Right: Cannot do mixed color shapes preattentively

Hue versus brightness

Is there a boundary?





Left: Varying brightness seems to interfere
Right: Boundary based on brightness can be done preattentively

Example Applet

- Nice on-line tutorial and example applet
  - http://www.csc.ncsu.edu/faculty/healey/PP/PP.html
  - Chris Healey, NC State

Potential Preattentive Features

length	hue
width	intensity
size	flicker
curvature	direction of motion
number	binocular lustre
terminators	stereoscopic depth
intersection	3-D depth cues
closure	lighting direction

Key Perceptual Properties

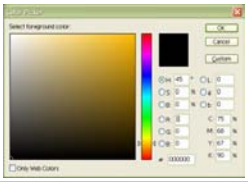
- Brightness
- Color
- Texture
- Shape

Luminance/Brightness

- Luminance
  - Measured amount of light coming from some place
- Brightness
  - *Perceived* amount of light coming from source

Color Models

- RGB
- HSB (HVS) model
  - Hue - what people think of color
  - Saturation - intensity, whiteness
  - Brightness (Value) - light/dark



Contrast

- Important for fg-bg colors to differ in brightness

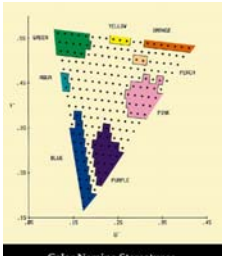
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?

Color for Categories

- Can different colors be used for categories of nominal variables?
  - Yes
  - Ware's suggestion: 12 colors
    - red, green, yellow, blue, black, white, pink, cyan, gray, orange, brown, purple

Color Categories

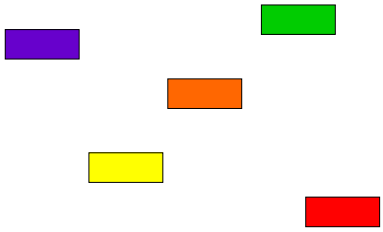
- Are there certain canonical colors?
  - Post & Greene '86 had people name different colors on a monitor
  - Pictured are ones with > 75% commonality



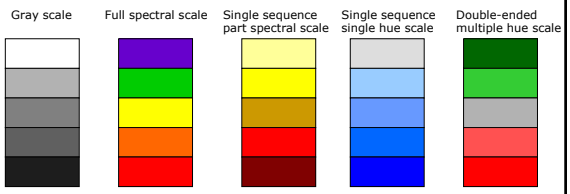
Color Naming Stereotypes

Color for Sequences

Can you order these (low->hi)



Possible Color Sequences

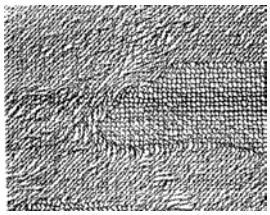


Color Purposes

- Call attention to specific data
- Increase appeal, memorability
- Increase number of dimensions for encoding data
  - Example, Ware and Beatty '88
    - x,y - variables 1 & 2
    - amount of r,g,b - variables 3, 4, & 5

Texture

- Appears to be combination of
  - orientation
  - scale
  - contrast
- Complex attribute to analyze



Shape, Symbol

- Symbols should be rapidly perceived and differentiated
- Application for maps, military, etc.

Basic Symbolic Displays

- Graphs
- Charts
- Maps
- Diagrams

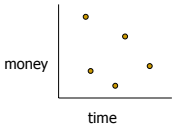
From:
S. Kosslyn, "Understanding charts and graphs", *Applied Cognitive Psychology*, 1989.

1. Graph

- Graph - Show the relationships between variables' values in a data table
  - Visual display that illustrates one or more relationships among entities
  - Shorthand way to present information
  - Allows a trend, pattern or comparison to be easily comprehended

Issues

- Critical to remain task-centric
  - Why do you need a graph?
  - What questions are being answered?
  - What data is needed to answer those questions?
  - Who is the audience?



