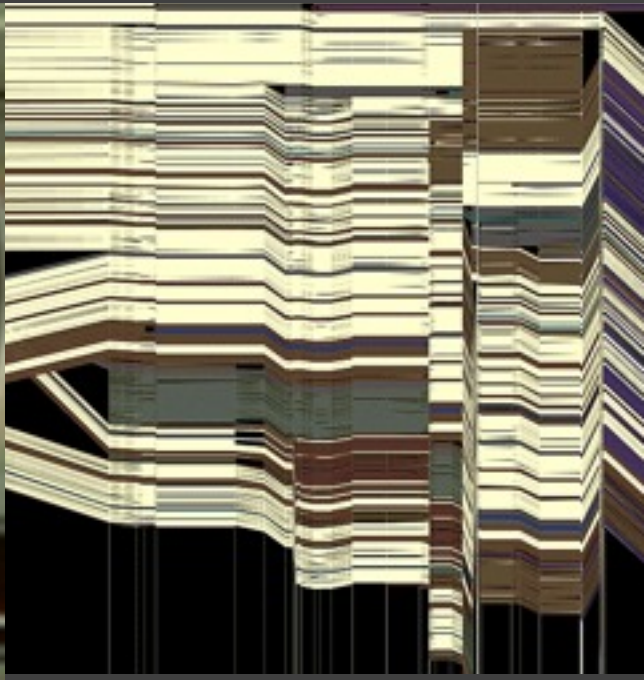
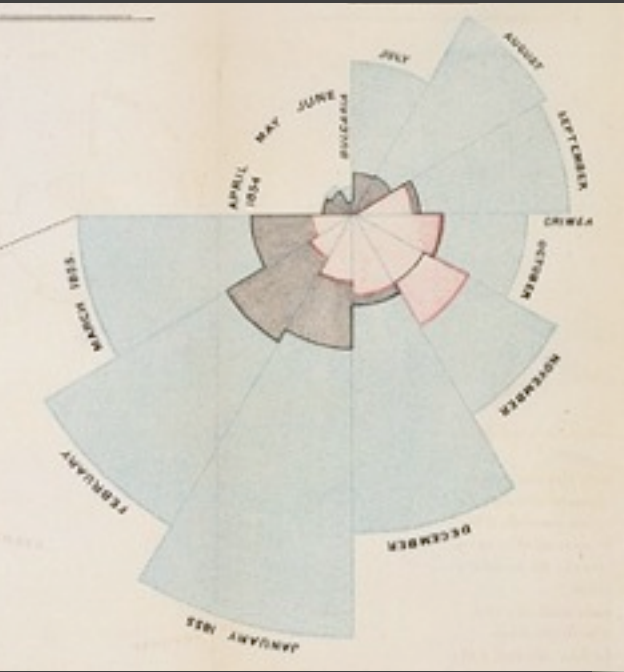


CSE512 :: 13 Mar 2014

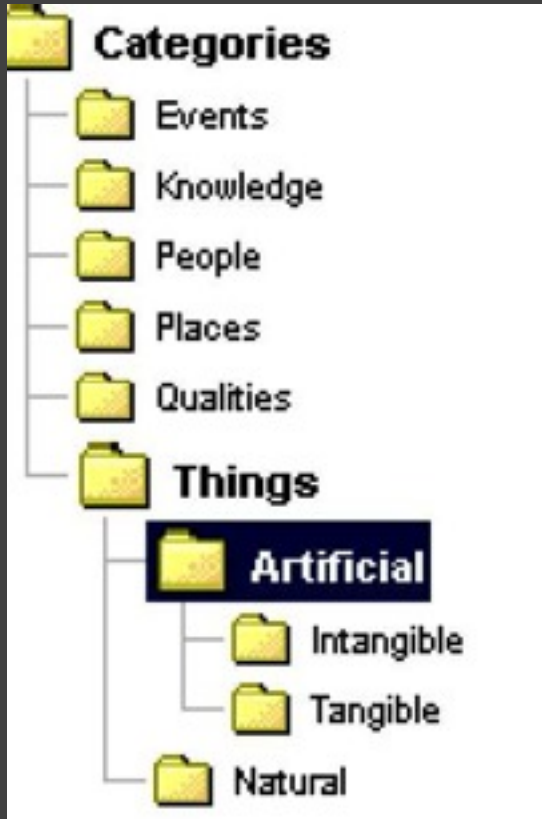
Evaluation



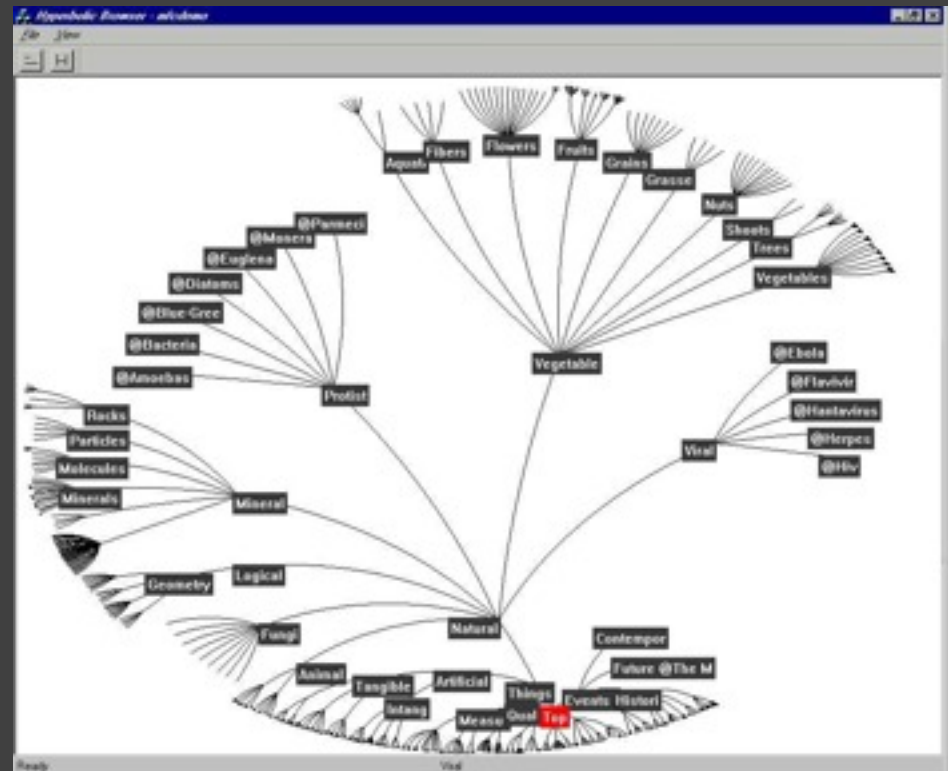
Jeffrey Heer University of Washington

How do we determine if a visualization is “*effective*”?

Example: Tree Browsers



vs.



Evaluation Methods

Inspection or Principled Rationale

Apply design heuristics, perceptual principles

Informal User Study

Have people use visualization, observe results

Controlled Experiment

Choose appropriate tasks / users to compare

Choose metrics (time, error, **what else?**)

Evaluation Methods

Field Deployment or Case Studies

Observation and Interview

Document effects on work practices

Theoretical Analysis

Algorithm time and space complexity

Benchmarks

Performance (e.g., interactive frame rates)

Scalability to larger data sets

Topics

Evaluating Focus + Context Visualizations

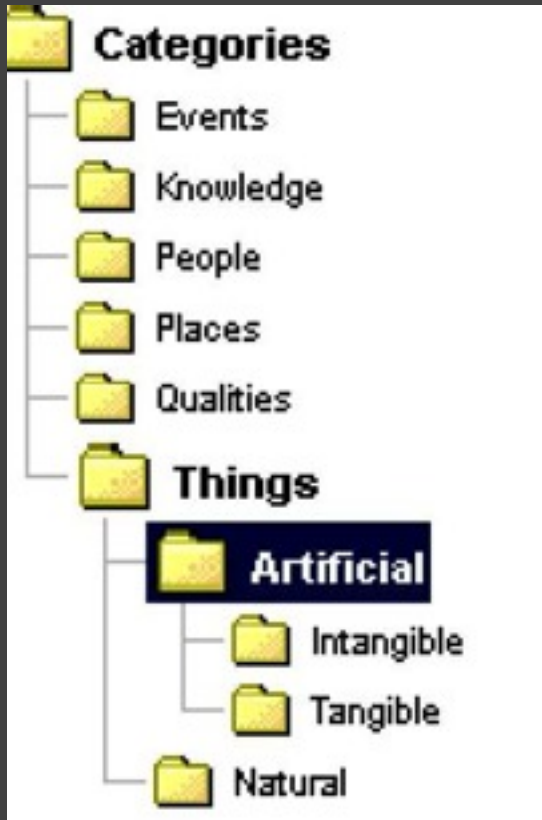
Evaluating Time-Series Visualizations

Perceptual Organization of Graphs

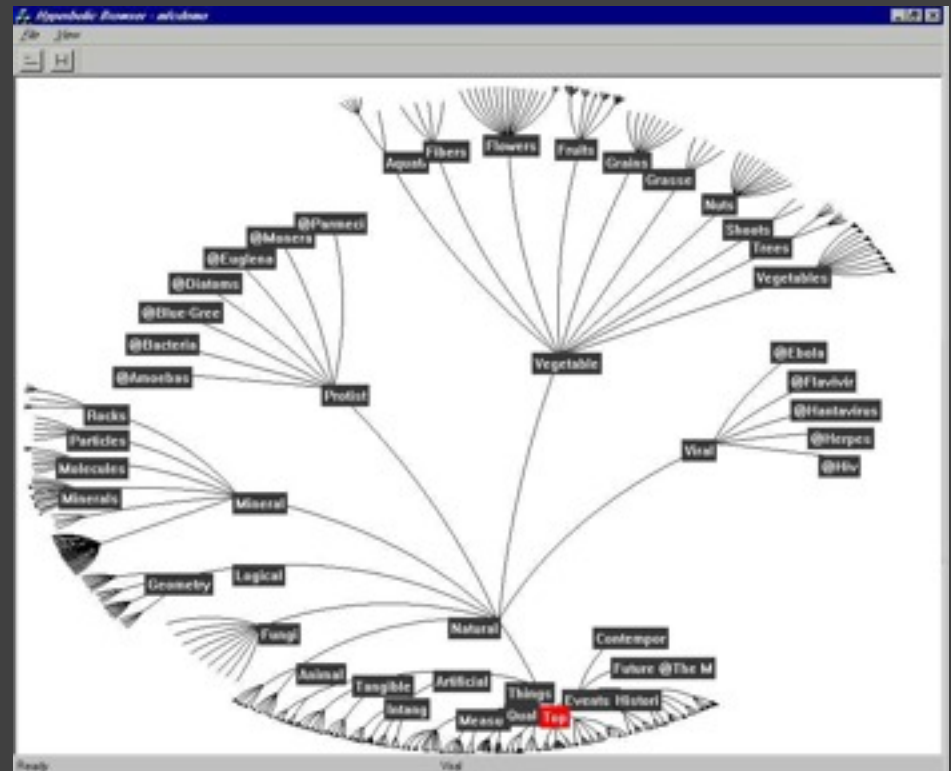
Contextual Effects on Visualization Use

Discussion and Course Evaluation

The Great CHI'97 Browse-Off



vs.



The Contest Winner!

Can we conclude that the hyperbolic tree is the better browser?

No. At least not yet...

Different people operating each browser.

“Is it the car or the driver?”

Tasks were not ecologically valid.

Xerox PARC researchers conducted eye-tracking studies to investigate...

[Pirolli, Card, & van der Wege, AVI 2000]

Task Types

Simple retrieval tasks – “Find Lake Victoria”

Complex retrieval tasks – “Which army is led by a Generalissimo?”

Local relational tasks – “Which religion has the most holidays?”

Complex relational tasks – “Which Greek deity has the same name as a space mission?”

Ambiguity and Information Scent

Which tree branch would you follow to answer these questions?

“Find a hammer”

vs.

“What's the highest rank in the British Royal Air Force?”

