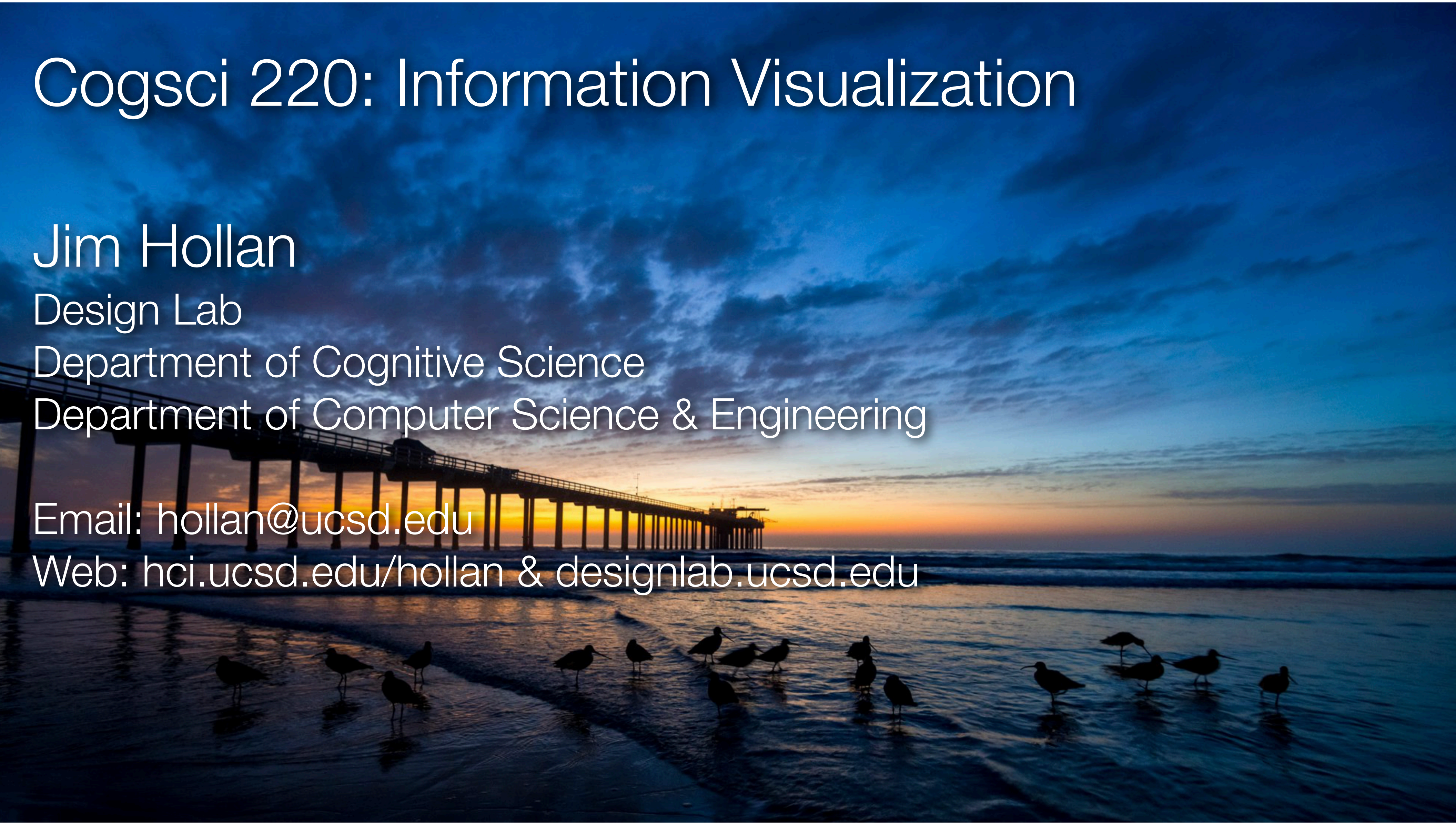


Cogsci 220: Information Visualization

Jim Hollan
Design Lab
Department of Cognitive Science
Department of Computer Science & Engineering

Email: hollan@ucsd.edu
Web: hci.ucsd.edu/hollan & designlab.ucsd.edu

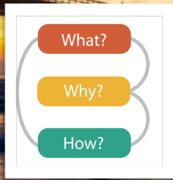
Cogsci 220: Information Visualization

Welcome to the third week. Hope everyone is well and safe in these challenging times.

Last week I discussed the value of visualization with the goal of convincing you it is crucially important and we tried using breakout rooms for you to start to get to know your classmates.

Today:

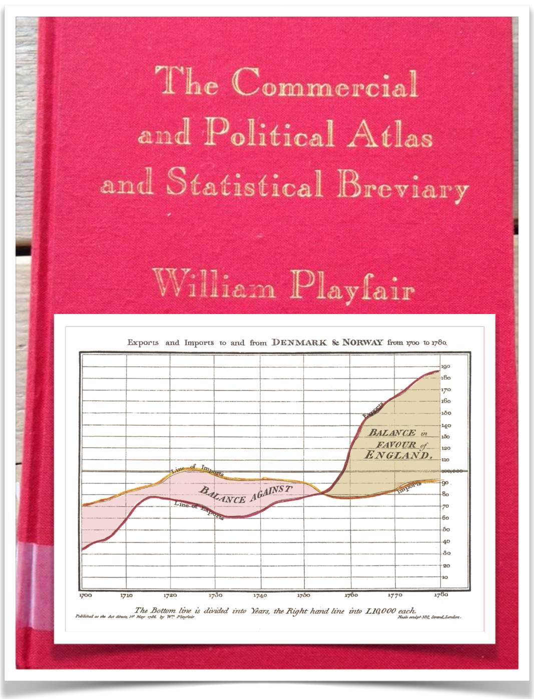
- Researchers of the week: John Tukey and Stu Card
- Fundamental importance of activities and tasks in designing visualizations
- What, why, and how of visualization

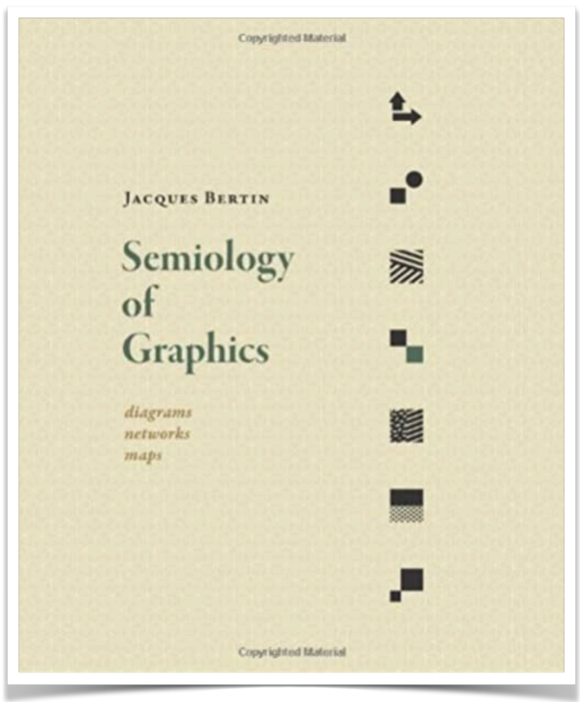


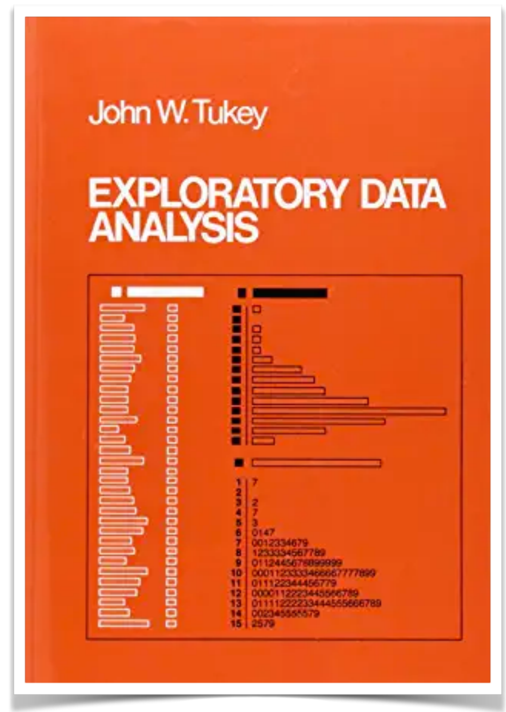
- If time your reactions to and comments on chapter on *Information Visualization* and *Cognitive Design of Tools of Thought* but want to begin two iterations of topic group breakouts starting at ~10

