Old Dominion University

CS 795/895
Information Visualization
Fall 2012

Cognition and User Tasks

Dr. Michele C. Weigle

http://www.cs.odu.edu/~mweigle/CS795-F12/

Outline

- Cognition
- User Tasks
- Analytic Tasks
- Resources
  - Card, Mackinlay, and Shneiderman, *Readings in Information Visualization*
    - aka CMS text
  - John Stasko's CS 7450 lectures (GaTech, Spr 11)

What is Cognition?

- "In science, cognition is a group of mental processes that includes attention, memory, producing and understanding language, solving problems, and making decisions."

- "Cognition is studied in various disciplines such as psychology, philosophy, linguistics, science and computer science."

http://en.wikipedia.org/wiki/Cognition

What is Cognition?

- Visualization is often thought of as process of making a graphic or an image

- Really is a cognitive process
  - form a mental image of something
  - internalize an understanding

- "The purpose of visualization is insight, not pictures" -Shneiderman
  - insight: discovery, decision making, explanation

Content courtesy of John Stasko, Ga Tech

Main Idea

Larkin & Simon '87
CMS text

- Visuals help us think
  - provide a frame of reference, a temporary storage area
- Cognition → Perception
- Pattern matching
- External cognition aid
  - role of external world in thinking and reason

Content courtesy of
John Stasko, Ga Tech

External Cognition
Reminder

paper	mental buffer
57 x 48 / 1 456 / 228 / 2736	[7 * 8 = 56]
	[5 * 8 = 40 + 5 = 45]
	[7 * 4 = 28]
	[5 * 4 = 20 + 2 = 22]
	[8 + 5 = 13]
	[4 + 2 + 1 = 7]

Tamara Munzner,
CPSC 533C, UBC

Definitions

- External Cognition
  - use of the *external world* to accomplish cognition
- Information Design
  - design of *external representations* to amplify cognition
- Data Graphics
  - use of *abstract, nonrepresentational* visual representations of data to amplify cognition

Definitions

- Visualization
  - use of *computer-based, interactive* visual representations of data to amplify cognition
- Scientific Visualization
  - use of interactive visual representations of *scientific data,* typically *physically based*, to amplify cognition
- Information Visualization
  - use of interactive visual representations of *abstract, nonphysically based data* to amplify cognition

Let's Back Up a Second

- Information Visualization
  - use of interactive visual representations of *abstract, nonphysically based data* to amplify cognition

- What does "amplify cognition" mean?

CMS text

Amplifying Cognition

Hollan, Hutchins, Kirsh
TOCHI '00

- A traditional cognitive science view is that tools amplify, or scaffold, cognition

- Hollan et al. argue that this is not correct
  - eg, our memory isn't amplified

- Instead, tools help transform the analytic process into another more doable one

Content courtesy of
John Stasko, Ga Tech

Distributed Cognition

Liu, Nersessian, Stasko
TVCG '08

- Cognitive system is composed of people *and* the artifacts they use
  - cognition isn't only internal
- Changes in external representation spur changes in internal representation and understanding
- It is *interaction* with the external representations that drives this process

Content courtesy of John Stasko, Ga Tech

Understanding

- People utilize an mental/internal model that is generated based on what is observed
- B. Tversky calls the internal model a *cognitive map*

Content courtesy of John Stasko, Ga Tech

Example

- You're taking the DC Metro to get to the National Science Foundation
  - You have some existing internal model of the system, stops, how to get there
  - On train, you glance at the Metro map for help
  - Refines your internal model, clarifying items and extending it
  - Note that it's still not perfect, no internal model ever is

Cognitive Map

- Just don't have one big one
- Have large number of these for all different kinds of things
- Collection of cognitive maps -> *Cognitive collage*

Process Models

- Process by which a person looks at a graphic and makes some use of it
  - a number of sub-steps probably exist
- Can you describe the process?

Navigation Process Model

Spence, *IJHCS* '99

- Navigation - creation and interpretation of an internal mental model

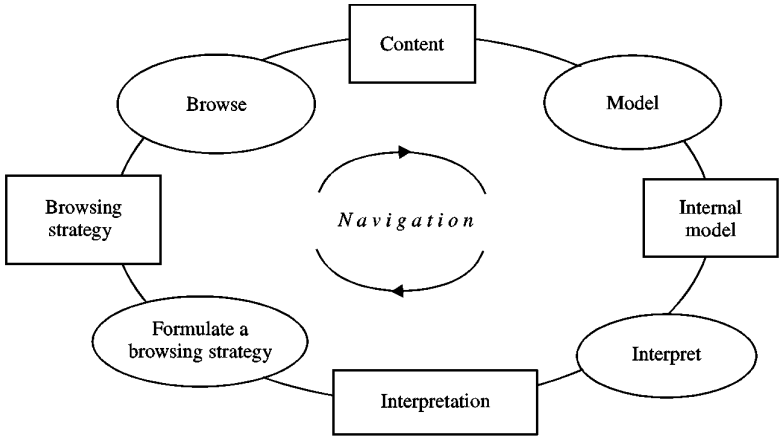
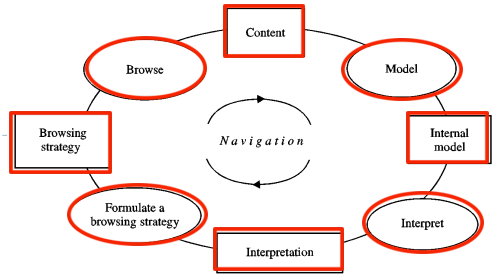


FIGURE 1. The proposed framework for navigation.

Navigation Process Model
Interpretation



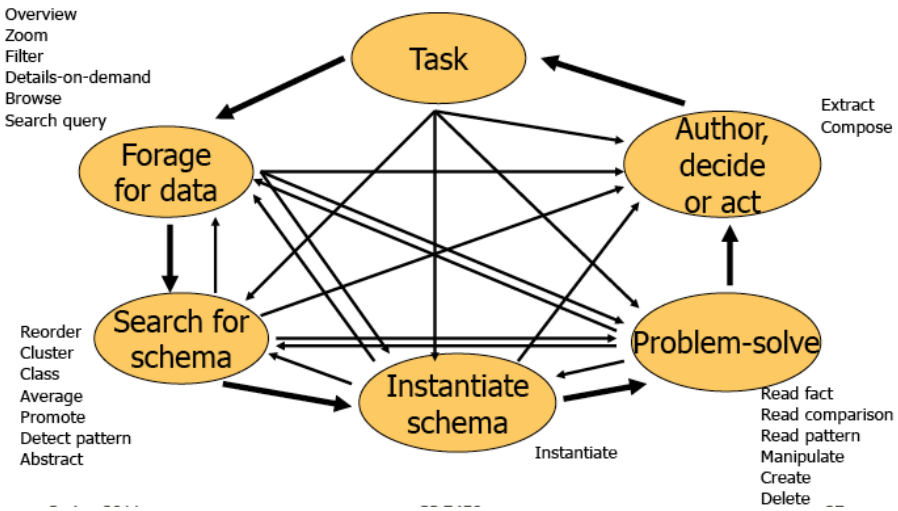
- *Content* is the display on screen
- *Modeling* of that pattern results in cognitive map, or *internal model*
- *Interpretation* ("ah, variables x and y are related") leads to new view, that generates an idea for a new browsing strategy
- Look at the display again (*browse*) with that in mind

CMS text

Knowledge Crystallization Process Model

- Knowledge crystallization task
  - gather info for some purpose
  - make sense of it by constructing a representational framework (*schema*)
  - package it into a form for communication or action