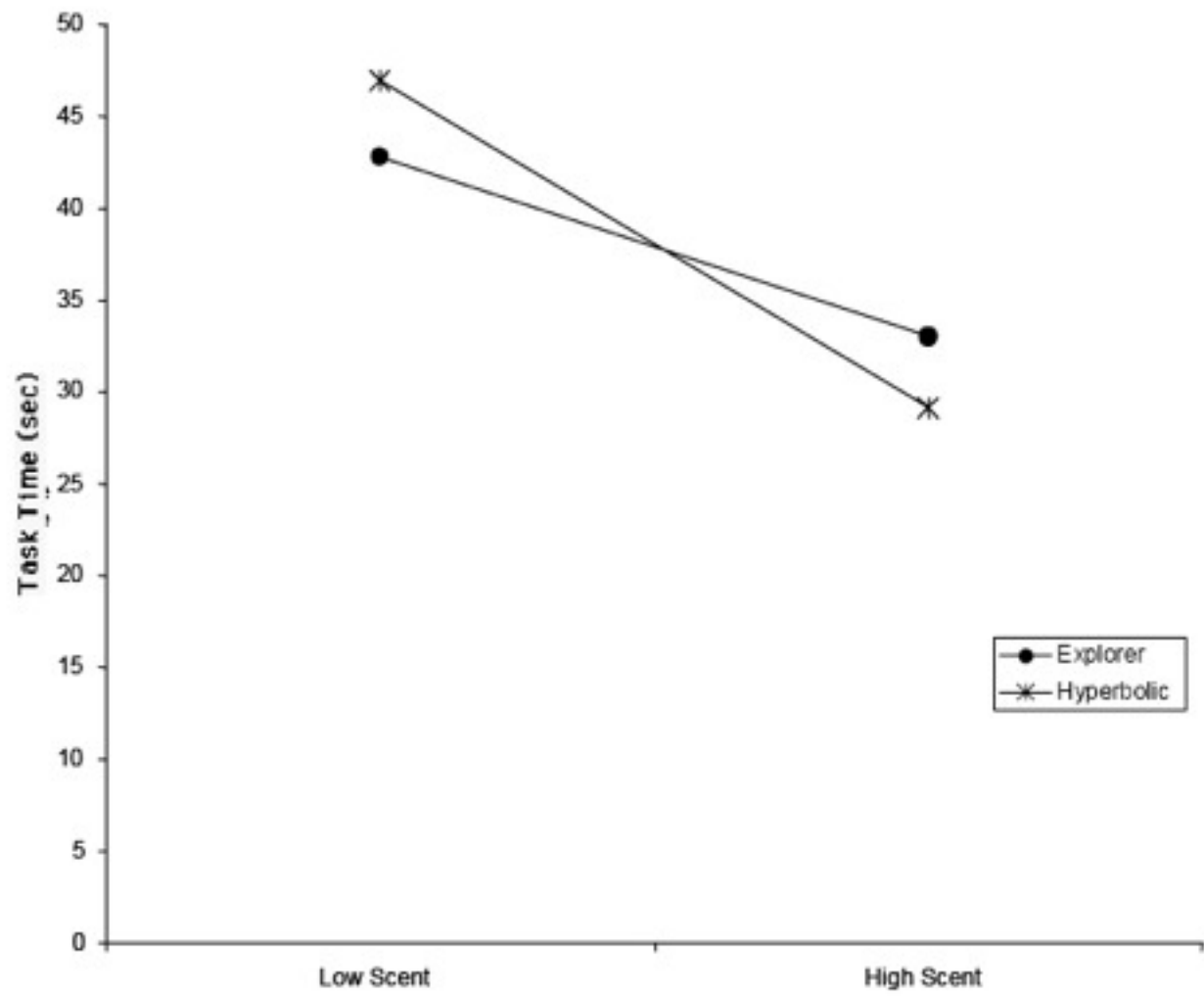
Initial Results: No Difference?

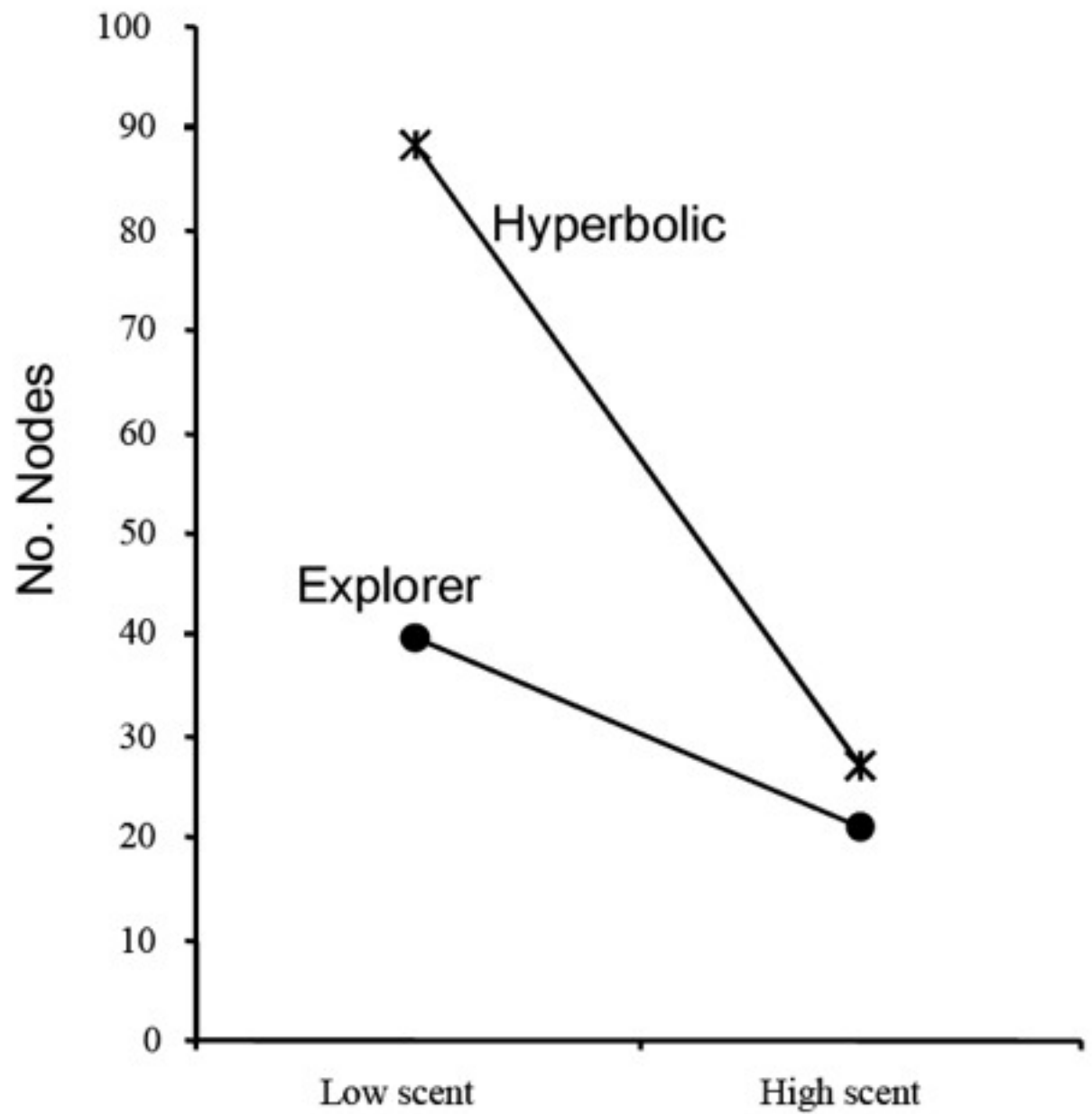
Question Type	Browser	
	Explorer (sec)	Hyperbolic (sec)
Retrieval Tasks		
Simple	35.55	34.37
Complex	41.55	42.02
All retrieval	38.55	38.20
Comparison Tasks		
Local	42.78	41.91
Global	71.07	73.19
All comparison	56.93	57.55
All questions	47.74	47.87

Operationalizing Scent

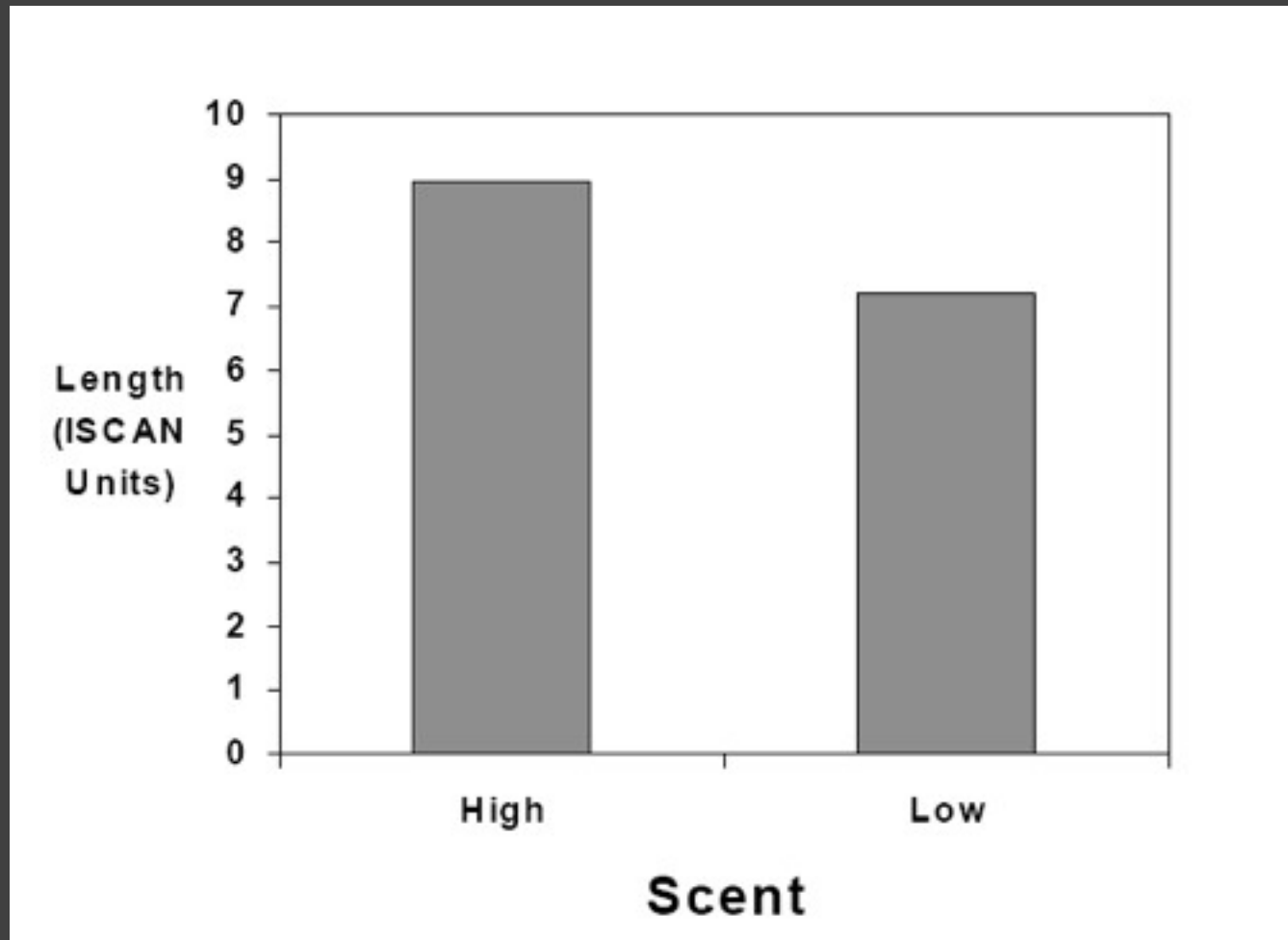
How useful is a text label on a node?

Information scent = the proportion of participants who correctly identified the location of the task answer from looking at upper branches in the tree.

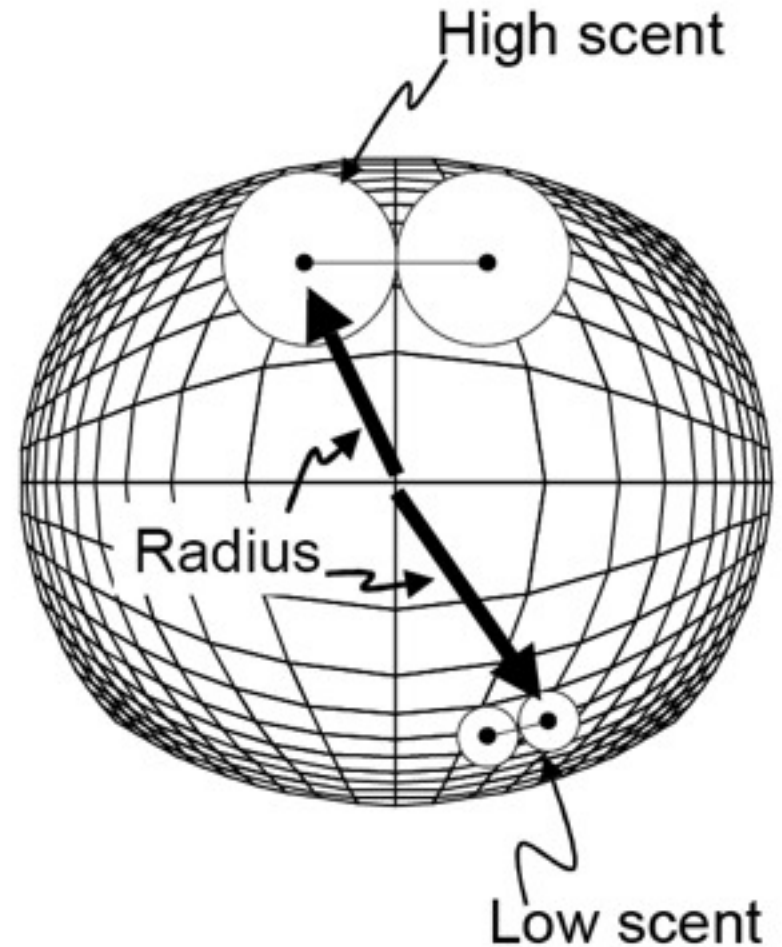
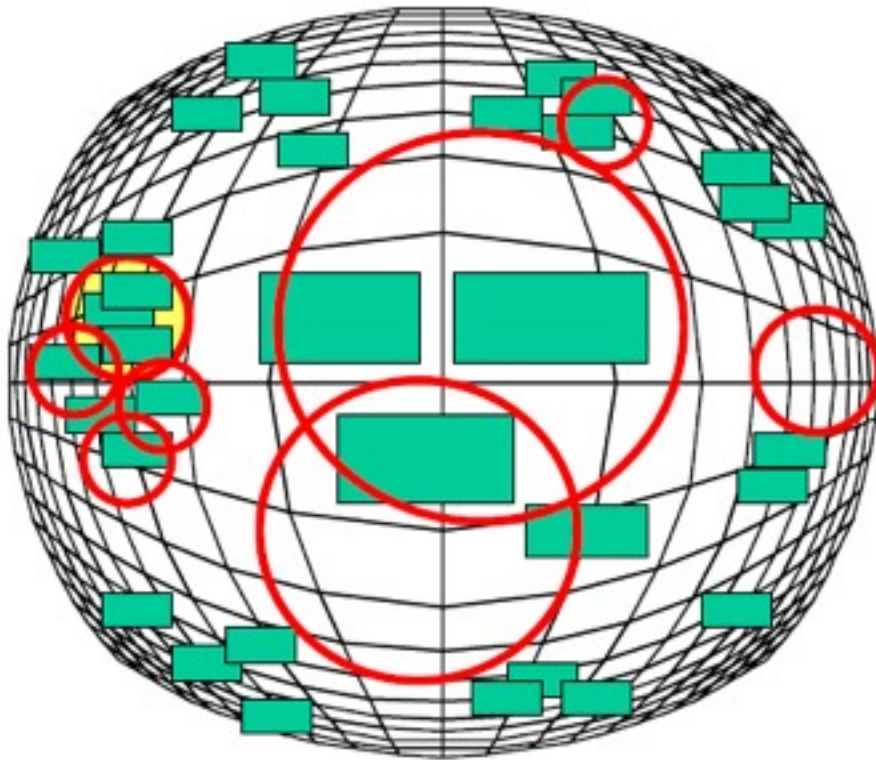




Length of eye movements

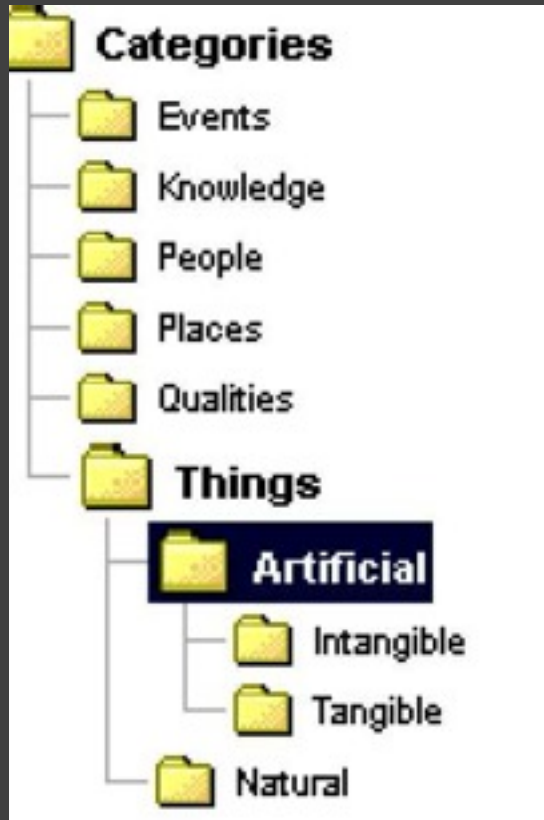


Adaptive Field of View?

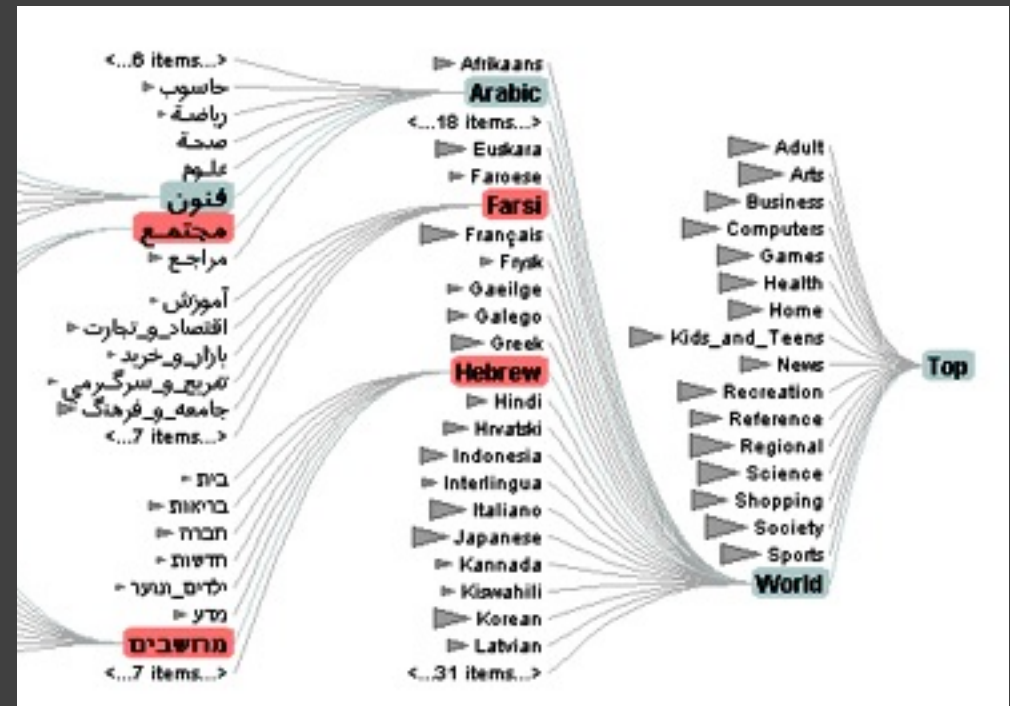


(c)

More Evaluations



vs.



Evaluation of DOI Trees

DOI Tree vs. Windows Explorer [Budiu, AVI 06]

Nodes visited (avg):	DOI 83	Exp 53	$p < .005$
Revisitation (avg):	DOI 6.60	Exp 8.15	$p < .005$
Divergence (avg):	DOI 4.57	Exp 3.96	$p < .001$

DOI Tree more forgiving to navigation errors
BUT no significant difference in task time

DOI Tree vs. Google Directory [Pirolli, CHI 06]

DOI Tree has superior task knowledge transfer

Lessons Learned

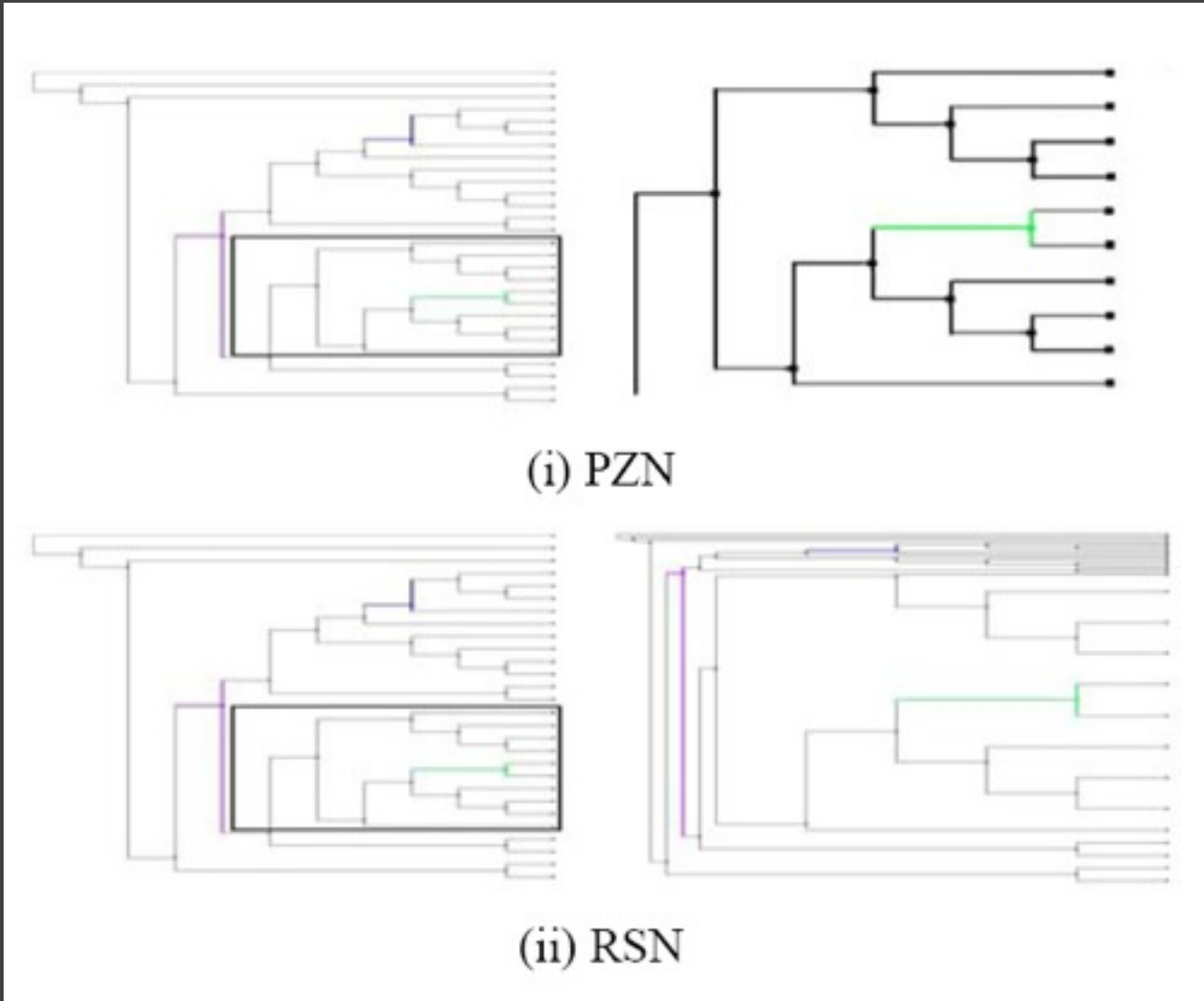
Both the **task** and **data properties** (in this case *information scent*) may interact with the visualization type in unexpected ways.

Equal **performance** in terms of accuracy or response time is **not the whole picture**. In this case, we saw differences in learning effects.

An Evaluation of Pan & Zoom and Rubber Sheet Navigation with and without an Overview

Dmitry Nekrasovski, Adam Bodnar, Joanna McGrenere,
François Guimbretière, Tamara Munzner

Pan and Zoom / Rubber Sheet Navigation



Experimental Task

Compare topological distance between nodes in a dendrogram.

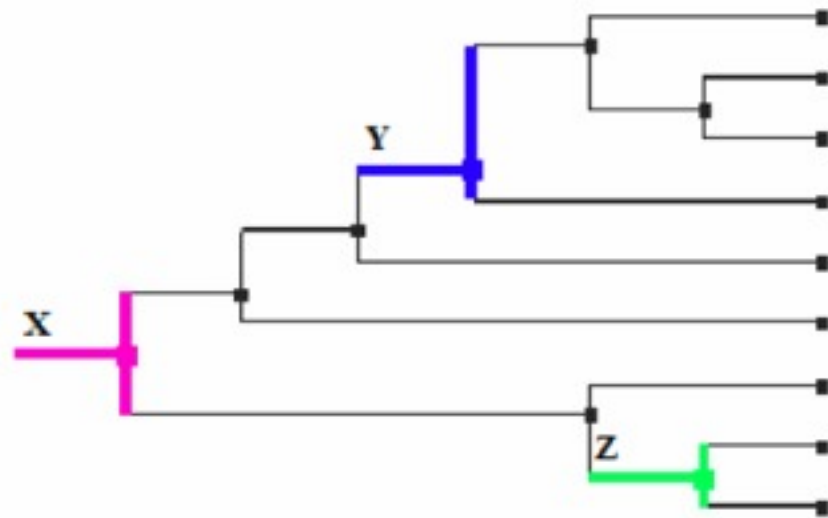


Figure 2: Illustration of the experimental task on a small tree. Subjects were asked to determine whether the pink node (labeled X) was closer to the blue node (labeled Y) or the green node (labeled Z) in terms of topological distance. In this case, the green node (Z) is closer.

Condition 1: Rubber Sheet - No Overview

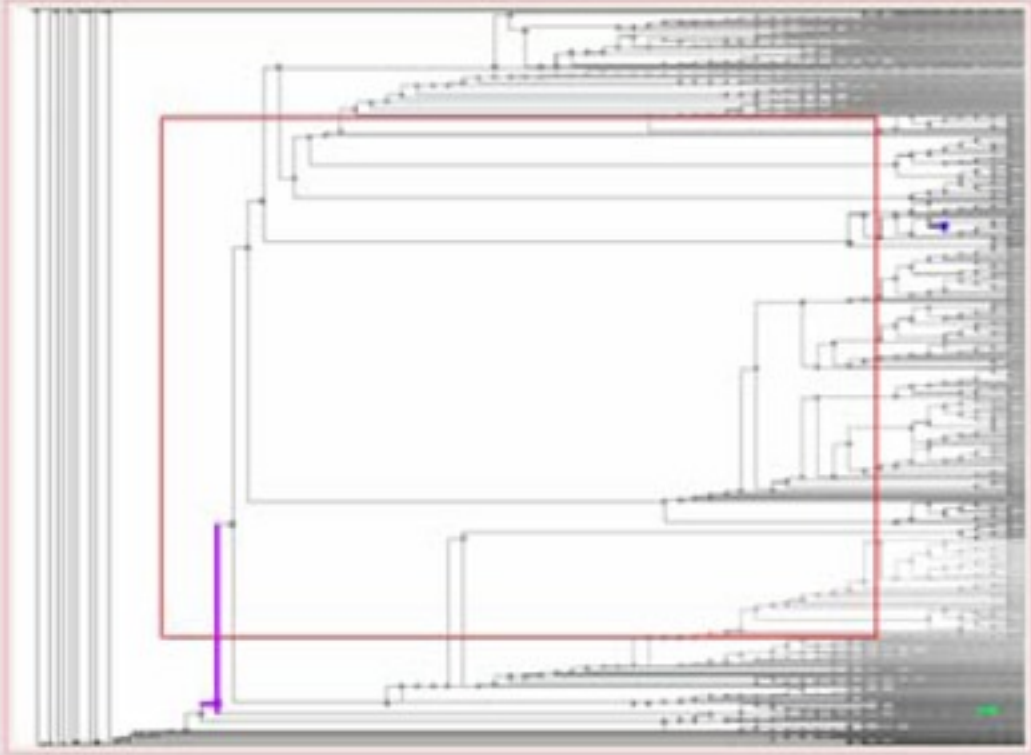
Evaluation 1.2 job, C1, level 1-8

File Edit Tools Help

Which node is the purple node closer to in terms of topological distance?

Blue Green

Drag with LEFT mouse button to ZOOM IN
Drag with RIGHT mouse button to PAN
Press R to RESET the visualization
Press ESCAPE to CLEAR the current mouse drag



The diagram shows a complex network of nodes and edges. A purple node is located at the bottom left, and a green node is located at the bottom right. A red rectangular box highlights a large portion of the network, including the purple node and several other nodes and edges. The network is dense and interconnected, with many nodes and edges visible. The purple node is connected to several other nodes, and the green node is also connected to several other nodes. The red box is positioned over the purple node and extends upwards and to the right, covering a significant portion of the network.

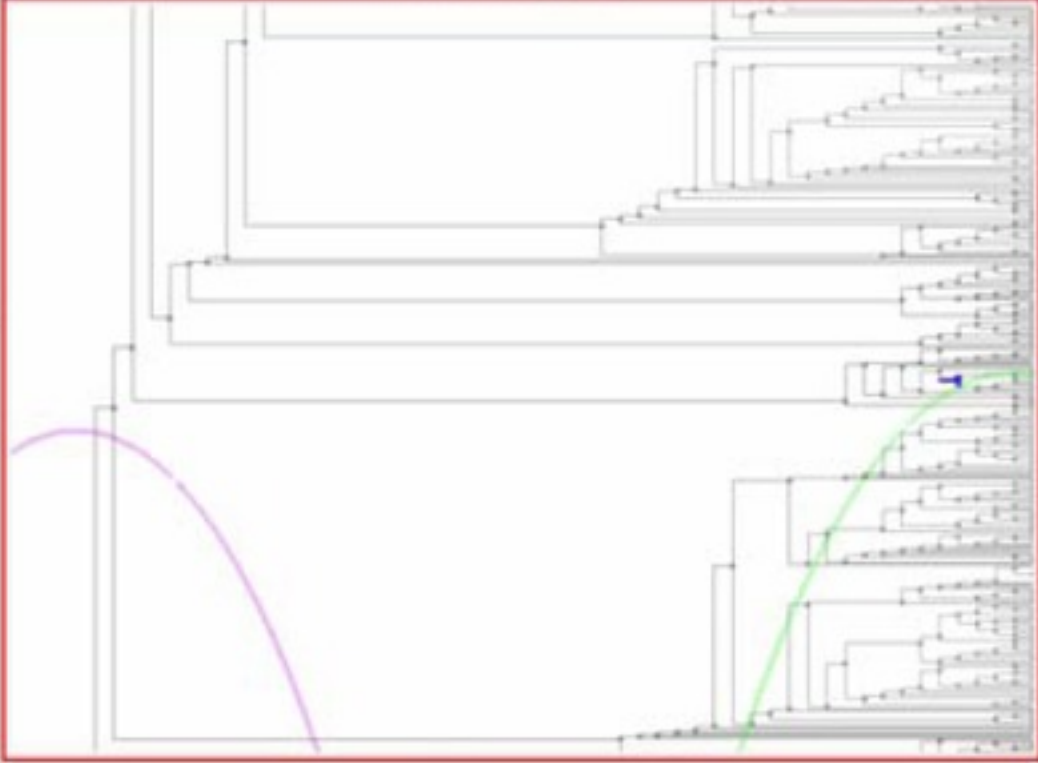
Condition 2: Pan & Zoom - No Overview

Evaluation 18 p6, C2, level = 8

File Edit View Help

Which node is the purple node closer to in terms of topology of distance?

Blue Green



Drag with LEFT mouse button to ZOOM IN
Drag with MIDDLE mouse button to ZOOM OUT
Drag with RIGHT mouse button to PAN

Press R to RESET the visualization
Press ESCAPE to CLEAR the current mouse drag

The image shows a complex phylogenetic tree with many nodes and branches. A purple arc highlights a specific node on the left side of the tree. A green line highlights a path from the bottom right towards the purple node. The tree is enclosed in a red rectangular border. The interface includes a title bar, a menu bar, a question, radio buttons for 'Blue' and 'Green', a 'Submit' button, and a set of instructions on the left side.

Condition 3: Rubber Sheet with Overview

Evaluation 11 pt, C3, level = 3
File View Tools Help

Which node is the purple node closer to in terms of topological distance?

Blue Green

Drag with LEFT mouse button to ZOOM IN
Drag with RIGHT mouse button to PAN
Press R to RESET the visualization
Press ESCAPE to CLEAR the current mouse drag

The screenshot shows a software window titled "Evaluation 11 pt, C3, level = 3" with a menu bar (File, View, Tools, Help). The main area displays a 3D architectural scene of a staircase. A purple path is highlighted on the stairs, and a red rectangular bounding box is drawn around a specific node on the path. A question asks: "Which node is the purple node closer to in terms of topological distance?" Below the question are two radio buttons labeled "Blue" and "Green", and a "Submit" button. On the left side of the interface, there are four instructions: "Drag with LEFT mouse button to ZOOM IN", "Drag with RIGHT mouse button to PAN", "Press R to RESET the visualization", and "Press ESCAPE to CLEAR the current mouse drag". A larger, zoomed-in view of the staircase scene is shown in a separate window on the right, with the same purple path and red bounding box.

Condition 4: Pan & Zoom with Overview

File Edit View Help

Which mode is the purple mode closer to in terms of topological distance?

Blue Green

Drag with LEFT mouse button to ZOOM IN
Drag with MIDDLE mouse button to ZOOM OUT
Drag with RIGHT mouse button to PAN

Press R to RESET the visualization
Press ESCAPE to CLEAR the current mouse drag

Experiment

Compare performance in 4 conditions:

1. Pan and Zoom (no overview)
2. Pan and Zoom (with overview)
3. Rubber Sheet (no overview)
4. Rubber Sheet (with overview)

40 subjects (24F/16M), between 18-39 years old.
Right-handed, normal vision.

Between-subjects design.

Hypotheses

1. RSN interfaces perform better than PZN interfaces independently of the presence or absence of an overview.
2. For RSN, the presence of an overview does not result in better performance.
3. For PZN, the presence of an overview results in better performance.

Results - H1 False

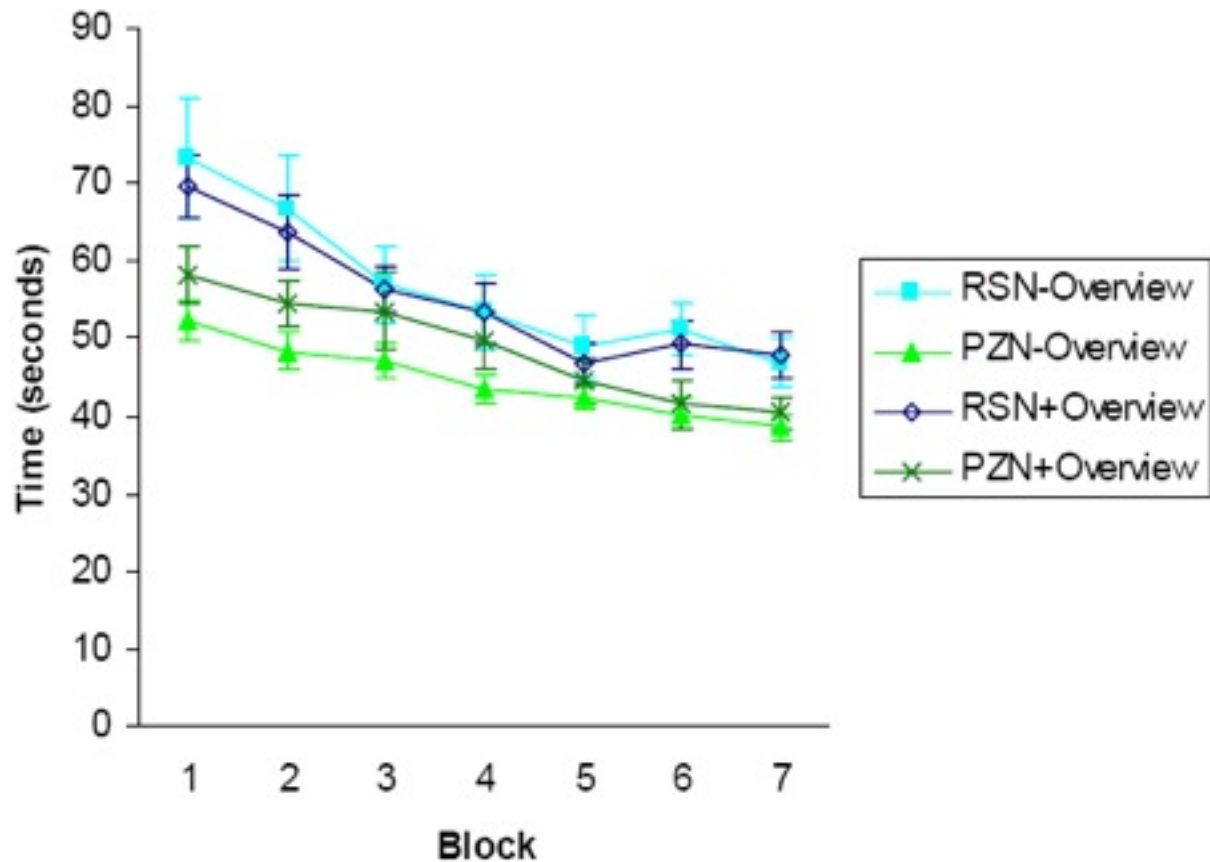


Figure 7: Mean completion times per trial for each interface by block in seconds (N=40).

Results - H2 True, H3 False

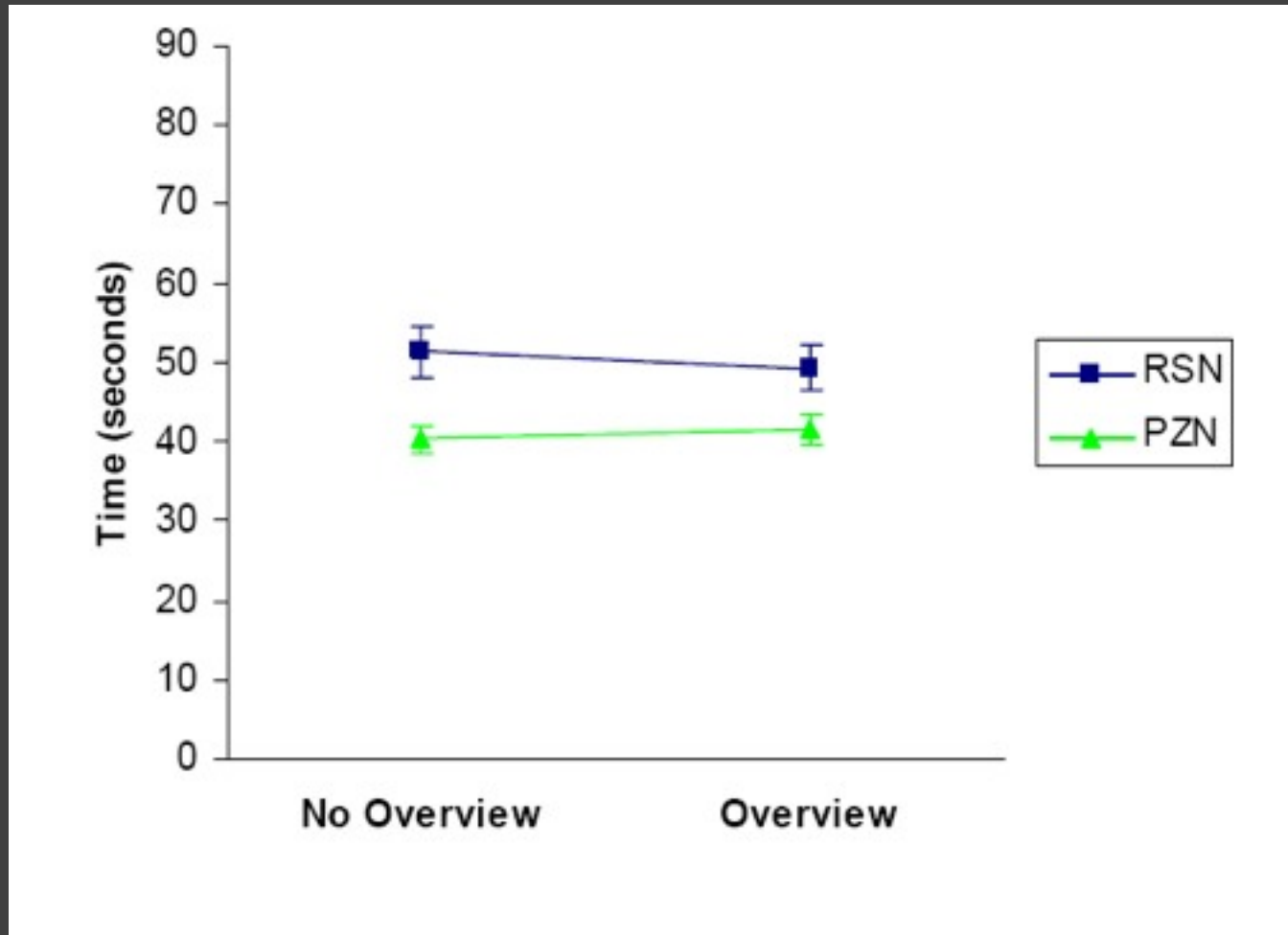


Figure 9: Block 7 mean per-trial completion times in seconds by navigation technique with and without an overview.

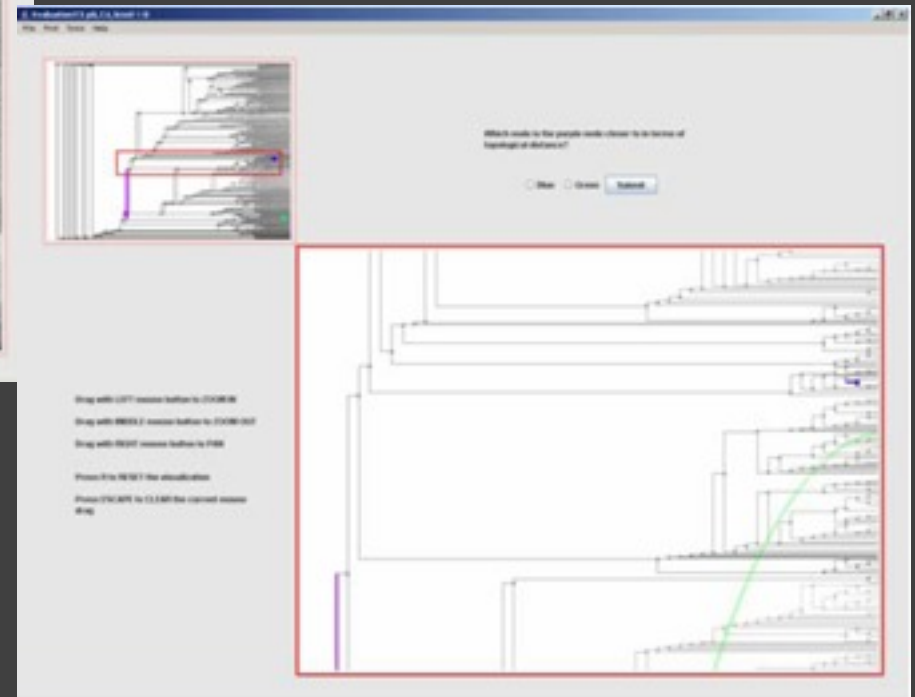
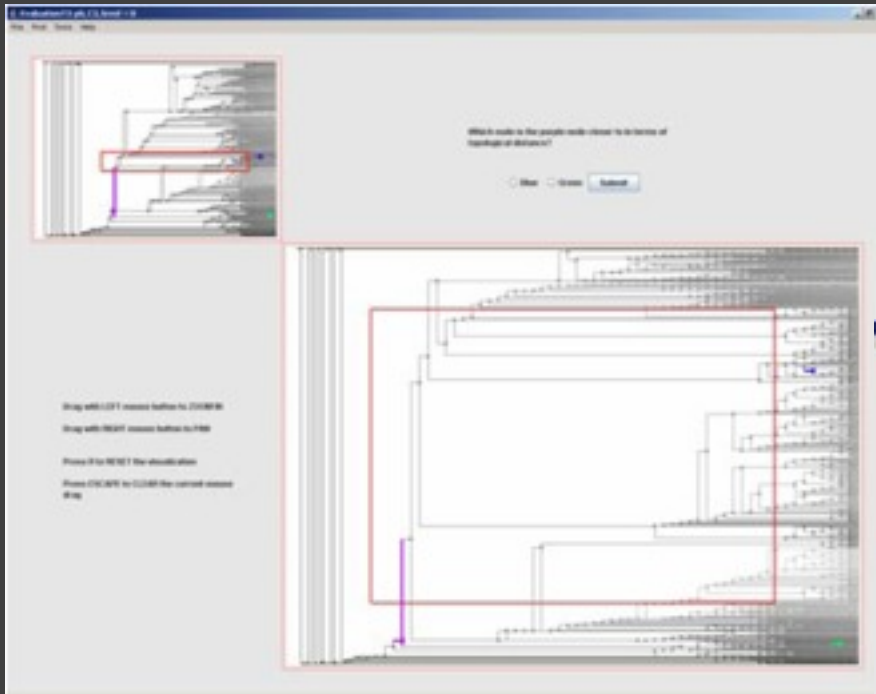
Results

R1. Pan & Zoom had lower completion times, navigation actions, resets, and reported mental demand.

R2. Overview has no significant impact on rubber sheet navigation, though it was reported to reduce physical demand.

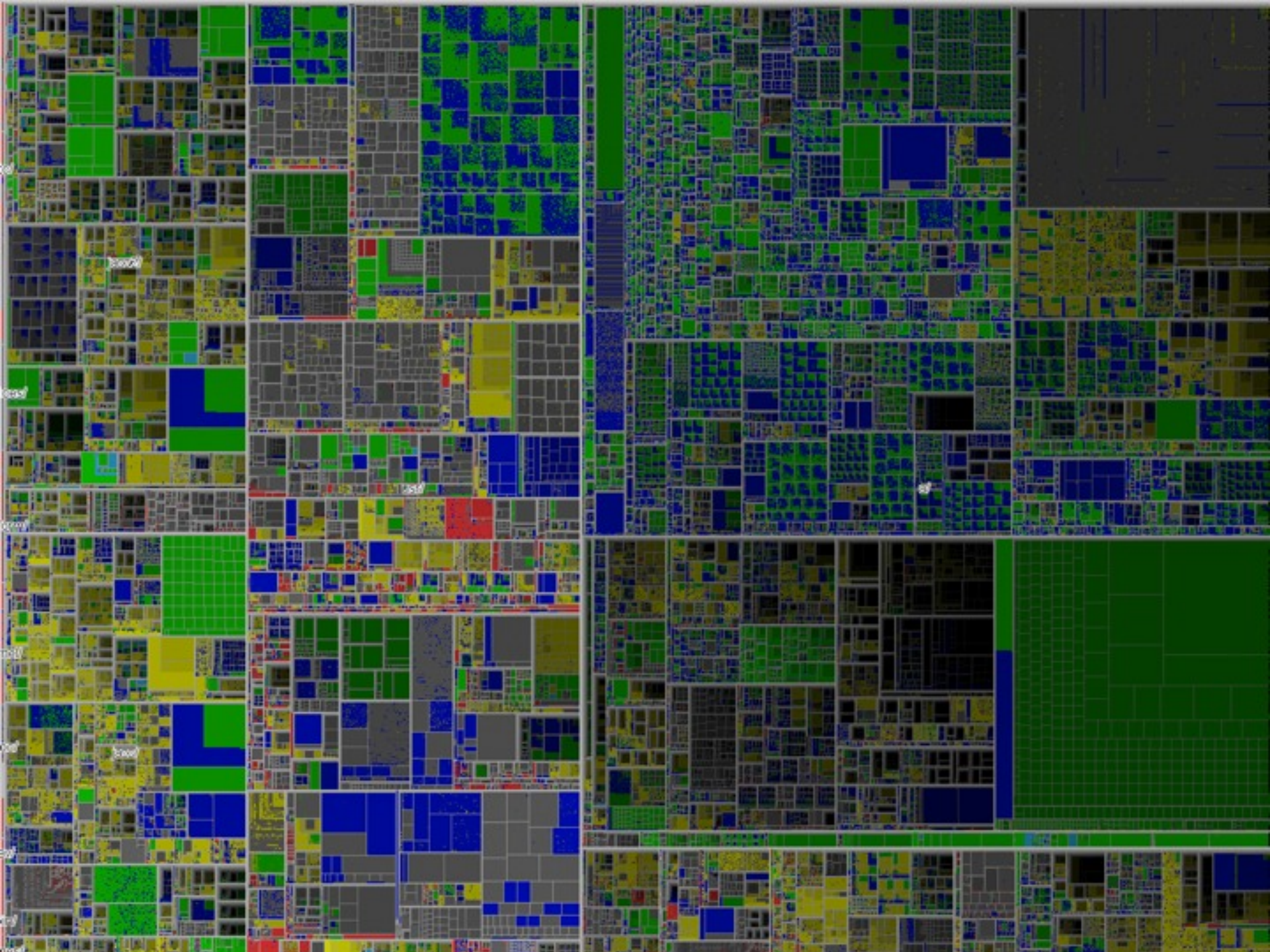
R3. Overview has no significant impact on pan & zoom navigation, though it was reported to reduce physical demand.

Thoughts?



Does this generalize
for overview displays?

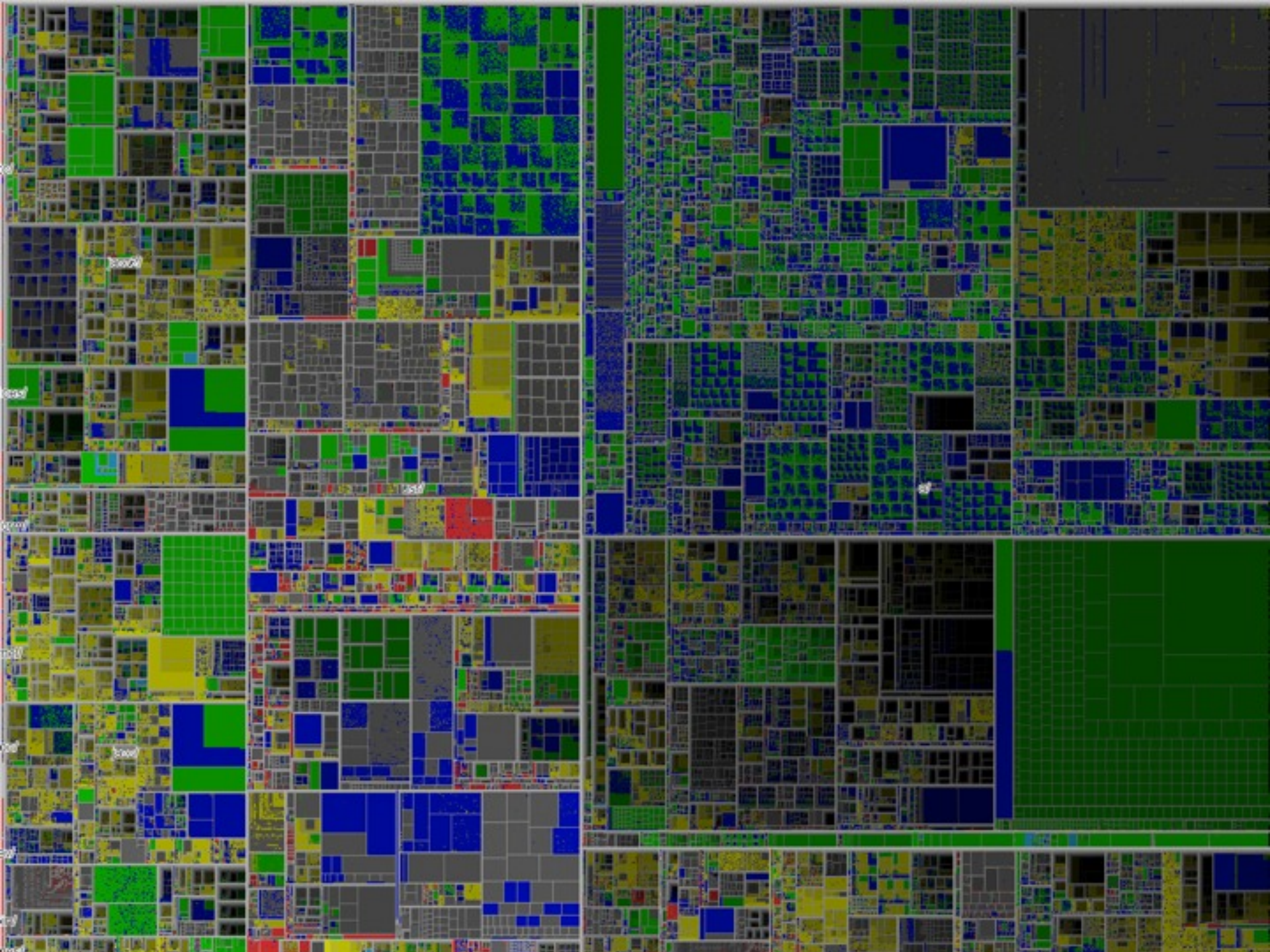
Evaluating Data-Dense Time Series Visualizations

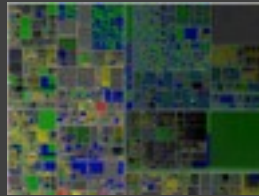


$$\text{Data Density} = \frac{(\# \text{ entries in data})}{(\text{area of graphic})}$$

“Graphical excellence... gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space”

Tufte '83



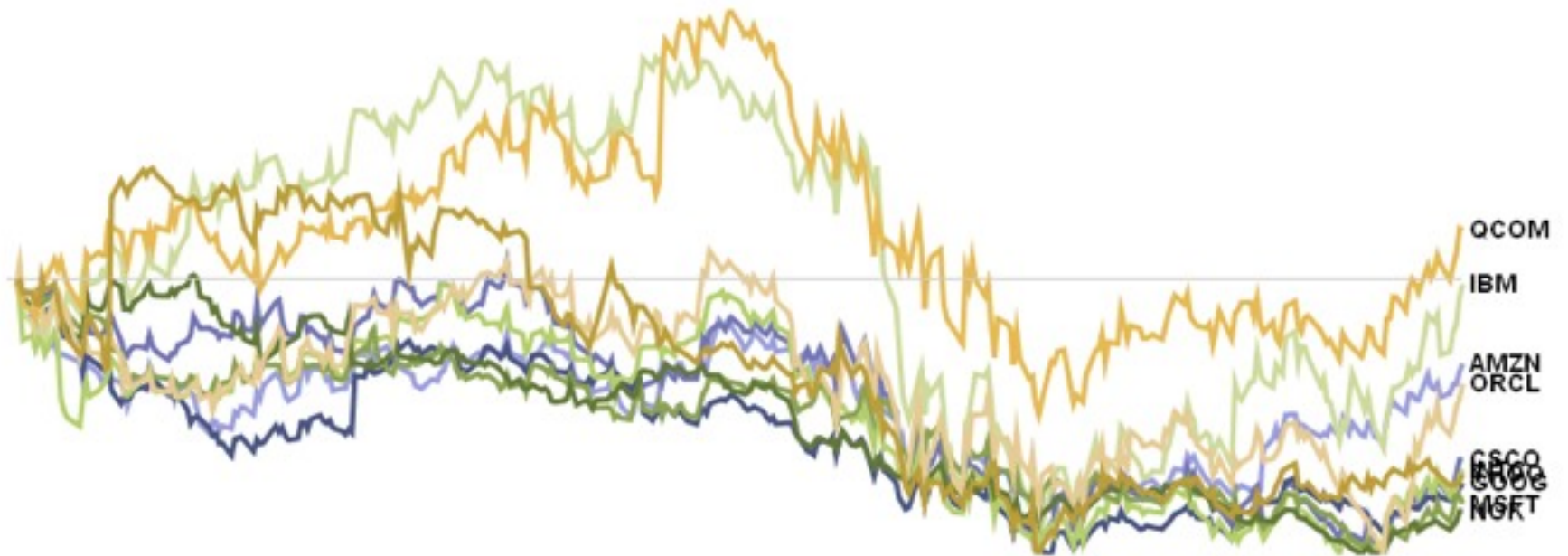


Time-Series Visualization

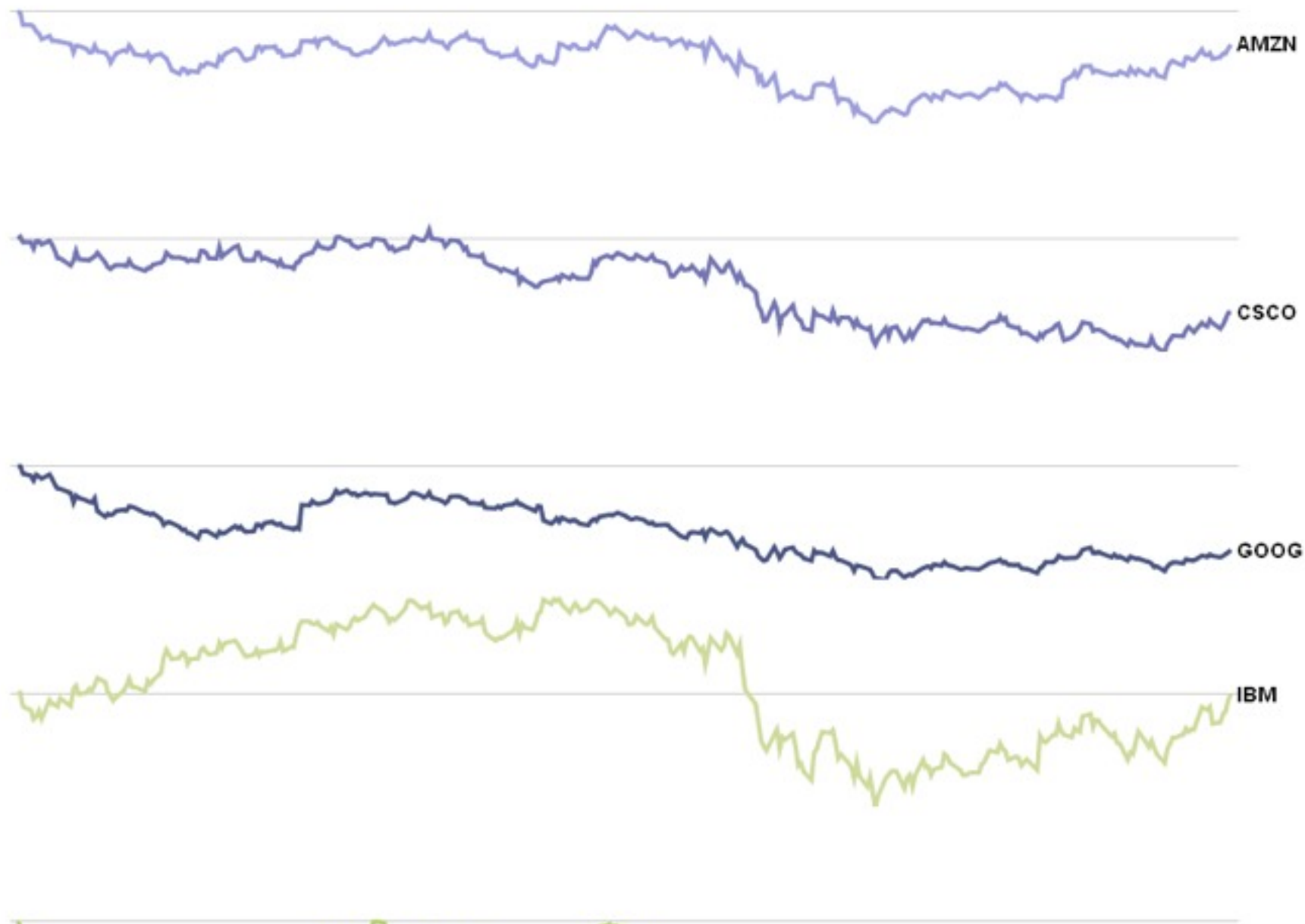
Relative Technology Stock Performance: Jan 2008 - Present



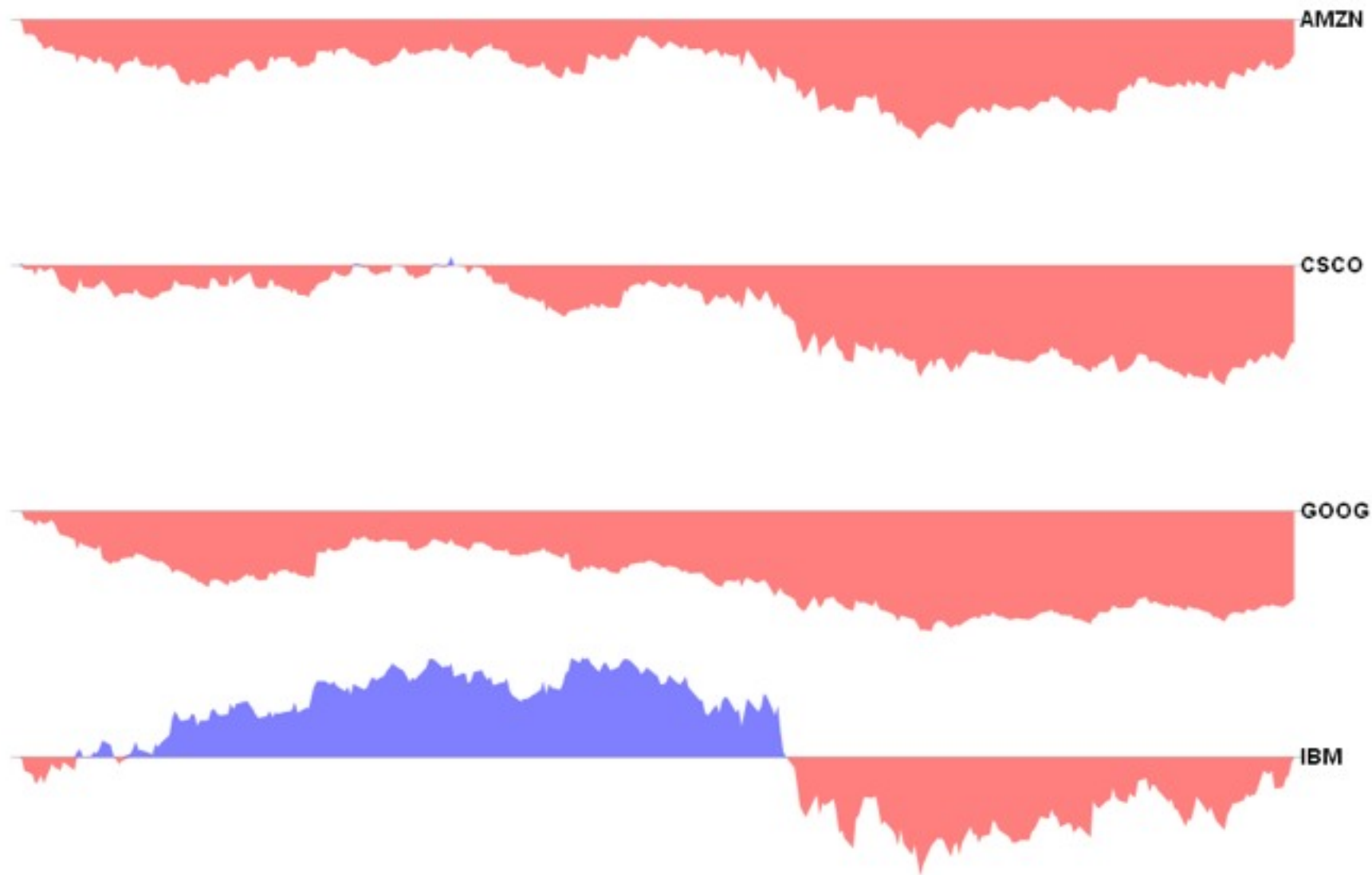
Relative Technology Stock Performance: Jan 2008 - Present



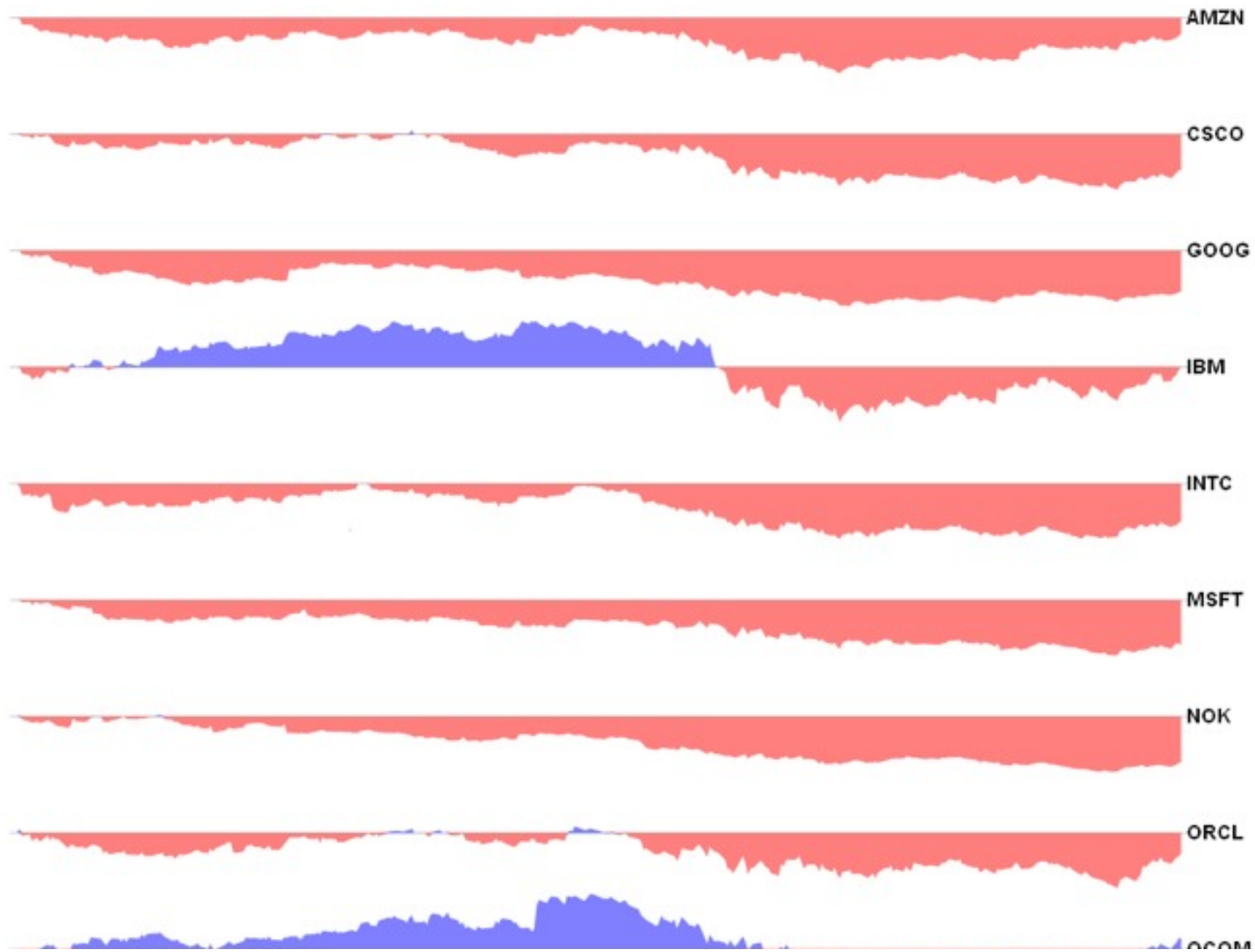
Relative Technology Stock Performance: Jan 2008 - Present



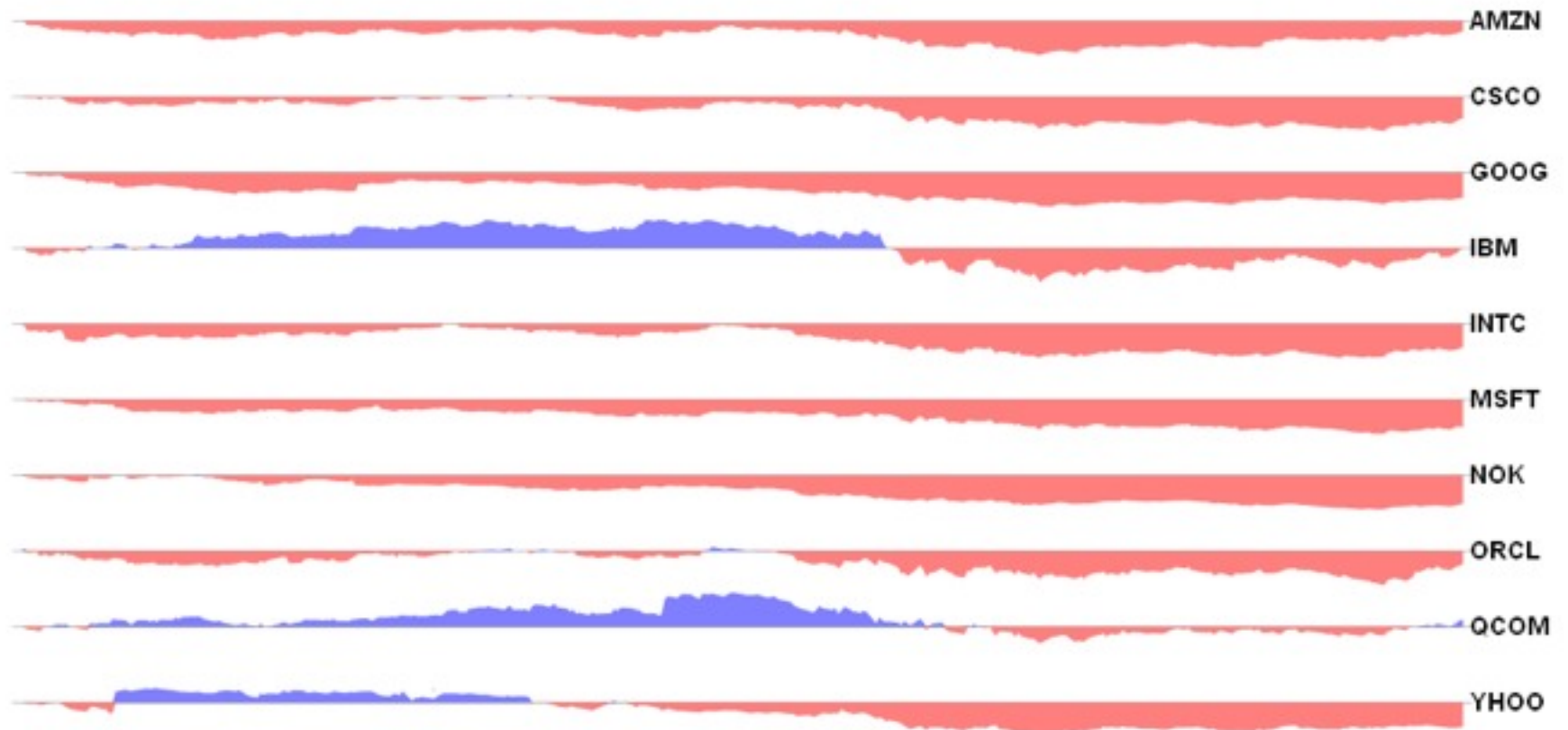
Relative Technology Stock Performance: Jan 2008 - Present



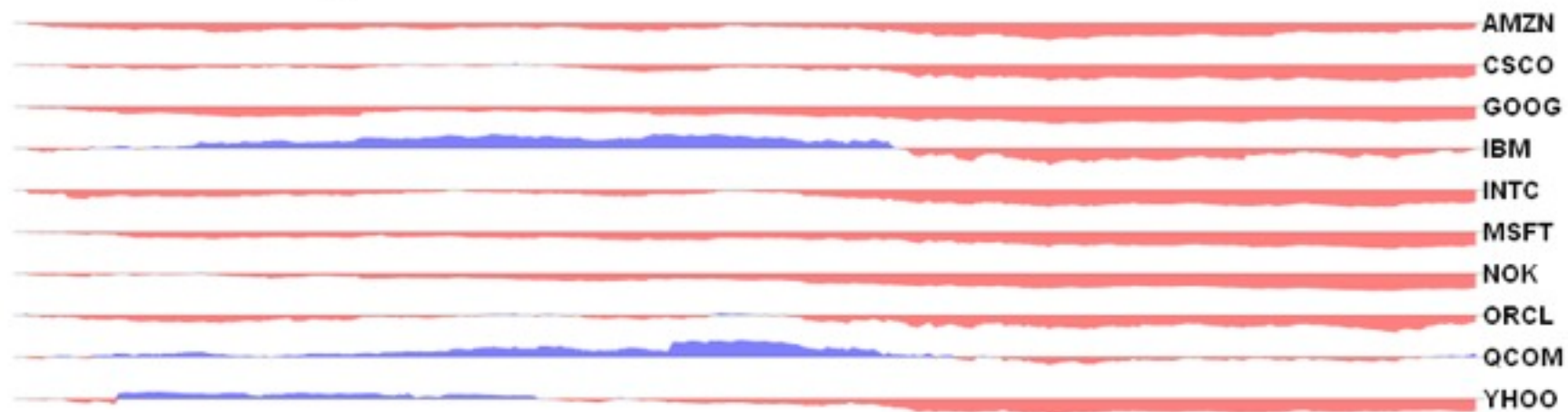
Relative Technology Stock Performance: Jan 2008 - Present



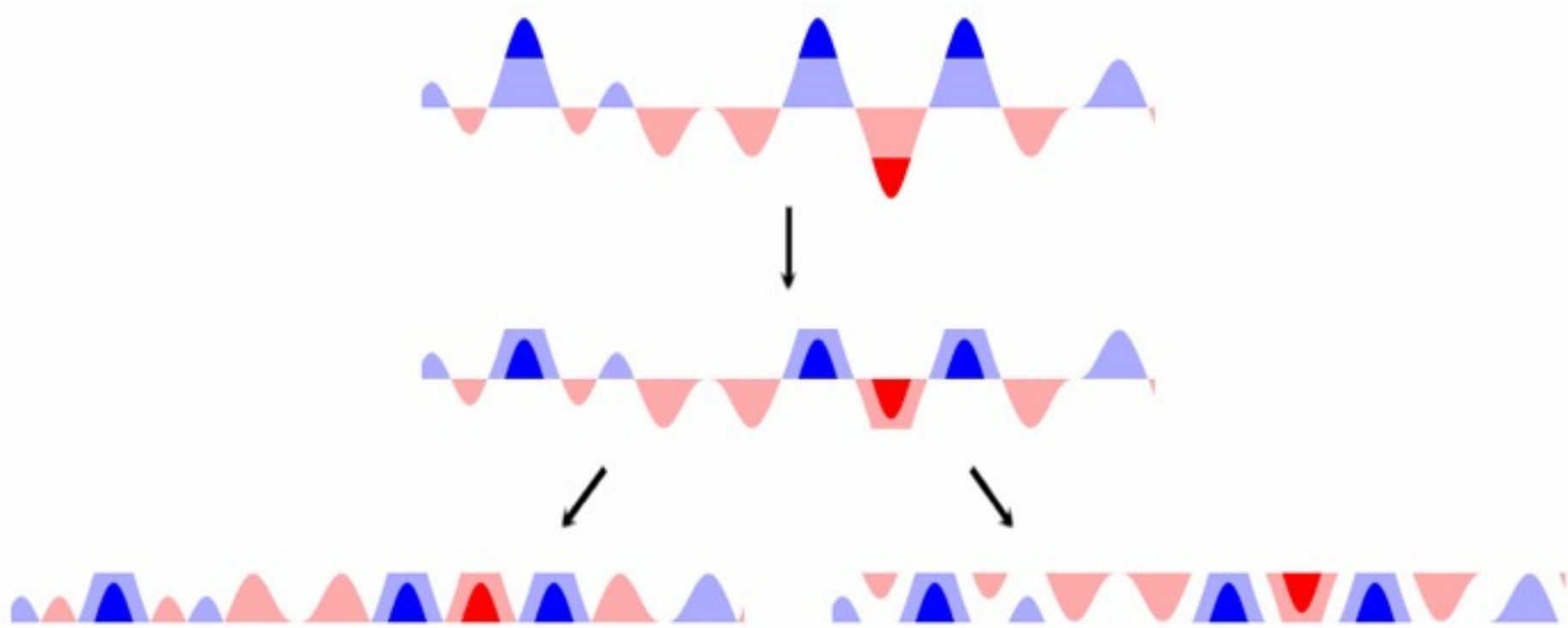
Relative Technology Stock Performance: Jan 2008 - Present



Relative Technology Stock Performance: Jan 2008 - Present



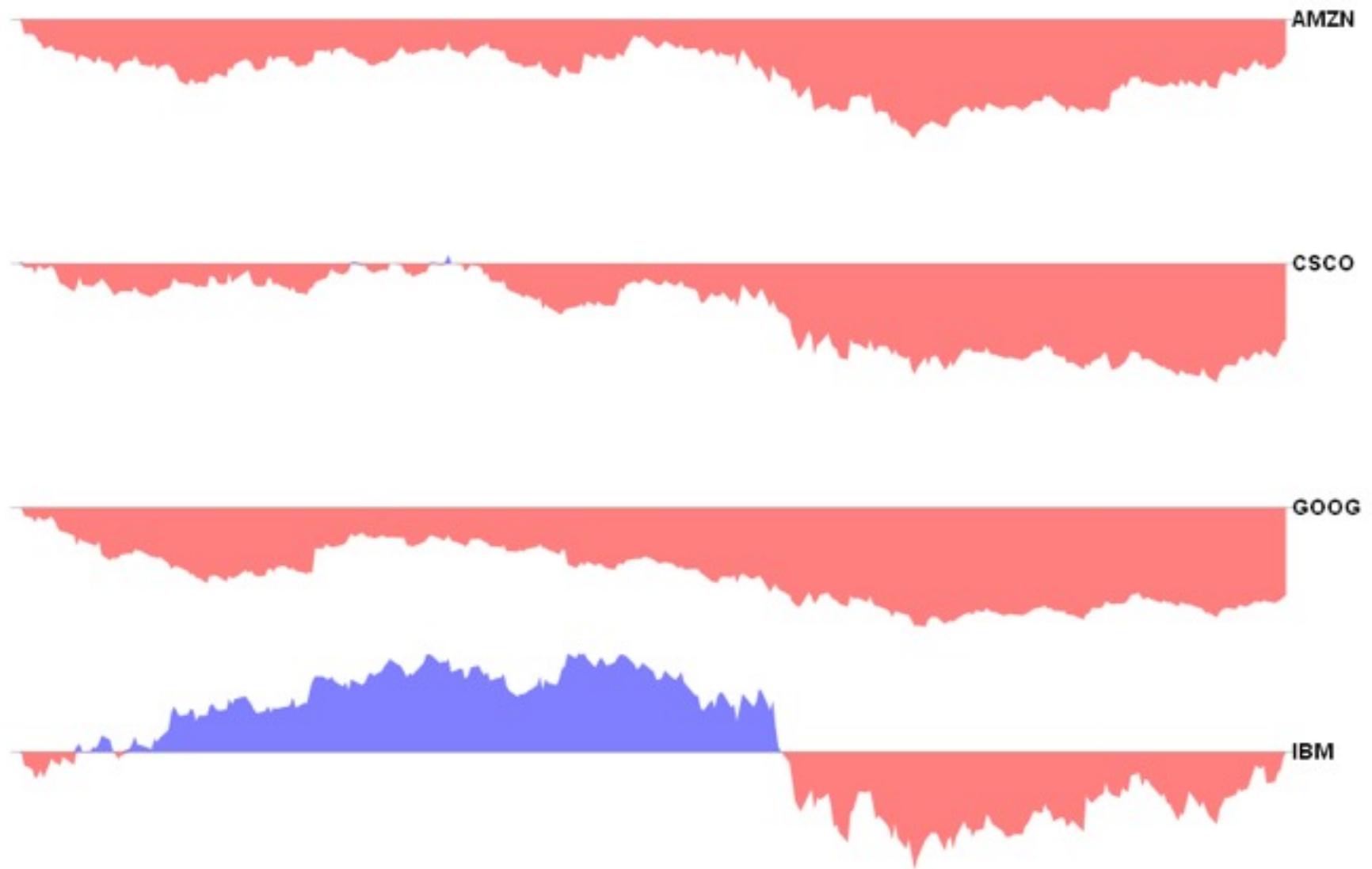
Horizon Graphs



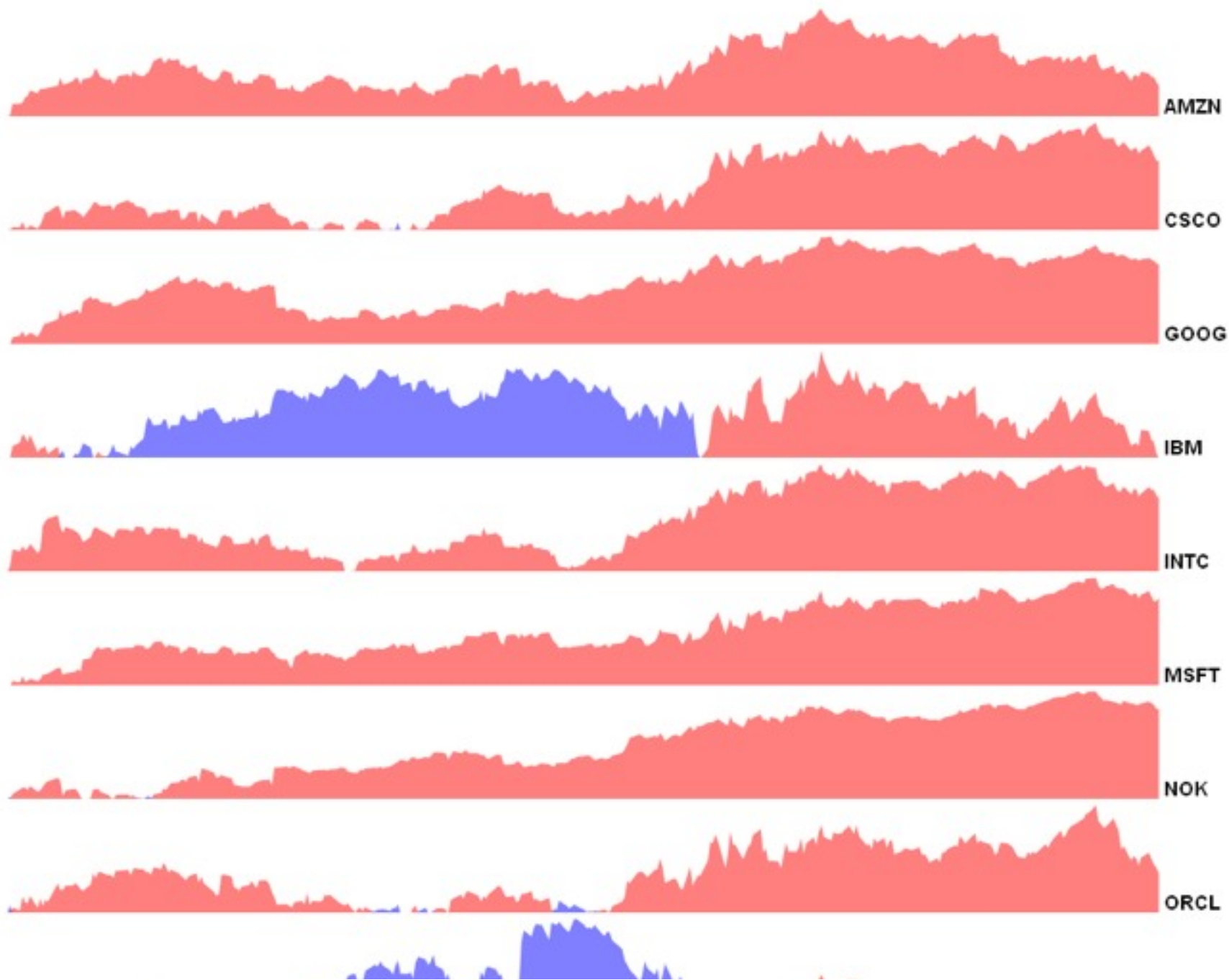
Mirror Horizon Graph

Offset Horizon Graph

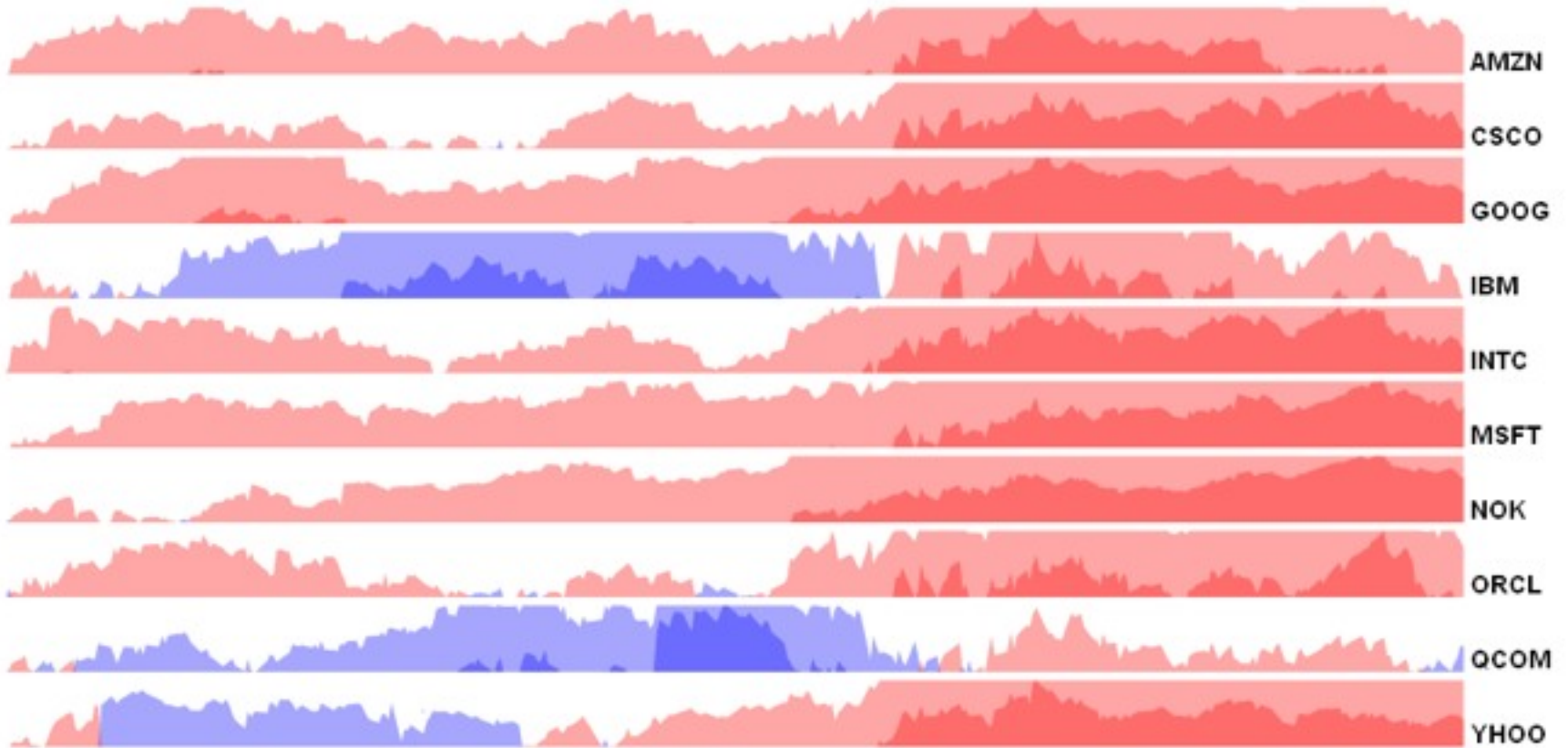
Relative Technology Stock Performance: Jan 2008 - Present



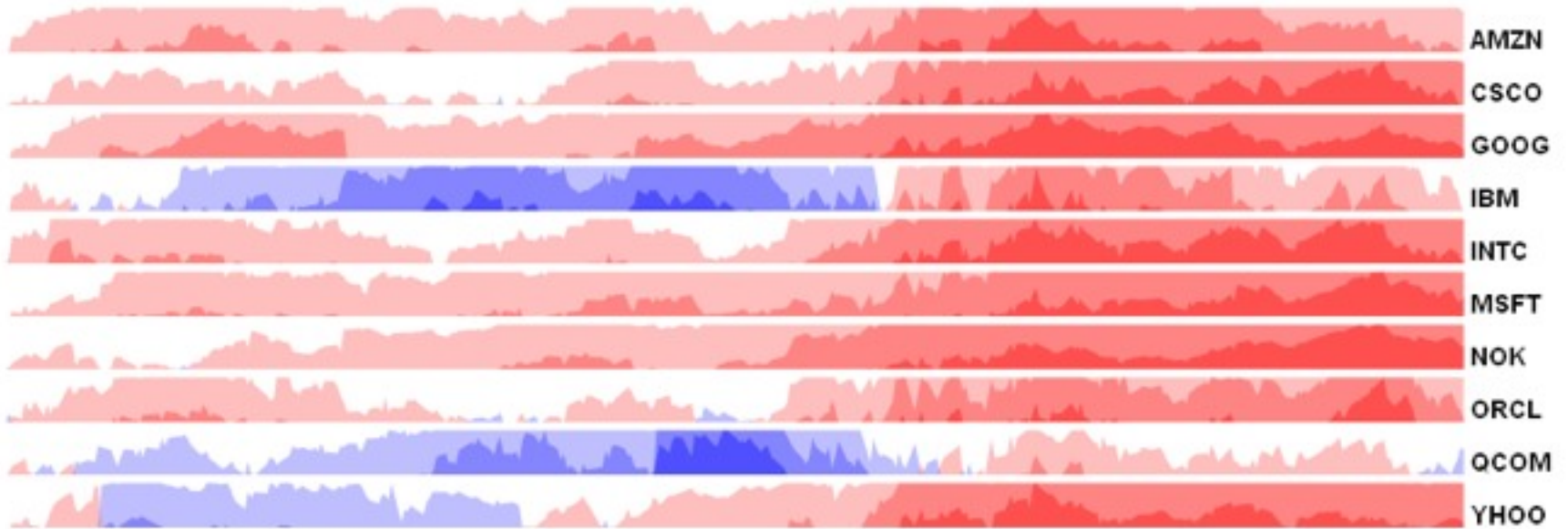
Relative Technology Stock Performance: Jan 2008 - