### Lecture 13: User Studies

**Information Visualization**

**CPSC 533C, Fall 2009**

**Tamara Munzner**

UBC Computer Science

**Wed, 28 October 2009**

---

### Readings Covered

- Ware, Appendix C: The Perceptual Evaluation of Visualization Techniques and Systems
- Cognitive Psychology
- Task-Centered User Interface Design, Clayton Lewis and John Rieman, Chapters 0-5.
- Ware: Evaluation Appendix
  - Task-Centered User Interface Design, Clayton Lewis and John Rieman, Chapters 0-5.
  - The challenge of information visualization evaluation. Catherine Plaisant.
  - Proc. Advanced Visual Interfaces (AVI) 2004

### Further Readings

- Task-Centered User Interface Design, Clayton Lewis and John Rieman, Chapters 0-5.

### Ware: Evaluation Appendix

- perceptual evaluation of infovis techniques and systems
- empirical research methods applied to vis
- difficult to isolate evaluation to perception
- research method depends on research question and object under study

---

### Psychophysics

- method of limits
- find limitations of human perceptions
- error detection methods
- find threshold of performance degradation
- staircase procedure to find threshold faster
- method of adjustment
- find optimal level of stimuli by letting subjects control the level

### Cognitive Psychology

- repeating simple, but important tasks, and measure reaction time or error
- Miller’s 7+/- 2 short-term memory experiments
- Fitts’ Law (target selection)
- Hick’s Law (decision making given n choices)
- interference between channels
- multi-modal studies
- using haptic feedback for interruption when the participants were visually (and cognitively) busy

### Structural Analysis

- requirement analysis, task analysis
- structured interviews
- can be used almost anywhere, for open-ended questions and answers
- rating/Likert scales
- commonly used to solicit subjective feedback
- ex: NASA-TLX (Task Load Index) to assess mental workload
- “it is frustrating to use the interface”
  - Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree
- measurement: faster
- objects of comparison:
  - coordinated O+D display
  - uncoordinated O display
  - uncoordinated D display
- condition of comparison: task requires reading details

### Comparative User Studies

- study design: factors and levels
- factors
  - independent variables
  - ex: interface, task, participant demographics
- levels
  - number of variables in each factor
  - limited by length of study and number of participants

- study design: within, or between?
  - within
    - everybody does all the conditions
    - can lead to ordering effects
    - can account for individual differences and reduce noise
    - thus can be more powerful and require fewer participants
    - combinatorial explosion
    - severe limits on number of conditions
  - between
    - divide participants into groups
    - each group does only some conditions

- study design: within, or between?
  - within
    - everybody does all the conditions
    - can lead to ordering effects
    - can account for individual differences and reduce noise
    - thus can be more powerful and require fewer participants
    - combinatorial explosion
    - severe limits on number of conditions
    - possible workaround is multiple sessions
    - possible order effects
  - between
    - divide participants into groups
    - each group does only some conditions

### Evaluation Throughout Design Cycle

- user/task centered design cycle
- iterative design process
- benchmarking
- deployment
- identify problems, go back to previous step

### Initial Assessments

- what kind of problems are the system aiming to address?
- ex: a large and complex dataset
- who are your target users?
- data analysts
- what are the tasks? what are the goals?
- find trends and patterns in the data via exploratory analysis
- what are their current practices
- statistical analysis
- why and how can visualization be useful?
- visual spotting of trends and patterns
- talk to the users, and observe what they do
- task analysis

### Iterative Design Process

- does your design address the users’ needs?
- can they use it?
- where are the usability problems?
- evaluate without users
  - cognitive walkthrough
  - action analysis
  - heuristics analysis
- evaluate with users
  - usability evaluations (think-aloud)
  - bottom-line measurements

### Benchmarking

- how does your system compare to existing ones?
- empirical, comparative studies
- ask specific questions
- compare an aspect of the system with specific tasks
  - Amar/Stake task taxonomy paper
  - quantitative, but limited
### Results
- Small multiples more accurate than animation
- Animation faster for presentation, slower for analysis
- Than small multiples and trends

### Design
- 2 use: presentation vs. analysis (between-subjects)
- 3 vis encodings: animation vs. traces vs. small mults
- 2 dataset size: small vs. large
- 3 encoding x 2 size: within-subjects
- 24 tasks per participant
  - 3 vs. 7 tasks x 2 trials

### Critique
- First study of macro/micro effects
- Breaking new ground
- Many possible followups
- Physical navigation vs. virtual navigation
  - The Effects of Peripheral Vision and Physical Navigation in Large Scale Visualization. GI 08
  - Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays. CHI 07

### Animation for Trends
- Gapminder: animated bubble charts + human
  - X/y position, size, color, animation
  - Is animation effective?
  - Presentation vs analysis
  - Trend vs transitions

### Trends
- Many countertrends lost in clutter
- Individual plots get small

### Small Multiples
- Many counter-trends lost in clutter
- Individual plots get small
Critique

- video coding is huge amount of work, but very illuminating
- untying complex story of real tool use
- methodology of CTA construction not discussed here
- often bottomup/topdown mix

Experiment 1

- how many bands? mirrored or offset?
- design: within-subjects
  - 2 chart types: mirrored, offset
  - 3 band counts: 2, 3, 4
  - 16 trials per condition
  - 96 trials per subject

Results

- found crossover point where 2-band better: 24 pixels
- virtual resolution: unmirrored unlayered height
  - line: 1x, 1band: 2x, 2band: 4x
- guidelines
  - mirroring is safe
  - layering (position) better than color alone
  - 24 pixels good for line charts, 1band mirrors
  - 12 or 16 pixels good for 2band

Cognitive Task Analysis

- initialize understanding of large scale weather
- build qualitative mental model (QMM)
- verify and adjust QMM
- write the brief
- task breakdown part of paper contribution

Coding Methodology

- interface
  - which interface used
  - whether picture/chart/graph
- usage (every utterance!)
  - goal
  - extract
    - quant/qual
    - goal-oriented/opportunistic
    - integrated/unintegrated
  - brief-writing
    - quant/qual
    - QMM/vis/notes

Proposals

- due 5pm this Fri (Oct 30) by emailing me a URL
- Subject: 533 submit proposal

Format: PDF great, HTML ok, Word acceptable

Credits

- Heidi Lam guest lecture
  - [http://www.cs.ubc.ca/~tmn/courses/cpsc533c-08-fall/#!lec10](http://www.cs.ubc.ca/~tmn/courses/cpsc533c-08-fall/#!lec10)
Presentations

- Days/topics now posted
- Seed papers posted for first day
- Rest up soon

- Slides required, PPT or PDF
  - If using my laptop, email me URL by 10am
  - If your own laptop, email me URL by 3:00pm
- You need both summary and critique/synthesis

- Important difference from me: audience hasn’t read papers!

- Grading (probably)
  - Summary 50%
  - Synthesis/critique 20%
  - Style 15%
  - Materials 15%

- 20 min total: 15-17 present, 3-5 questions
  - Must practice to get timing right!