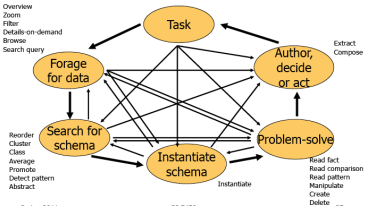
Knowledge Crystallization Process Model



Content courtesy of John Stasko, Ga Tech

Knowledge Crystallization Process Model

- Information foraging
- Search for schema (representation)
- Instantiate schema with data
- Problem solve to trade off features
- Search for a new schema that reduces problem to a simple trade-off
- Package the patterns found in some output product



Content courtesy of John Stasko, Ga Tech

Knowledge Crystallization Process Model
Example

- Forage: Collect articles and data about laptops
- Representation: Id attributes to compare
- Instantiate schema: Make table of laptops and attributes
- Problem solve: Reorder rows and columns, make plots, mark laptops that are out of the running
- Search for a new schema: Cluster into 3 groups (power, multimedia capability, portability), delete all but top 1-2 laptops in each
- Package the patterns: Create a briefing document

CMS text

How Vis Amplifies Cognition

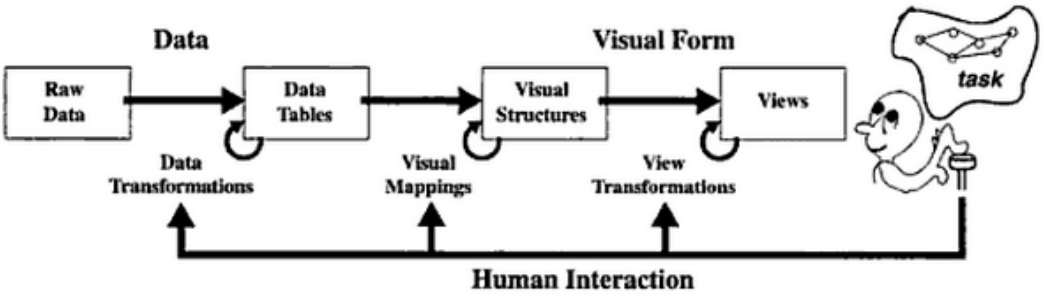
- "Purpose of information visualization is to use *perception* to amplify cognition"

- Increased resources (memory and processing)
- Reduced search
- Enhanced recognition of patterns
- Perceptual inference
- Perceptual monitoring
- Manipulable medium

CMS text
Table 1.3, pg. 16

Content courtesy of John Stasko, Ga Tech

Visualization Reference Model



CMS text
Figure 1.23, pg. 17

Components of the Reference Model

TABLE **1.24**

The components of the reference visualization model shown in Figure 1.23. Specific techniques are also included in the table. The specific techniques for Data Tables, discussed in the text, are a list of common data types that have well-known Data Tables. Tasks are operations that a user may want to do with the visualization.

DATA TABLES	VISUAL STRUCTURES	VIEWS	HUMAN INTERACTION	TASKS	LEVEL
Cases Variables Values Metadata	Spatial Substrate Marks Graphical properties	Location Probes Viewpoint Controls Distortion	Data Tables Visual Structures Views	Forage for Data Problem Solving Search for Schema Instantiate Schema Author, Decide, or Act	Infosphere Workspace Visual Knowledge Tools Visual Objects
Specific Techniques					
Spatial (Scientific) Geographic Documents Time Database Hierarchies Networks World Wide Web	Position: NOQ Marks: PLAV Properties: Connection, Enclosure, Retinal, Time Axes: Composition Alignment Folding Recursion Overloading	Brushing Zooming Overview + Detail Focus + Context	Dynamic Queries Direct Manipulation Magic Lens	Overview Zoom Filter Details-on-Demand Browse Search Read Fact Read Comparison Read Pattern Manipulate Create	Delete Reorder Cluster Class Promote Average Abstract Instantiate Extract Compose Organize

CMS text
Table 1.24, pg. 33

Outline

- Cognition
- User Tasks
- Analytic Tasks

User Tasks

- What things will people want to accomplish using information visualizations?
- Earlier, we briefly discussed search vs. browsing

Content courtesy of John Stasko, Ga Tech

Browsing vs. Search

- Important difference in activities
- Appears that information visualization may have more to offer to browsing
- But...browsing is a softer, fuzzier activity
- So, how do we articulate utility?
  - maybe describe when it's useful
  - when is browsing useful?

Browsing is Useful When…

Lin '97

- Good underlying structure so that items close to one another can be inferred to be similar
- Users are unfamiliar with collection contents
- Users have limited understanding of how system is organized and prefer less cognitively loaded method of exploration
- Users have difficulty verbalizing underlying information need
- Information is easier to recognize than describe

Thought

- Maybe infovis isn't about answering questions or solving problems...
- Maybe it's about asking better questions

Tasks

- Browsing and searching are very high level
- Let's be more specific…
- What things will people want to accomplish using information visualizations?

Task Taxonomies

- Number of different ones exist, important to understand what process they focus on
  - creating an artifact
  - human tasks
  - tasks using the visualization system

User Task Taxonomy

Wehrend & Lewis Vis '90

- Wehrend & Lewis created a low-level, domain independent taxonomy of user tasks in visualization environments

- Eleven basic actions
  - identify, locate, distinguish, categorize, cluster, distribution, rank, compare within relations, compare between relations, associate, correlate

Task x Data Type Taxonomy

Shneiderman VL '96 *assigned*

- What do people do with visualization?
- Mantra: "Overview first, zoom and filter, then details on demand"
  - design paradigm for infovis systems

Content courtesy of John Stasko, Ga Tech

Task x Data Type Taxonomy

Tasks

1. overview
2. zoom
3. filter
4. details-on-demand
5. relate
6. history
7. extract

Data Types

1. 1-D linear
2. 2-D map
3. 3-D world
4. N-D (multi-dimensional)
5. temporal
6. tree
7. network

Shneiderman, IEEE VL 1996

Analytic Task Taxonomy

Amar, Eagan, & Stasko, InfoVis '05

- Retrieve Value
  - given a set of specific cases, find attributes of those cases
- Filter
  - given some concrete conditions on attribute values, find data cases satisfying those conditions
- Compute Derived Value
  - given a set of data cases, compute an aggregate numeric representation of those data cases
- Find Extremum
  - find data cases possessing an extreme value of an attribute over its range within the data set
- Sort
  - given a set of data cases, rank them according to some ordinal metric

Analytic Task Taxonomy

Amar, Eagan, & Stasko, InfoVis '05

- Determine Range
  - given a set of data cases and an attribute of interest, find the span of values within the set
- Characterize Distribution
  - given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute's values over the set
- Find Anomalies
  - identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers
- Cluster
  - given a set of data cases, find clusters of similar attribute values
- Correlate
  - given a set of data cases and two attributes, determine useful relationships between the values of those attributes

What Questions Were Left Out?

- Basic math
  - "Which cereal has more sugar, Cheerios or Special K?"
  - "Compare the average MPG of American and Japanese cars."
- Uncertain criteria
  - "Does cereal (X, Y, Z...) sound tasty?"
  - "What are the characteristics of the most valued customers?"
- Higher-level tasks
  - "How do mutual funds get rated?"
  - "Are there car aspects that Toyota has concentrated on?"
- More qualitative comparison
  - "How does the Toyota RAV4 compare to the Honda CRV?"
  - "What other cereals are most similar to Trix?"