Information Visualization

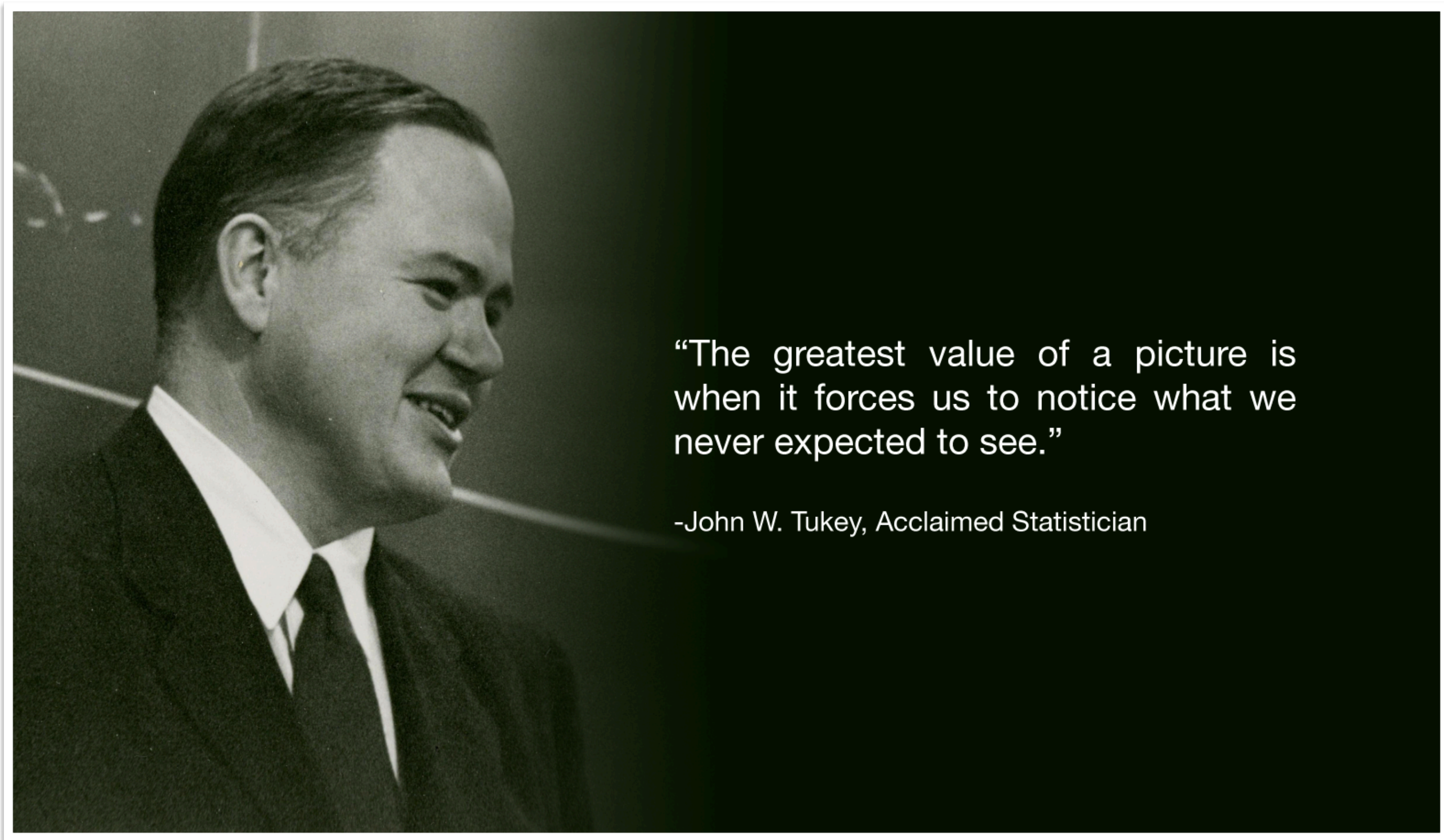
Early seminal work: Playfield (1786), Bertin (1963), Tukey (1977)



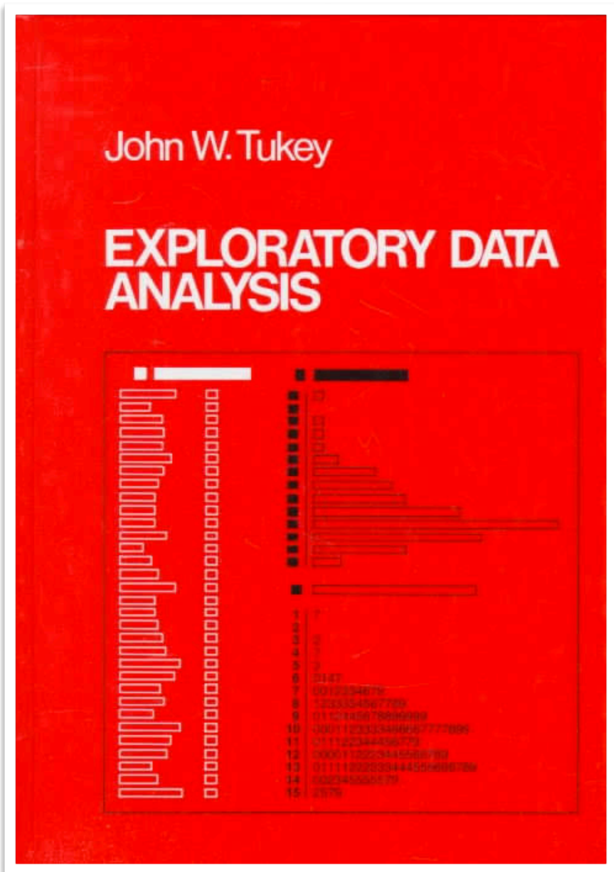




John Tukey


"The greatest value of a picture is when it forces us to notice what we never expected to see."
-John W. Tukey, Acclaimed Statistician

John Tukey



This book is based on an important principle:

It is important to understand what you CAN DO before you learn to measure how WELL you seem to have DONE it.

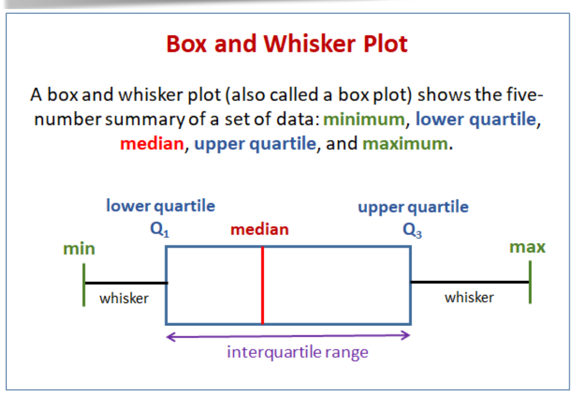
Learning first what you can do will help you to work more easily and effectively.

This book is about exploratory data analysis, about looking at data to see what it seems to say. It concentrates on simple arithmetic and easy-to-draw pictures. It regards whatever appearances we have recognized as partial descriptions, and tries to look beneath them for new insights. Its concern is with appearance, not with confirmation.

A basic problem about any body of data is to make it more easily and effectively handleable by minds—our minds, her mind, his mind. To this general end:

◇ anything that makes a simpler description possible makes the description more easily handleable.

◇ anything that looks below the previously described surface makes the description more effective.

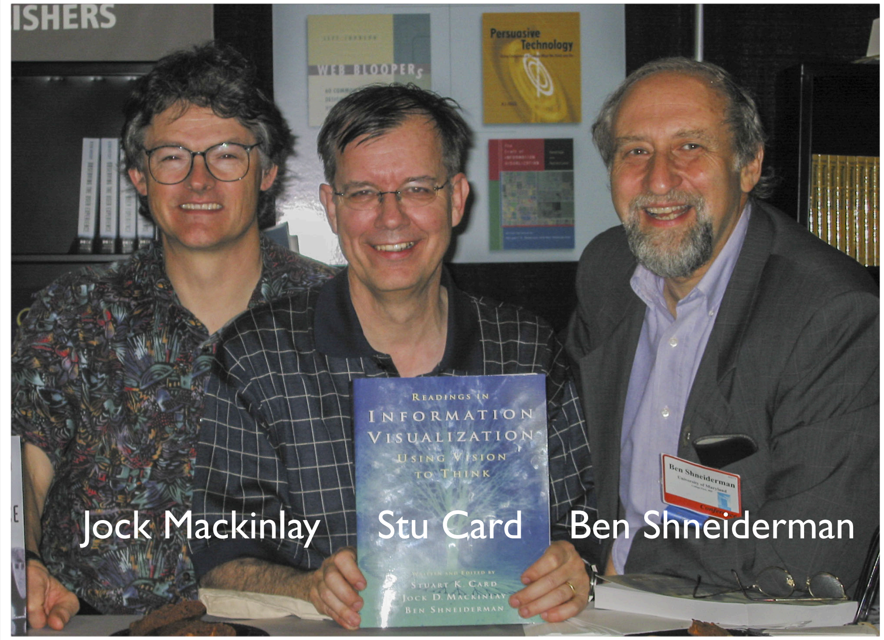


It's better to solve the right problem approximately than to solve the wrong problem exactly.

I know of no person or group that is taking nearly adequate advantage of the graphical potentialities of the computer.


Tukey Prim9 at SLAC

Stu Card



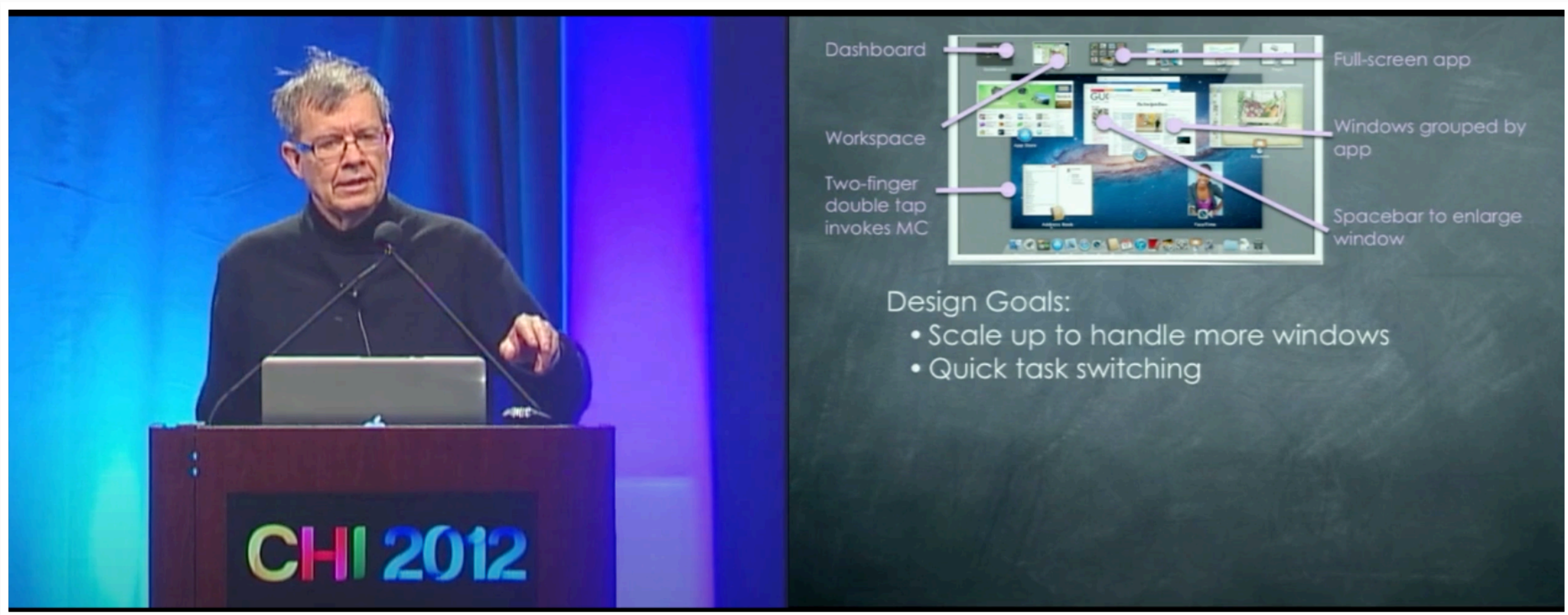
Google Scholar h-index: 91

51,505 Citations

Stuart K. Card is one of the pioneers of applying human factors in human–computer interaction. His study of input devices led to the Fitts's Law characterization of the computer mouse and was a major factor leading to the mouse's commercial introduction by Xerox, most notably in the Alto and Star projects, some of the very earliest GUI systems employing a desktop metaphor. The 1983 book *The Psychology of Human-Computer Interaction*, which he co-wrote with Thomas P. Moran and Allen Newell, became seminal work in the HCI field. Further research into the theoretical characterizations of human–machine interaction led to developments like "the Model Human Processor, the GOMS theory of user interaction, information foraging theory, and statistical descriptions of Internet use”. His work on information visualization and the “information visualizer” helped create the field.

Stu Card: Interaction Science

https://www.youtube.com/watch?v=3w12AbzHw8E SIGCHI Lifetime Research Award



Information Visualizer

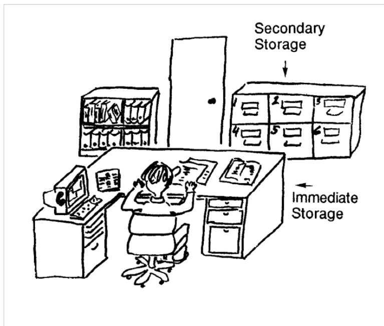


Fig. 1. An office organized to have an efficient information cost structure.

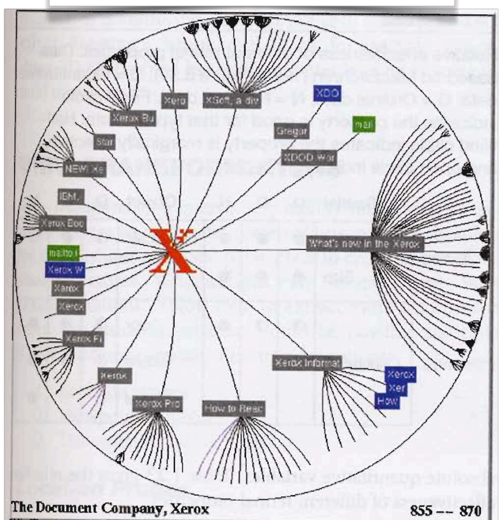


FIGURE **1.36**

Hyperbolic tree. See Lamping and Rao (1996●). Courtesy of Xerox Corporation.

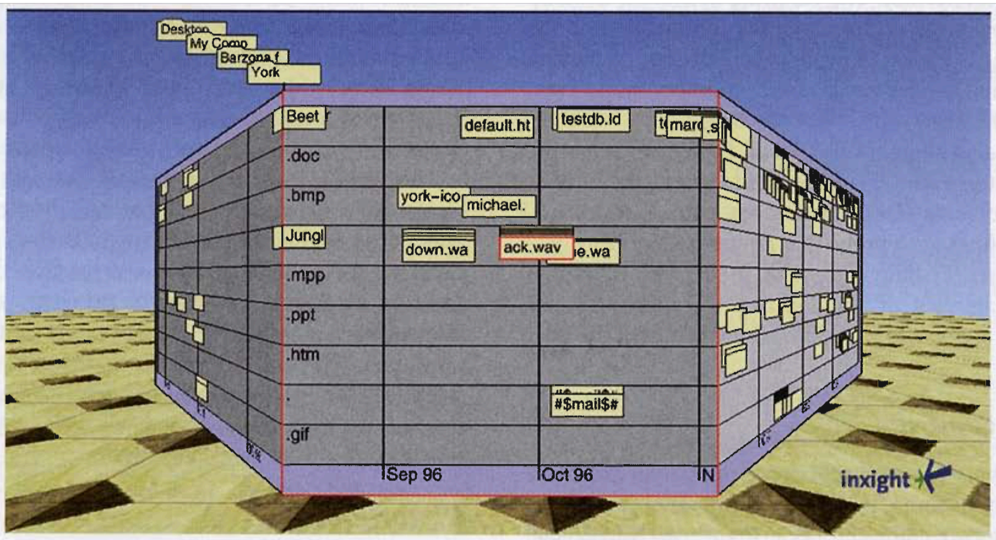
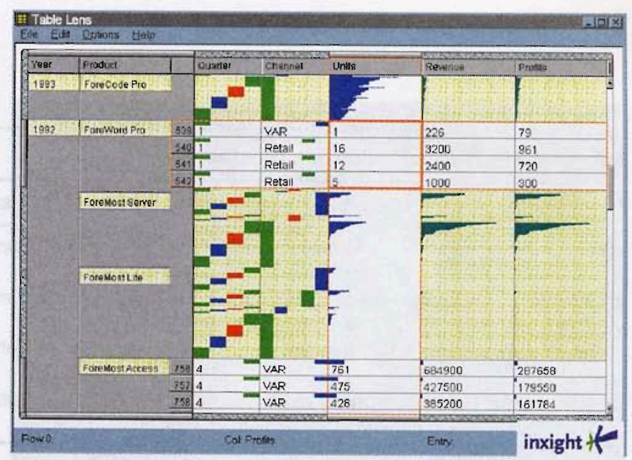
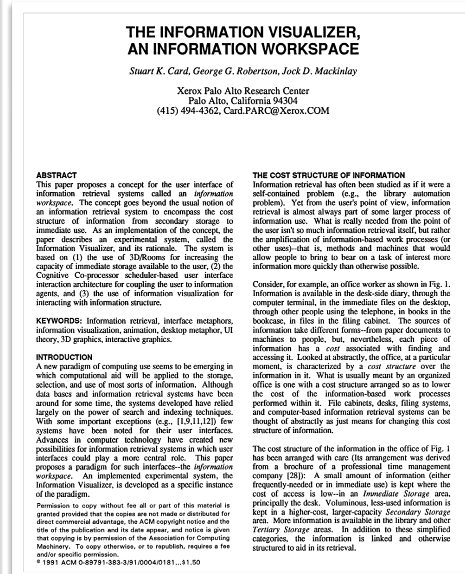


FIGURE **1.40**

Perspective wall. Courtesy of Inxight (1991).





THE INFORMATION VISUALIZER, AN INFORMATION WORKSPACE

Stuart K. Card, George G. Robertson, Jock D. Mackinlay

Xerox Palo Alto Research Center
Palo Alto, California 94304
(415) 494-4362, Card.PARC@Xerox.COM

ABSTRACT

This paper proposes a concept for the user interface of information retrieval systems called an *information workspace*. The concept goes beyond the usual notion of an information retrieval system to encompass the cost structure of information from secondary storage to immediate use. As an implementation of the concept, the paper describes an experimental system, called the Information Visualizer, and its rationale. The system is based on (1) the use of 3D/Rooms for increasing the capacity of immediate storage available to the user, (2) the Cognitive Co-processor scheduler-based user interface interaction architecture for coupling the user to information agents, and (3) the use of information visualization for interacting with information structure.

KEYWORDS: Information retrieval, interface metaphors, information visualization, animation, desktop metaphor, UI theory, 3D graphics, interactive graphics.

INTRODUCTION

A new paradigm of computing use seems to be emerging in which computational aid will be applied to the storage, selection, and use of most sorts of information. Although data bases and information retrieval systems have been around for some time, the systems developed have relied largely on the power of search and indexing techniques. With some important exceptions (e.g., [1,9,11,12]) few systems have been noted for their user interfaces. Advances in computer technology have created new possibilities for information retrieval systems in which user interfaces could play a more central role. This paper proposes a paradigm for such interfaces--the *information workspace*. An implemented experimental system, the Information Visualizer, is developed as a specific instance of the paradigm.

THE COST STRUCTURE OF INFORMATION

Information retrieval has often been studied as if it were a self-contained problem (e.g., the library automation problem). Yet from the user's point of view, information retrieval is almost always part of some larger process of information use. What is really needed from the point of the user isn't so much information retrieval itself, but rather the amplification of information-based work processes (or other uses)--that is, methods and machines that would allow people to bring to bear on a task of interest more information more quickly than otherwise possible.

Consider, for example, an office worker as shown in Fig. 1. Information is available in the desk-side diary, through the computer terminal, in the immediate files on the desktop, through other people using the telephone, in books in the bookcase, in files in the filing cabinet. The sources of information take different forms--from paper documents to machines to people, but, nevertheless, each piece of information has a *cost* associated with finding and accessing it. Looked at abstractly, the office, at a particular moment, is characterized by a *cost structure* over the information in it. What is usually meant by an organized office is one with a cost structure arranged so as to lower the cost of the information-based work processes performed within it. File cabinets, desks, filing systems, and computer-based information retrieval systems can be thought of abstractly as just means for changing this cost structure of information.

The cost structure of the information in the office of Fig. 1 has been arranged with care (Its arrangement was derived from a brochure of a professional time management company [28]): A small amount of information (either frequently-needed or in immediate use) is kept where the cost of access is low--in an *Immediate Storage* area, principally the desk. Voluminous, less-used information is kept in a higher-cost, larger-capacity *Secondary Storage* area. More information is available in the library and other *Tertiary Storage* areas. In addition to these simplified categories, the information is linked and otherwise structured to aid in its retrieval.

Information Visualizer: Table Lens

Visualization: A Computational Medium

Computation provides the most plastic medium we have ever know for representation.

Alan Kay refers to computers as a metamedium

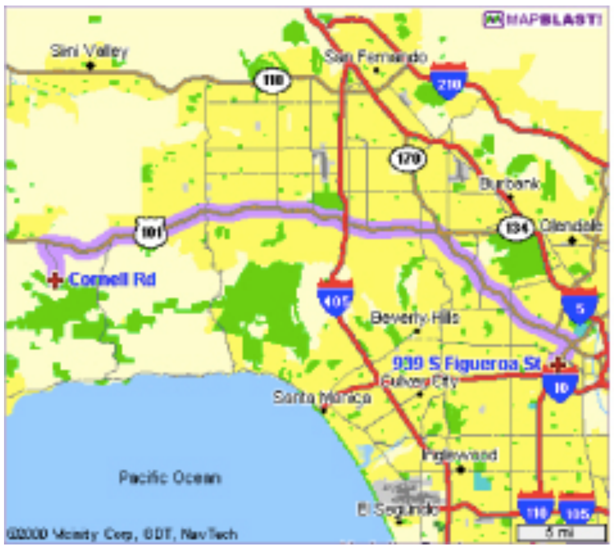
“It is the first metamedium, and as such it has degrees of freedom for representation and expression never before encountered and as yet barely investigated.”

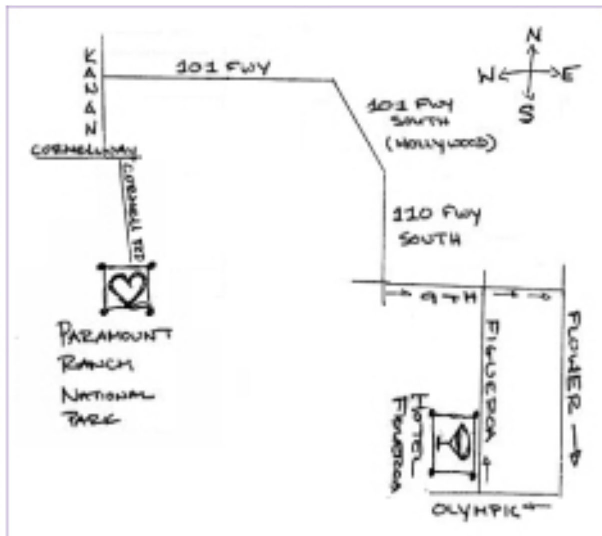
Fundamental importance of understanding tasks and activities

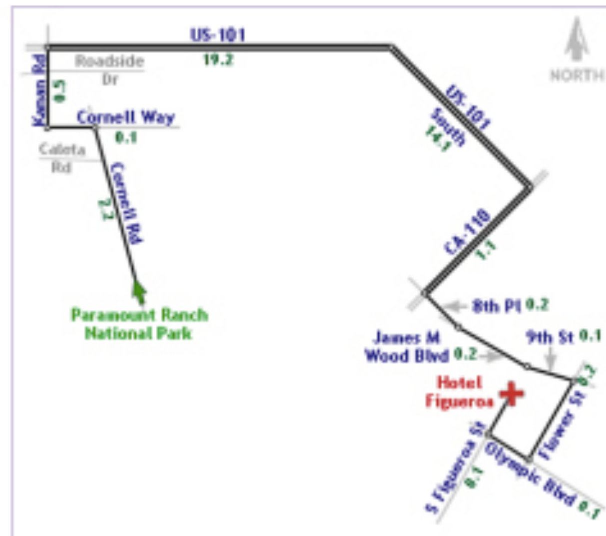
Understanding Tasks

Rendering Effective Route Maps, Maneesh Agrawala

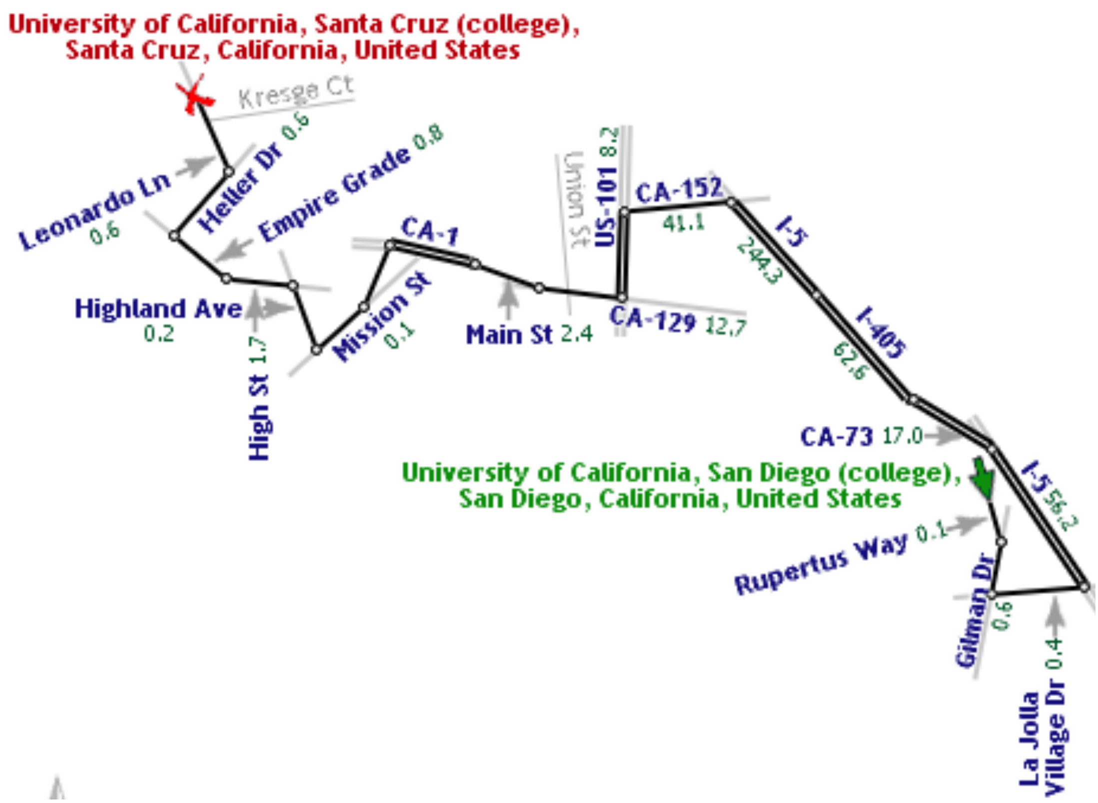
Distortions in Hand-Drawn Maps: Exaggeration (e.g. road length), Regularization (e.g. turning angle), Simplification (e.g. road shape), Debugging







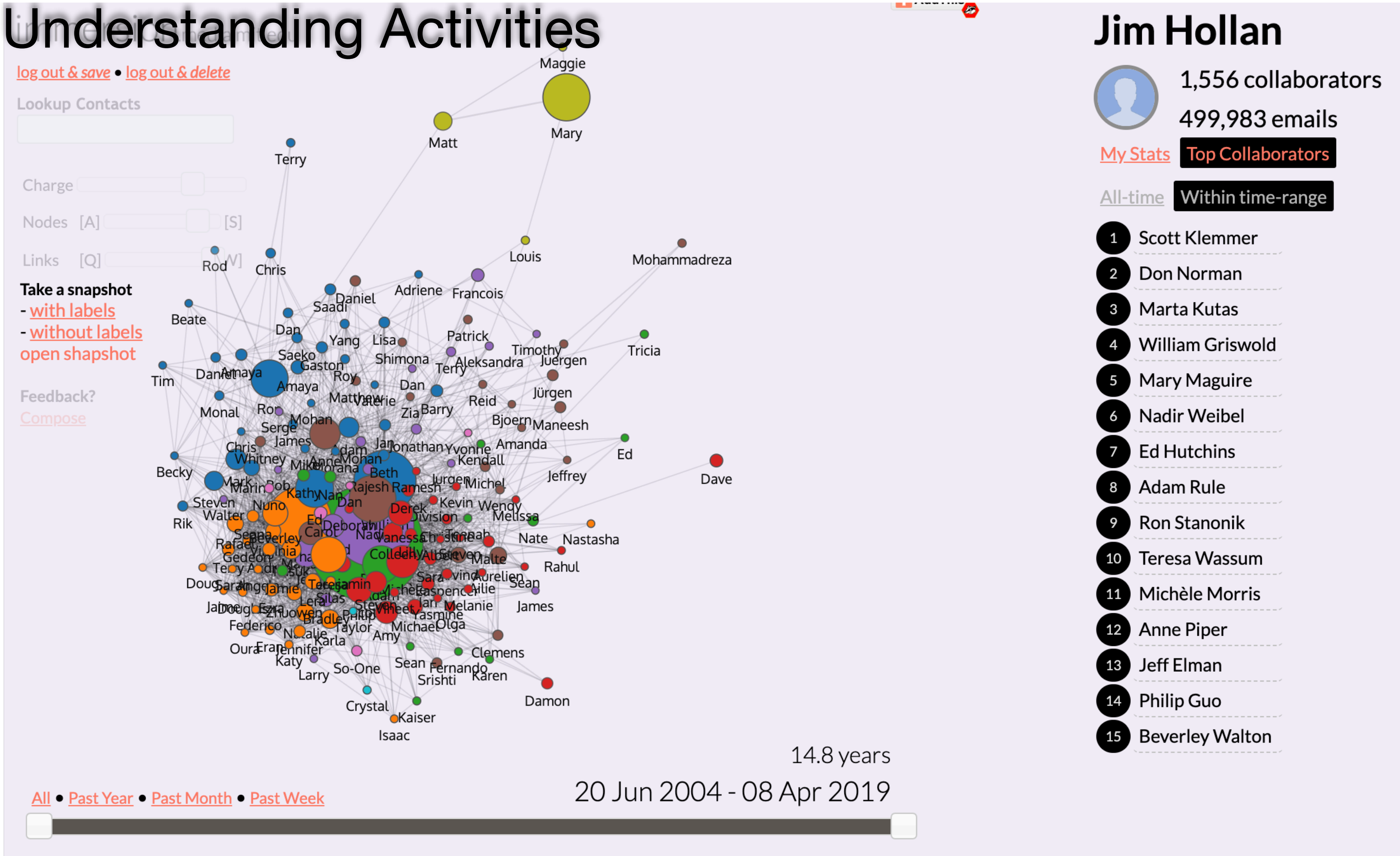
Understanding Tasks



Understanding Activity Over Time

https://www.nytimes.com/interactive/2016/08/20/sports/olympics/decisive-moments-rio-olympics-composite-pictures.html

Understanding Activities



Understanding Activities

log out & save • log out & delete

Lookup Contacts

Charge

Nodes [A] [S]

Links [Q] [W]

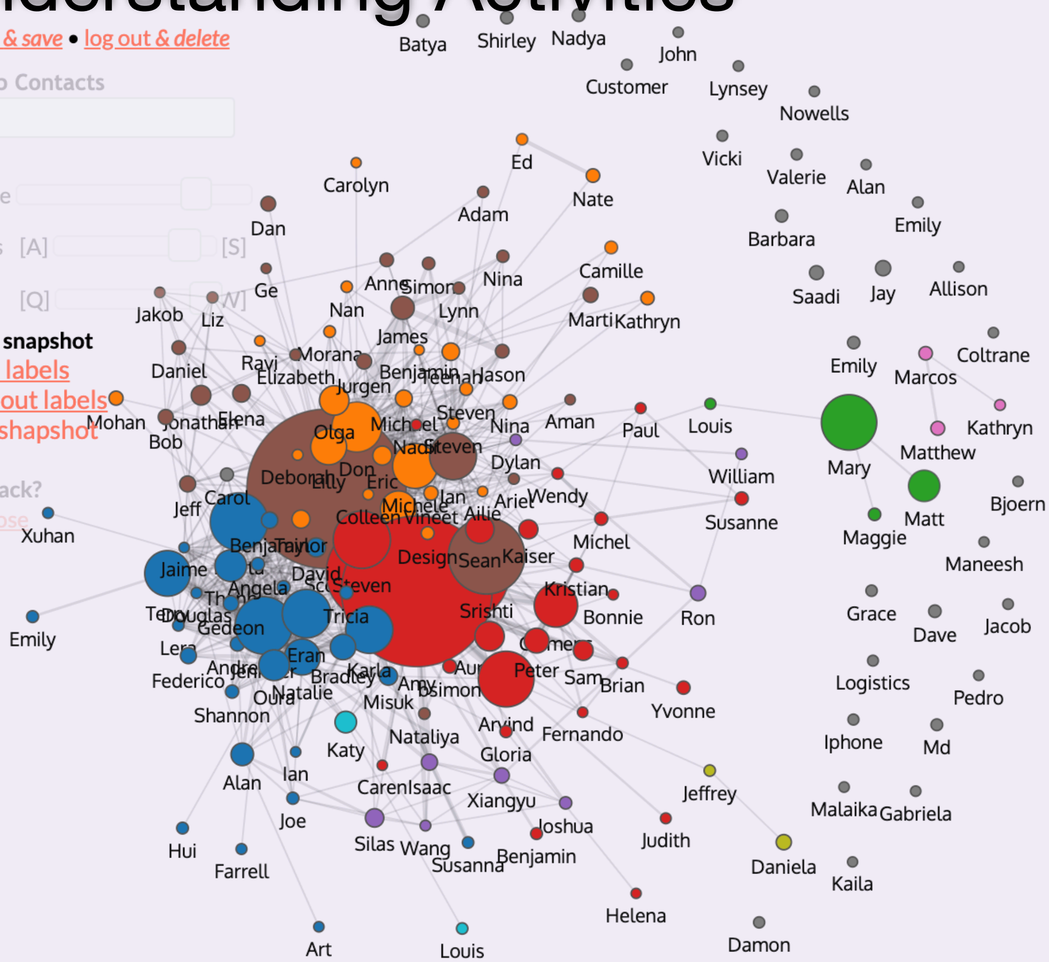
Take a snapshot

- with labels
- without labels

open snapshot

Feedback?

Compose



12.0 months

All • Past Year • Past Month • Past Week

08 Apr 2018 - 08 Apr 2019

Jim Hollan

1,556 collaborators

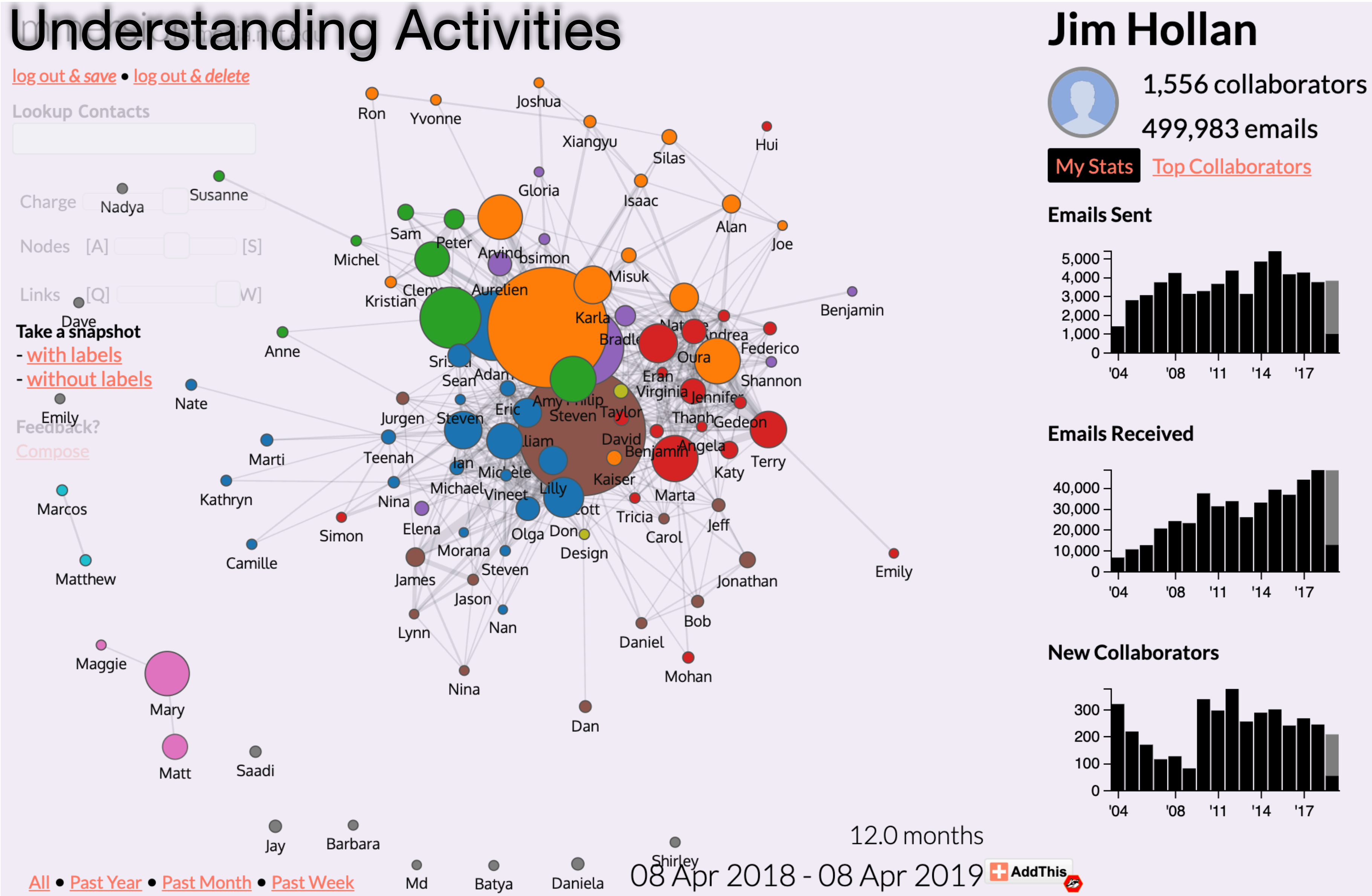
499,983 emails

My Stats | Top Collaborators

All-time | Within time-range

1. Scott Klemmer
2. Amy Fox
3. Philip Guo
4. Adam Rule
5. Srishti Palani
6. Marta Kutas
7. Jennifer Griffin
8. Steven Dow
9. Arvind Satyanarayan
10. Mary Maguire
11. Don Norman
12. Eran Mukamel
13. Karla Rivera
14. Ian Strelsky
15. Terry Jernigan

Understanding Activities



Tasks and Efficiency

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks and activities more effectively.

Tasks and activities serve as constraint on design (as does data)

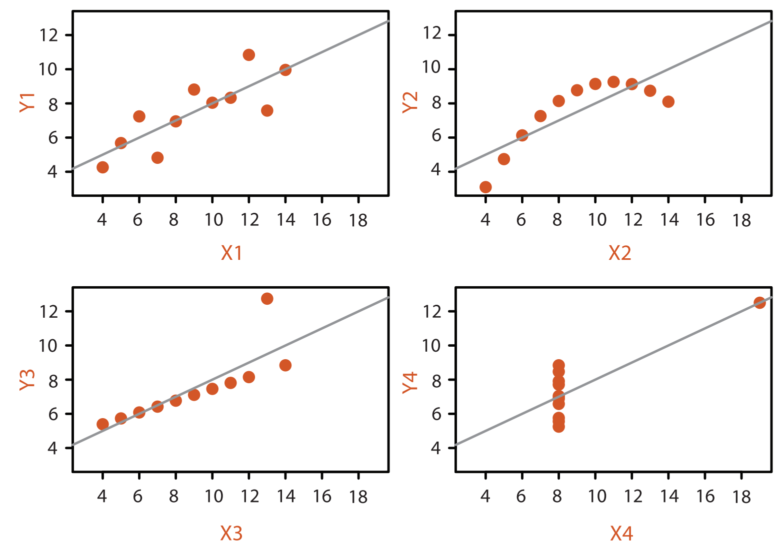
Challenge: recast tasks and activities from domain-specific vocabulary to visual forms

Exploit Perception and Details Matter

Replace cognition with perception

More than summaries, details matter

Identical statistics	
x mean	9
x variance	10
y mean	7.5
y variance	3.75
x/y correlation	0.816



Resource Limitations

Vis designers must take into account three very different kinds of resource limitations: those of computers, of humans, and of displays.