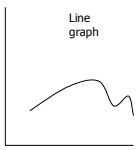
Graph Components

- Framework
  - Measurement types, scale
- Content
  - Marks, lines, points
- Labels
  - Title, axes, ticks

Basic Data Types

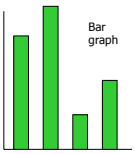
- Nominal (qualitative)
  - (no inherent order)
  - city names, types of diseases, ...
- Ordinal (qualitative)
  - (ordered, but not at measurable intervals)
  - first, second, third, …
  - cold, warm, hot
- Interval (quantitative)
  - list of numbers

Common Graph Formats



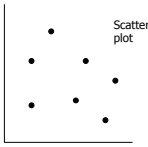
Y-axis is quantitative variable

See changes over consecutive values



Y-axis is quantitative variable

Compare relative point values



Two variables, want to see relationship

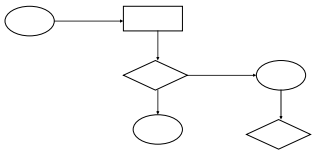
Is there a linear, curved or random pattern?

Graphing Guidelines

- Independent vs. dependent variables
  - Put independent on x-axis
  - See resultant dependent variables along y-axis
- If there are two independent variables, often place them along the 2 axes (you choose which) and then the mark may encode the dependent variable

2. Chart

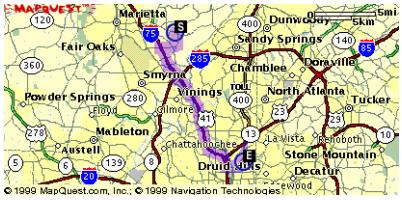
- Structure is important, relates entities to each other
- Primarily uses lines, enclosure, position to link entities



Examples: flowchart, family tree, org chart, ...

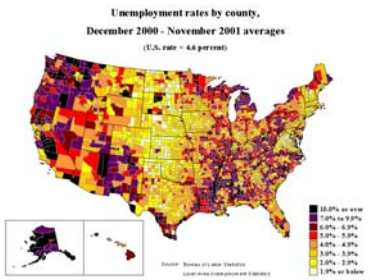
3. Map

- Representation of spatial relations
- Locations identified by labels



Choropleth Map

Areas are filled and colored differently to indicate some attribute of that region

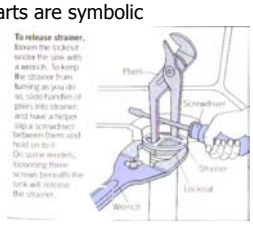


Cartography

- Cartographers and map-makers have a wealth of knowledge about the design and creation of visual information artifacts
  - Labeling, color, layout, …
- Information visualization researchers should learn from this older, existing area

4. Diagram

- Schematic picture of object or entity
- Parts are symbolic



Examples: figures, steps in a manual, illustrations,...

Tufte's Design Principles

- 1. Tell the truth
  - Graphical integrity
- 2. Do it effectively with clarity, precision…
  - Design aesthetics

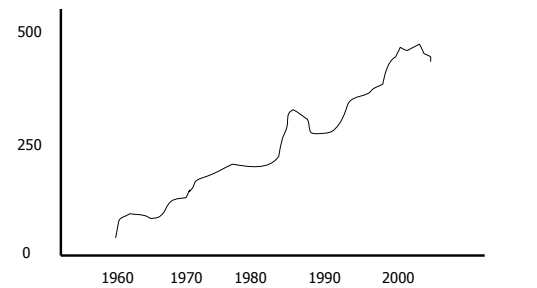
E. Tufte, *The Visual Display of Quantitative Information* (1983)
E. Tufte, *Envisioning Information* (1990)
E. Tufte, *Visual Explanations* (1997)

1. Graphical Integrity

- Your graphic should tell the truth about your data

Show entire scale

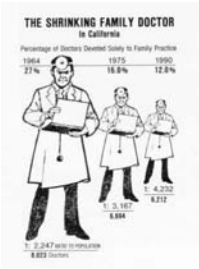
Show in context



Measuring Misrepresentation

"Lie factor" = 2.8

- Visual attribute value should be directly proportional to data attribute value
- Height/width vs. area vs. volume



$$\text{Lie factor} = \frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}}$$

2. Design Principles

- Maximize data-ink ratio

$$\text{Data ink ratio} = \frac{\text{Data ink}}{\text{Total ink used in graphic}}$$

= proportion of graphic's ink devoted to the non-redundant display of data-information

Design Principles

- Avoid chartjunk
  - Extraneous visual elements that detract from message

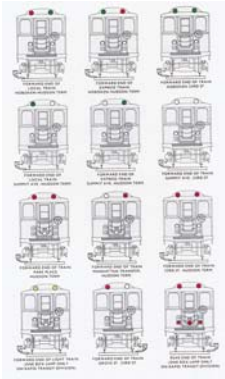
don't be the duck of architecture

Design Principles

- Utilize multifunctioning graphical elements
  - Graphical elements that convey data information and a design function

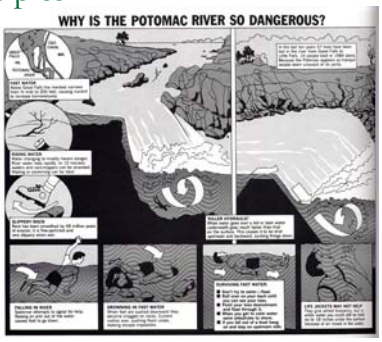
Design Principles

- Use small multiples
  - Repeat visually similar graphical elements nearby rather than spreading far apart



Design Principles

- Show mechanism, process, dynamics, and causality
  - Cause and effect are key

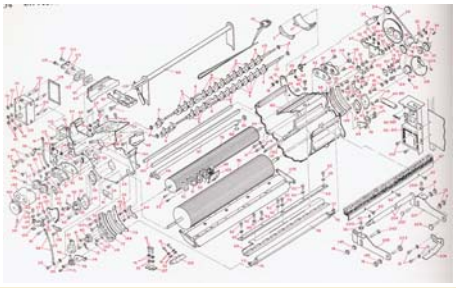


Design Principles

- Escape flatland
  - Data is multivariate
  - Doesn't necessarily mean 3D projection

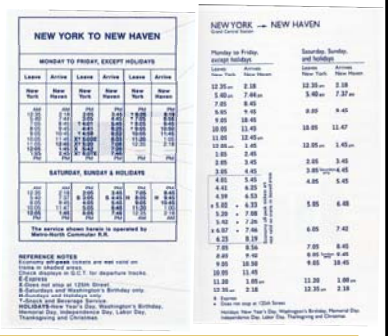
Design Principles

- Utilize layering and separation



Design Principles

- Utilize narratives of space and time
  - Tell a story of position and chronology through visual elements



Design Principles

- Content is king
  - Quality, relevance and integrity of the content is fundamental
  - What's the analysis task? Make the visual design reflect that
  - Integrate text, chart, graphic, map into a coherent narrative

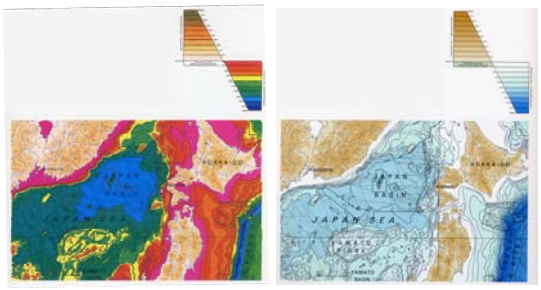
Graph and Chart Tips

- Avoid separate legends and keys – Put that in the graphic
- Make grids & labeling faint so that they recede into background

Proper Color Use

- To label
- To measure
- To represent or imitate reality
- To enliven or decorate

Color Examples



Guides for Enhancing Visual Quality

- Attractive displays of statistical info
  - have a properly chosen format and design
  - use words, numbers and drawing together
  - reflect a balance, a proportion, a sense of relevant scale
  - display an accessible complexity of detail
  - often have a narrative quality, a story to tell about the data
  - are drawn in a professional manner, with the technical details of production done with care
  - avoid content-free decoration, including chartjunk

Graphical Displays Should

- Show the data
- Induce the viewer to think about substance rather than about methodology, graphic design the technology of graphic production, or something else
- Avoid distorting what the data have to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several levels of detail, from a broad overview to the fine structure
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- Be closely integrated with statistical and verbal descriptions of a data set