Content courtesy of John Stasko, Ga Tech

Outline

- Cognition
- User Tasks
- Analytic Tasks

High-Level Analysis Tasks

- Complex decision-making, especially under uncertainty
- Learning a domain
- Identifying the nature of trends
- Predicting the future
- …

Content courtesy of John Stasko, Ga Tech

Providing Better Analysis

Shneiderman
Information Visualization '02

- Combine computational analysis approaches such as data mining with infovis
  - too often viewed as competitors in past
- Each has something to contribute

Content courtesy of John Stasko, Ga Tech

Issues Influencing Design of Discovery Tools

- Statistical algorithms vs. visual data presentation
  - "Visual presentations can give users a richer sense of what is happening in the data and suggest possible directions for further study." -Shneiderman
- Hypothesis testing vs. exploratory data analysis
  - controlled testing vs. collecting large data sets and searching for interesting patterns

Differing Views

- Hypothesis testing
  - Advocates:
    - By stating hypotheses up front, limit variables and sharpens thinking, more precise measurement
  - Critics:
    - Too far from reality, initial bias toward finding hypotheses
- Exploratory Data Analysis
  - Advocates:
    - Find the interesting things this way, we now have computational capabilities to do this
  - Skeptics:
    - Not generalizable, everything is a special case, detecting statistical relationships does not infer cause and effect

Recommendations

- Integrate data mining and information visualization
- Allow users to specify what they are seeking
- Recognize that users are situated in a social context
- Respect human responsibility

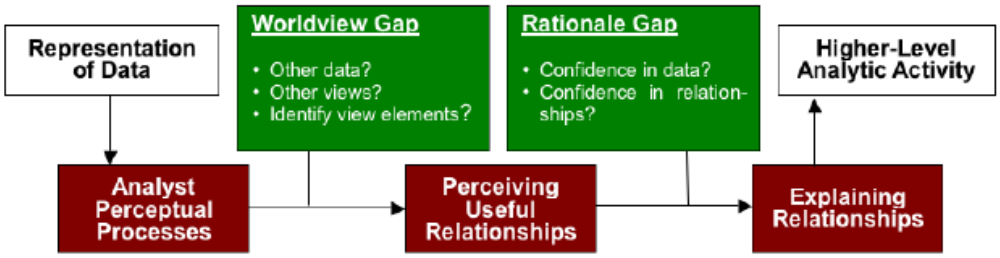
Status Quo Limitations

Amar & Stasko *TVCG* '05

- Current information visualization systems inadequately support decision making
  - limited affordances
  - predetermined representations
  - decline of determinism in decision-making
- "Representational primacy" versus "Analytic primacy"

Analytic Gaps

- Analytic gaps – "obstacles faced by visualizations in facilitating higher-level analytic tasks, such as decision making and learning."
  - Worldview Gap
  - Rationale Gap



Content courtesy of John Stasko, Ga Tech

Knowledge Precepts

For narrowing these gaps

- Worldview-Based Precepts ("Did we show the right thing to the user?")
  - Determine Domain Parameters
  - Expose Multivariate Explanation
  - Facilitate Hypothesis Testing
- Rationale-Based Precepts ("Will the user believe what they see?")
  - Expose Uncertainty
  - Concretize Relationships
  - Expose Cause and Effect

Content courtesy of John Stasko, Ga Tech

Application of Precepts
Expose Uncertainty

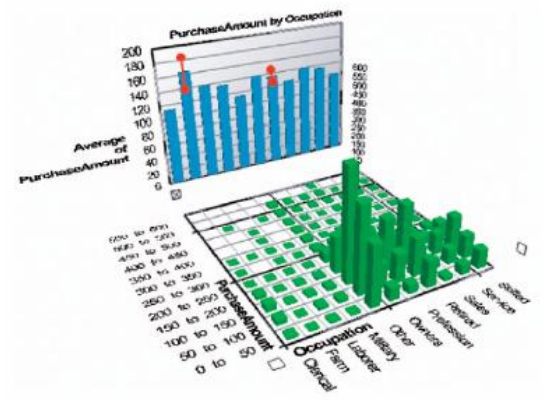


Fig. 2. Error bars (which we have added in red) would be a simple way to increase confidence in the degree of difference between two aggregations. (Picture taken from the SeeIt system by Visible Decisions, Inc.)

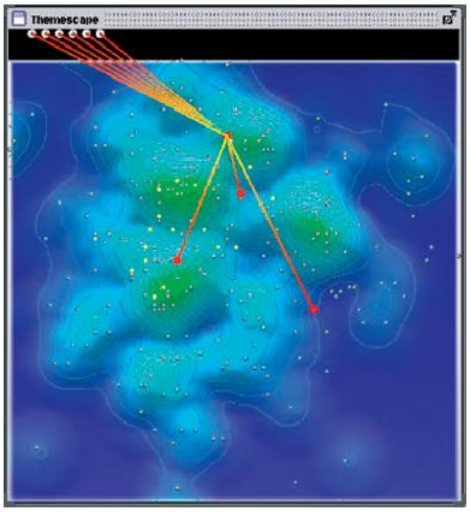


Fig. 3. This themescape variation allows documents with missing metadata, shown as dots in the upper black region, to participate in analysis, such as the reference relationship shown. (Picture courtesy of Nicholas Diakopoulos.)

Content courtesy of John Stasko, Ga Tech

Application of Precepts
Expose Cause and Effect

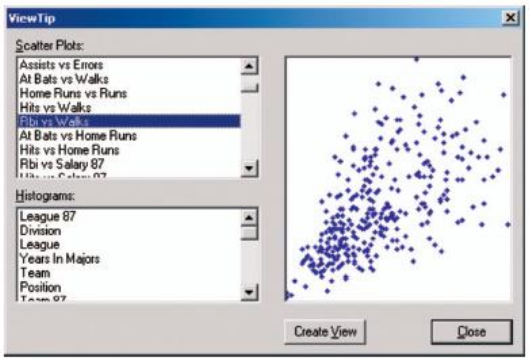


Fig. 4. The View Tips in SpotFire Pro 4.0 allow users to quickly examine possible sources of correlation for further examination.

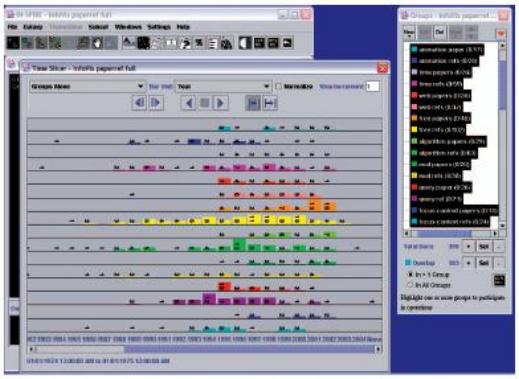


Fig. 5. IN-SPIRE uses horizontal scrolling to navigate time slices of user-defined content groups. (Picture produced at and provided with permission of Pacific Northwest National Laboratory, which is managed and operated by the Battelle Memorial Institute on behalf of the US Department of Energy.)

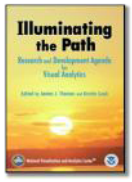
Content courtesy of John Stasko, Ga Tech

Put Them Together

- Combine the ideas:
  - Use computational, statistical analysis more
  - Cater to the user's analytic reasoning needs
- And put together with infovis
- Leads to…

Visual Analytics

- "The science of analytical reasoning facilitated by interactive visual interfaces"
- Combines
  - Data analysis
  - Infovis
  - Analytical reasoning



Thomas & Cook
Illuminating the Path

Visual Analytics

- Grew from view that infovis was neglecting these other aspects
- Grew from stimulus in the homeland security area
  - Need for better data analysis methods
  - Really big data
- For more info, see last year's seminar
  - http://www.cs.odu.edu/~mweigle/CS796-S11/

Content courtesy of John Stasko, Ga Tech

Outline

- Cognition
- User Tasks
- Analytic Tasks

Next Up:
Time Series

Reading: TBD