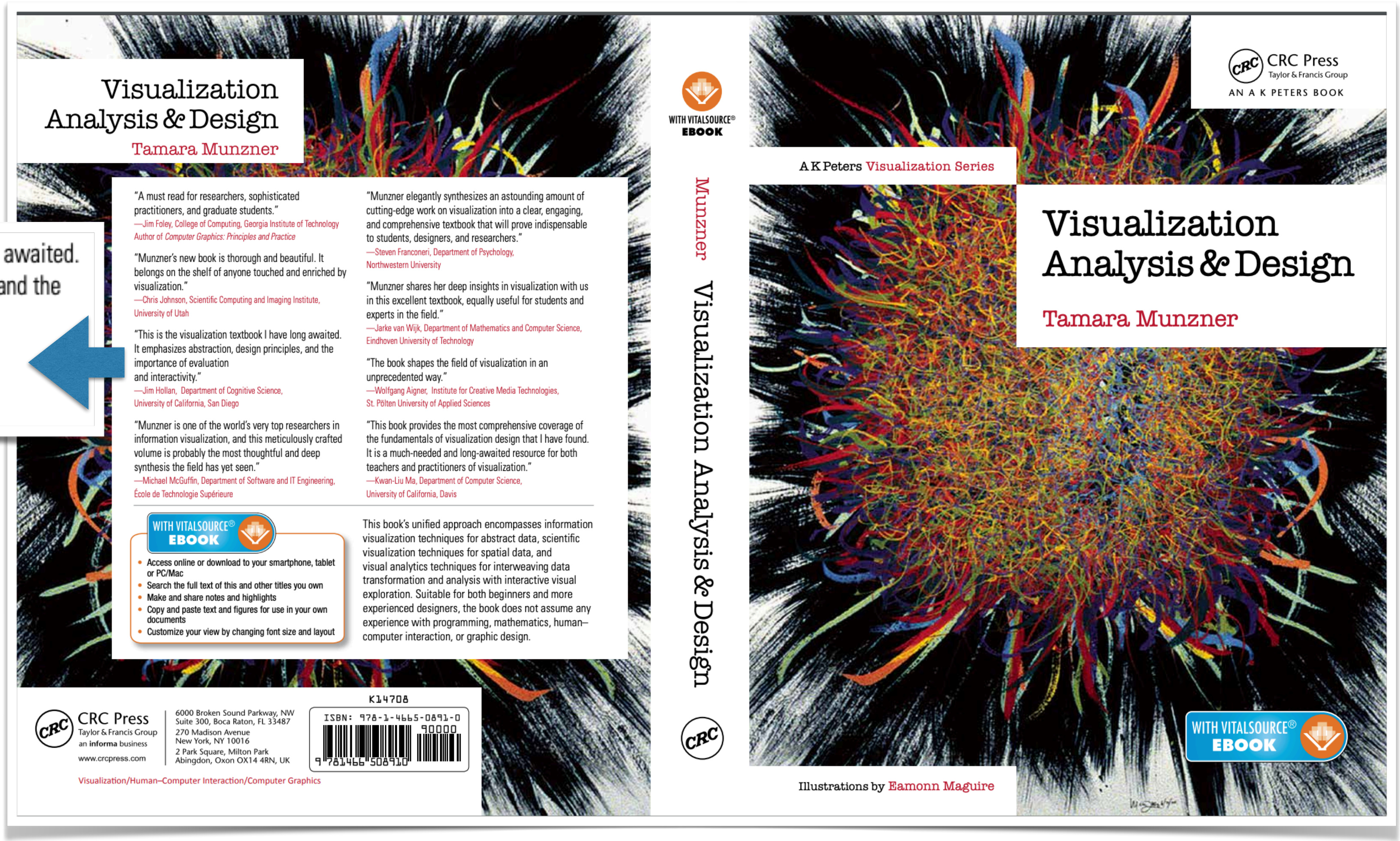
Computational limits: processing time, system memory, …

Human limits: human attention, memory, …

Display limits and Information density: pixels are precious resource, ratio of space used to encode info vs whitespace, …

Tamara Munzner: Visualization Analysis & Design

"This is the visualization textbook I have long awaited. It emphasizes abstraction, design principles, and the importance of evaluation and interactivity."
—Jim Hollan, Department of Cognitive Science, University of California, San Diego



What, Why, and How

What is shown?

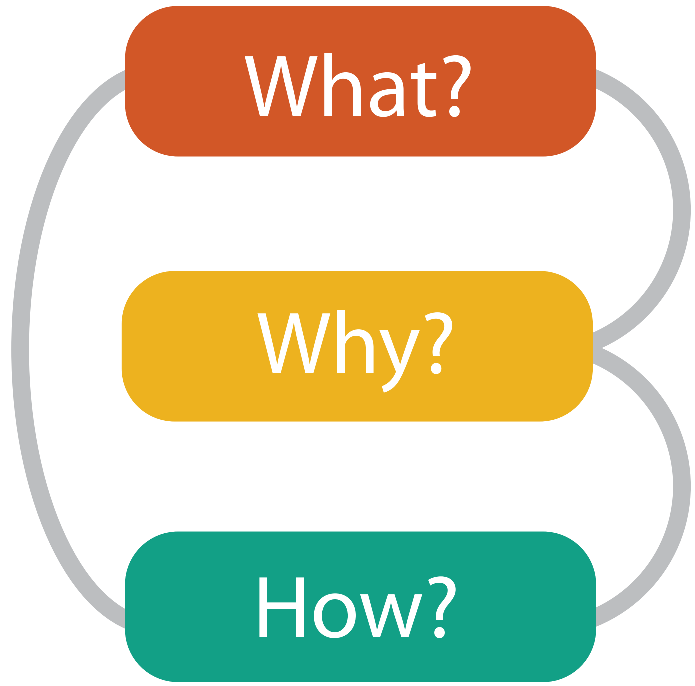
Data abstraction

Why is the user looking at it?

Task abstraction

How is it shown?

Idiom: visual encoding and interaction



What?

Datasets

Data Types

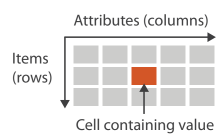
- Items
- Attributes
- Links
- Positions
- Grids

Data and Dataset Types

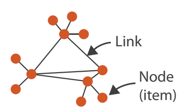
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Dataset Types

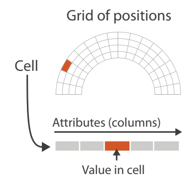
- Tables



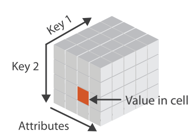
- Networks



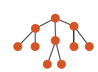
- Fields (Continuous)



- *Multidimensional Table*



- *Trees*



- Geometry (Spatial)

Attributes

Attribute Types

- Categorical

- Ordered
  - *Ordinal*

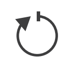
  - *Quantitative*

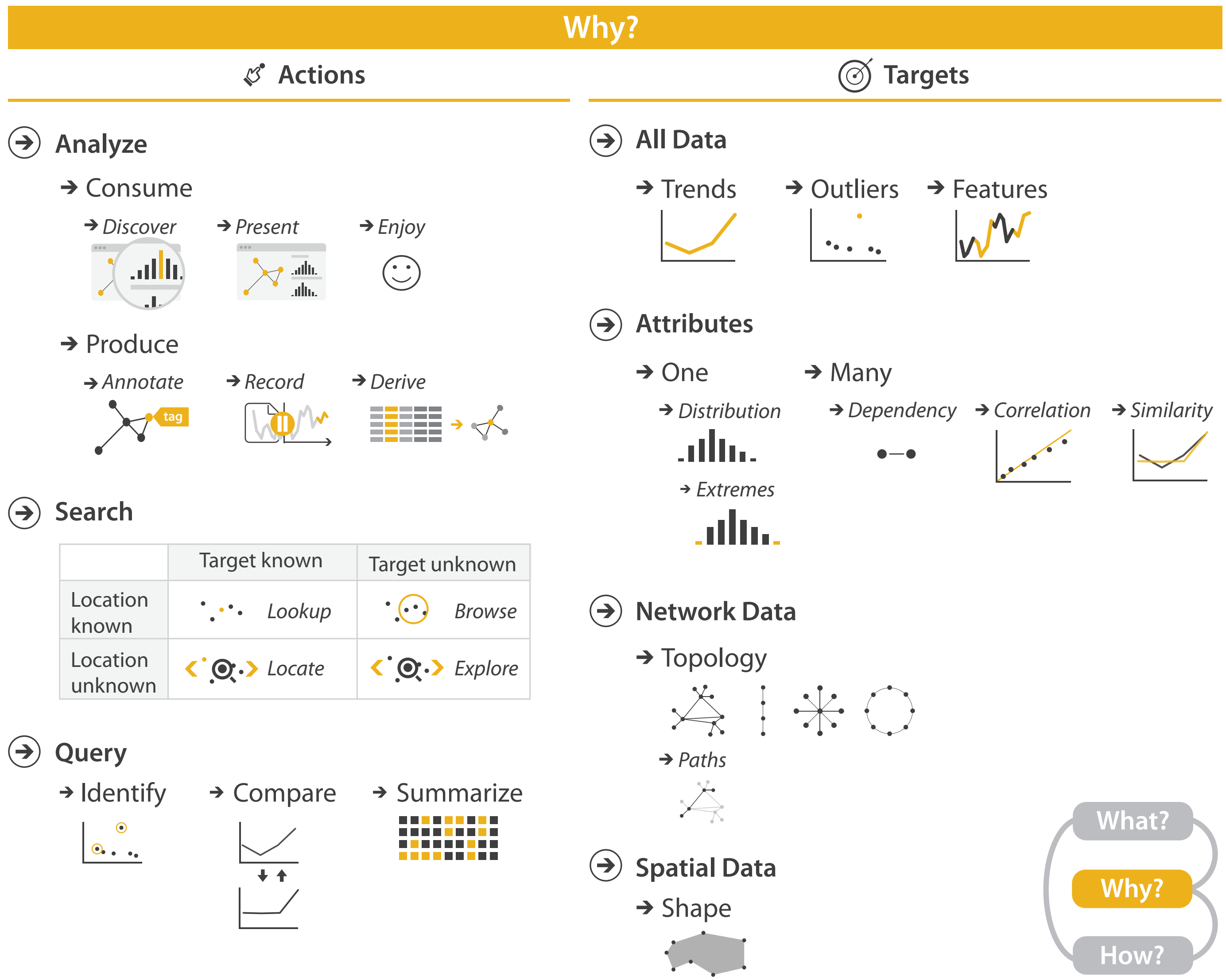
Ordering Direction

- Sequential

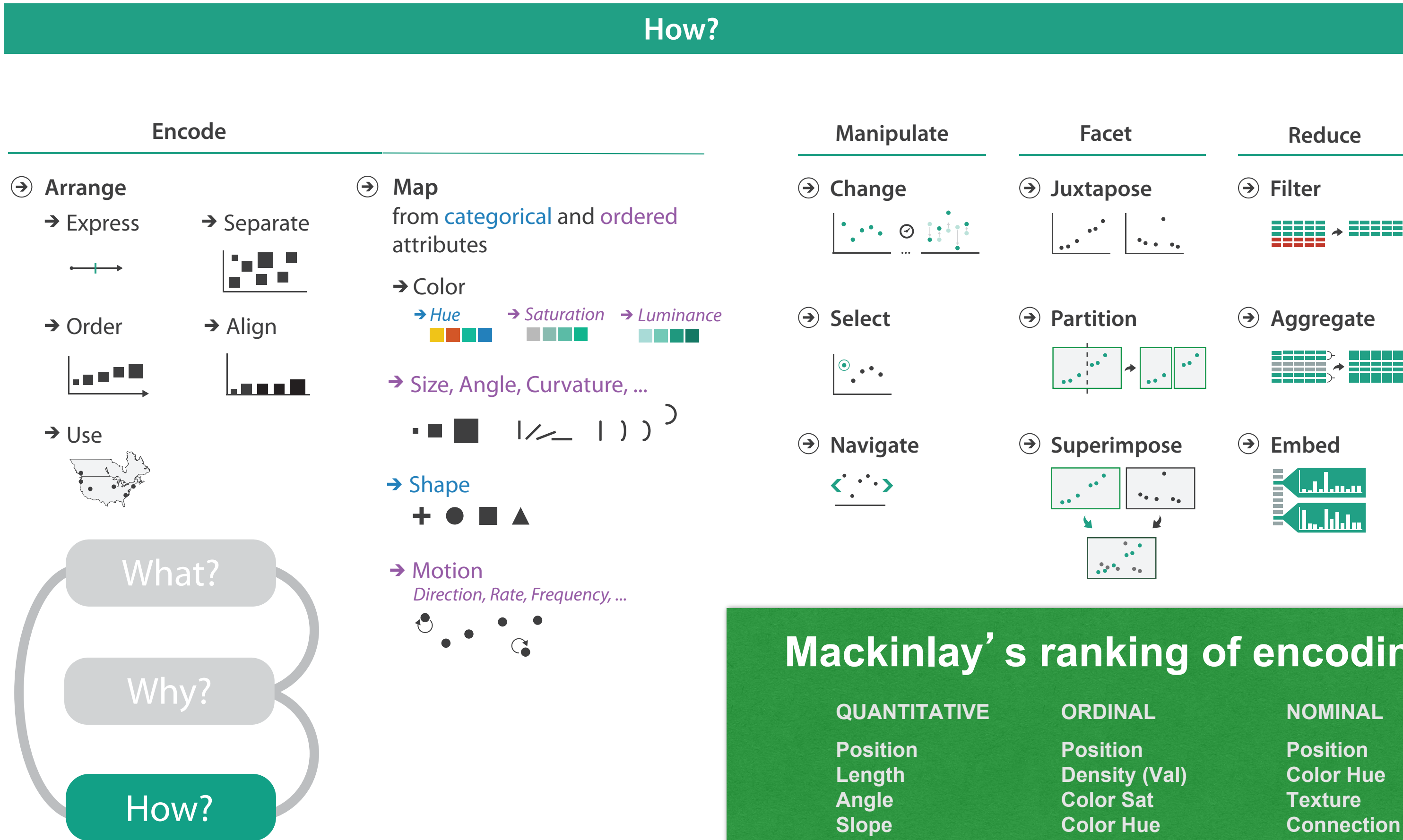
- Diverging

- Cyclic





Why?
Actions
Targets
Analyze
Consume
Discover
Present
Enjoy
Produce
Annotate
tag
Record
Derive
Search
Target known
Target unknown
Location known
Lookup
Browse
Location unknown
Locate
Explore
Query
Identify
Compare
Summarize
All Data
Trends
Outliers
Features
Attributes
One
Many
Distribution
Dependency
Correlation
Similarity
Extremes
Network Data
Topology
Paths
Spatial Data
Shape
What?
Why?
How?



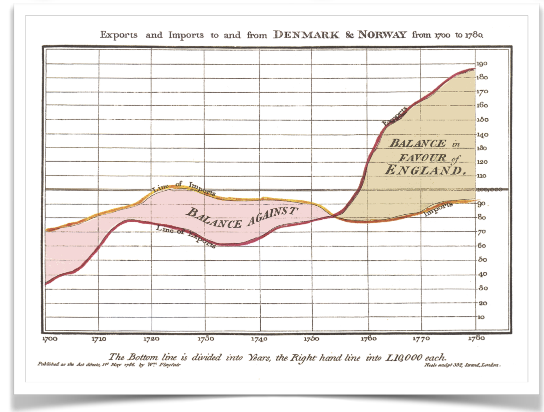
Mackinlay's ranking of encodings

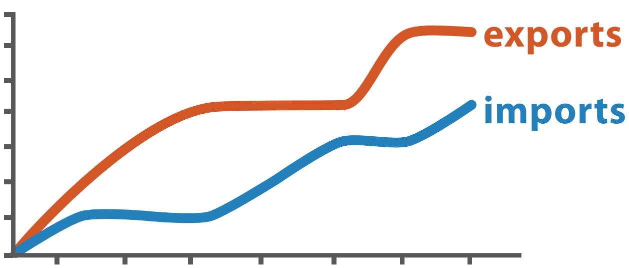
QUANTITATIVE	ORDINAL	NOMINAL
Position	Position	Position
Length	Density (Val)	Color Hue
Angle	Color Sat	Texture
Slope	Color Hue	Connection
Area (Size)	Texture	Containment
Volume	Connection	Density (Val)
Density (Val)	Containment	Color Sat
Color Sat	Length	Shape
Color Hue	Angle	Length
Texture	Slope	Angle
Connection	Area (Size)	Slope
Containment	Volume	Area
Shape	Shape	Volume

Goal: Reduce Task and Activity Complexity

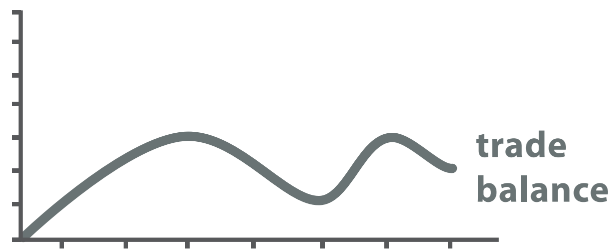
Don't just draw what you're given

- Decide what the right thing to show is
- Create it with a series of transformations from the original dataset
- Draw that "right thing"





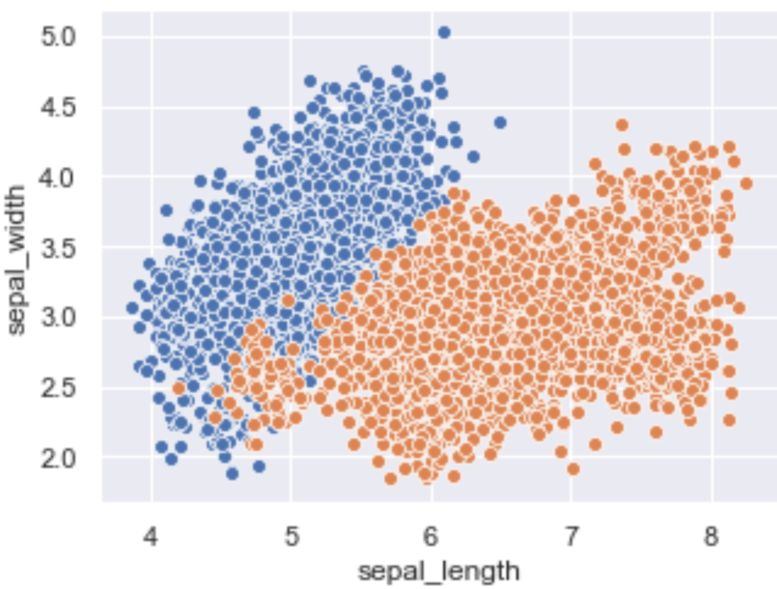
Original Data

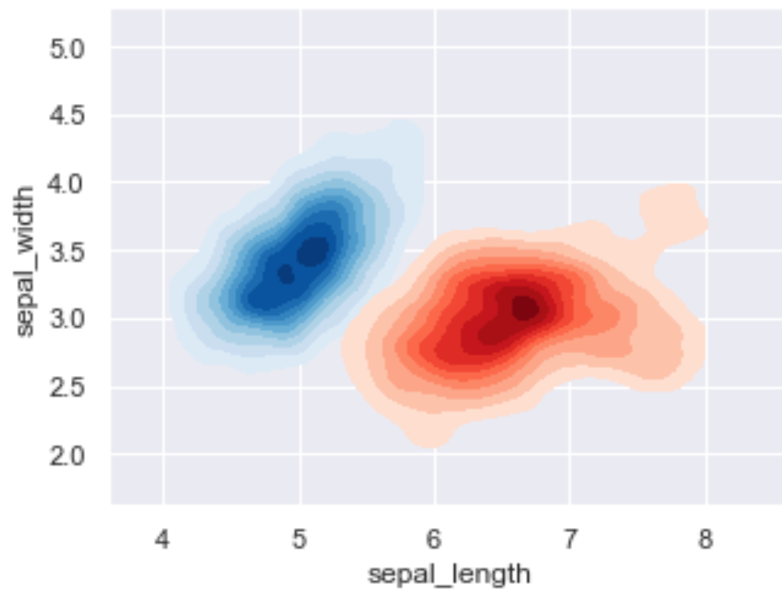


trade balance = *exports* − *imports*

Derived Data

Today Face Many Big Data Issues





Commonly used plots fail to scale to large datasets. Many interesting issues result. How might we address this? Perhaps by using averaging, binning, and smoothing.

Information Visualization

The power of the unaided mind is highly overrated. Without external aids, memory, thought, and reasoning are all constrained. But human intelligence is highly flexible and adaptive, superb at inventing procedures and objects that overcome its own limits. The real powers come from devising external aids that enhance cognitive abilities. How have we increased memory, thought, and reasoning? By the invention of external aids: It is things that make us smart. (Norman, 1993, p. 43)

How information visualization amplifies cognition.

Increased Resources	
High-bandwidth hierarchical interaction	The human moving gaze system partitions limited channel capacity so that it … resolution and wide aperture in sensing visual environments (Resnikoff, 1987)
Parallel perceptual processing	Some attributes of visualizations can be processed in parallel compared to te…
Offload work from cognitive to perceptual system	Some cognitive inferences done symbolically can be recoded into inferences … perceptual operations (Larkin and Simon, 1987).
Expanded working memory	Visualizations can expand the working memory available for solving a problem…
Expanded storage of information	Visualizations can be used to store massive amounts of information in a quick… maps).
Reduced Search	
Locality of processing	Visualizations group information used together, reducing search (Larkin and S…
High data density	Visualizations can often represent a large amount of data in a small space (Tu…
Spatially indexed addressing	By grouping data about an object, visualizations can avoid symbolic labels (L…
Enhanced Recognition of Patterns	
Recognition instead of recall	Recognizing information generated by a visualization is easier than recalling th… user.
Abstraction and aggregation	Visualizations simplify and organize information, supplying higher centers with aggregated forms of information through abstraction and selective omission (Card, Robertson, and Mackinlay, 1991; Resnikoff, 1987).
Visual schemata for organization	Visually organizing data by structural relationships (e.g., by time) enhances patterns.
Value, relationship, trend	Visualizations can be constructed to enhance patterns at all three levels (Bertin, 1977/1981).
Perceptual Inference	
Visual representations make some problems obvious	Visualizations can support a large number of perceptual inferences that are extremely easy for humans (Larkin and Simon, 1987).
Graphical computations	Visualizations can enable complex specialized graphical computations (Hutchins, 1996).
Perceptual Monitoring	Visualizations can allow for the monitoring of a large number of potential events if the display is organized so that these stand out by appearance or motion.
Manipulable Medium	Unlike static diagrams, visualizations can allow exploration of a space of parameter values and can amplify user operations.

Information Visualization

To understand something is called "seeing" it. We try to make our ideas "clear," to bring them into "focus," to "arrange" our thoughts. The ubiquity of visual metaphors in describing cognitive processes hints at a nexus of relationships between what we see and what we think. When we imagine someone hard at mental work, we might picture a scholar drawing a diagram, a book of sources open at her side. Or we might imagine a stockbroker, watching computer displays of financial data, rushing to act on events. Whatever the activity, mental work and perceptual interactions of the world are likely to be interwoven.

This interweaving of interior mental action and external perception (and manipulation) is no accident. It is the essence of how we achieve expanded intelligence. As Norman says,

> *The power of the unaided mind is highly overrated. Without external aids, memory, thought, and reasoning are all constrained. But human intelligence is highly flexible and adaptive, superb at inventing procedures and objects that overcome its own limits. The real powers come from devising external aids that enhance cognitive abilities. How have we increased memory, thought, and reasoning? By the invention of external aids: It is things that make us smart.* (Norman, 1993, p. 43)

An important class of the external aids that make us smart are graphical inventions of all sorts. These serve two related but quite distinct purposes. One purpose is for communicating an idea, for which it is sometimes said, "A picture is worth ten thousand words."[1] Communicating an idea requires, of course, already having the idea to communicate. The second purpose is to use graphical means to create or discover the idea itself: using the special properties of visual perception to resolve logical problems, as Bertin (1977/1981) would say. *Using vision to think.* This second sense of graphics is the subject of this book.

Graphic aids for thinking have an ancient and venerable history. What is new is that the evolution of computers is making possible a medium for graphics with dramatically improved rendering, real-time interactivity, and dramatically lower cost. This medium allows graphic depictions that automatically assemble thousands of data objects into pictures, revealing hidden patterns. It allows diagrams that move, react, or even initiate. These, in turn, create new methods for amplifying cognition, new means for coming to knowledge and insight about the world. A few years ago, the power of this new medium was applied to science, resulting in scientific visualization. Now it is possible to apply the medium more generally to business, to scholarship, and to education. This broader application goes under the name of *information visualization*. The purpose of this book is to introduce information visualization, to collect some of the important papers in the field, and to give samples of some of the latest work.

EXTERNAL COGNITION

To understand the intuition behind information visualization, it is useful to gain an appreciation for the important role of the external world in thought and reasoning. This notion is sometimes called *external cognition* (Scaife and Rogers, 1996) to express the way in which internal and external representations and processing weave together in thought. As Norman suggests, the use of the external world, and especially the use of cognitive artifacts or physical inventions to enhance cognition, is all around us.

Multiplication Aids

Take multiplication, one of the most mental of activities. Have a person multiply a pair of two-digit numbers, such as 34 × 72, in his or her head and time how long it takes. Now repeat the experiment with another pair of numbers, in longhand using pencil and paper.

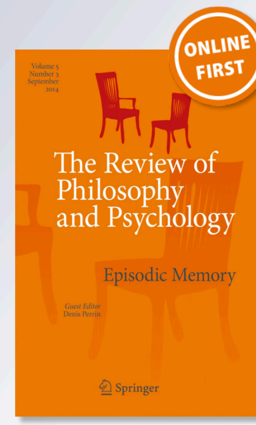
$$\begin{array}{r} 34 \\ \times 72 \\ \hline 68 \\ 23^{2}80 \\ \hline 24^{1}48 \end{array}$$

[1]According to Paul Martin Lester, professor of communications at the University of California at Fullerton, this quotation was simply made up by ad writer Frederick R. Barnard and included as an invented "Chinese proverb" in a streetcar advertisement for Royal Baking Powder. The ad writer wanted to make the point that pictures can attract attention faster than other media. See *http://www5.Fullerton.edu/les/ad.html* and *Printers' Ink*, March 10, 1927.

The Cognitive Design of Tools of Thought

Barbara Tversky

Review of Philosophy and Psychology
ISSN 1878-5158
Rev.Phil.Psych.
DOI 10.1007/s13164-014-0214-3

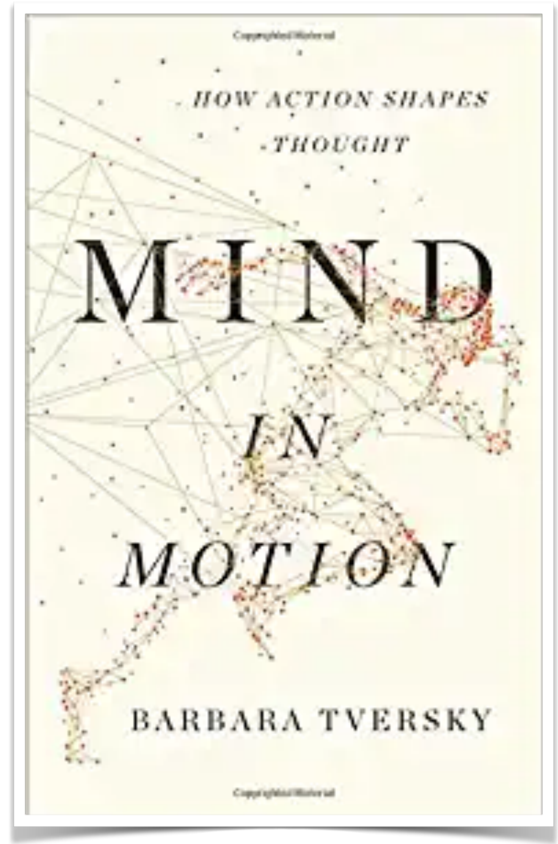


Barbara Tversky

In her new book, Mind in Motion, Barbara Tversky cogently describes research on ***how we think about space—and how we use space to think.*** Based on decades of empirical work in spatial cognition and external representation, Tversky formulates two principles for cognitively-driven design:

Principle of Correspondence: The content and form of the representation should match the content and form of the targeted concepts.

Principle of Use: The representation should promote efficient accomplishment of the targeted tasks.



Contact Info

chc004@ucsd.edu	Chen	Chen
mgahl@ucsd.edu	Martha	Gahl
mhaupt@ucsd.edu	Michael	Haupt
shillin@ucsd.edu	Sloan	Hill-Lindsay
tlatonatequida@ucsd.edu	Talia	LaTona-Tequida
cmeehan@eng.ucsd.edu	Casey	Meehan
kenakai@ucsd.edu	Kendall	Nakai
nrizvi@ucsd.edu	Naba	Rizvi
tsharkey@ucsd.edu	Tommy	Sharkey
msmart@eng.ucsd.edu	Mary Anne	Smart
bstringer@ucsd.edu	Brinley	Stringer
l5sun@ucsd.edu	Lu	Sun
ktenney@ucsd.edu	Kristin	Tenney
iswhite@ucsd.edu	Isabel	White
gwwolfe@ucsd.edu	Garrett	Wolfe
xuy021@ucsd.edu	Xuhan	Yang
myarmand@eng.ucsd.edu	Matin	Yarmand

Breakout Groups

First

Stringer, Latona-Tequida, Tenney, White

Sharkey, Rizvi, Yarmand, Chen

Representation
Yarmand, Yang, Haupt, Rizzi, Sharkey, Gahl, Nakai

Activity Histories, Pad++, Lenses

Wolfe, Hill-Lindsay, Chen, Mehan, Smart, Sun

Second

Stringer, Latona-Tequida, Tenney, White

Sharkey, Rizvi, Yarmand. Chen

Representation
Wolfe, Hill-Lindsay

Organization
Yang, Haupt

Cogsci 220: Information Visualization (Winter 2021)

<< Back to home

During our meeting in week 3 (1/19) you will meet in two 30 minute breakout/discussion groups. Please refer back to the description of topics before class to prepare.

First Breakout Session

- *Representation and Behavior of Information Entities*
  White, Stringer, Tenney, Yarmand, Yang, Haupt, Rizvi, Sharkey, Gahl, Nakai, LaTona-Tequida
- *Activity Histories. Pad++, and Lenses*
  Wolfe, Hill-Lindsay, Chen, Meehan, Smart, Sun

Reconvene Class for Feedback

Second Breakout Session

- *Representation and Behavior of Information Entities*
  Wolfe, Hill-Lindsay, White, Stringer, Tenney, LaTona-Tequida
- *Organization and Behavior of Groups of Information Entities*
  Yarmand, Yang, Haupt, Sharkey, Rizvi
- *Activity Histories and Webstrates*
  Gahl, Nakai, Meehan, Smart, Chen, Sun

Department of Cognitive Science - UC San Diego

Ideally by the end of the breakout groups you should be close to forming a group with a general topic focus.

By the end of the week (Friday 1/22), send a note to 220-g@ucsd.edu with a list of your team members and your general topic focus. Be prepared next Tuesday (1/26) to give a short (5 minute) description of your focus.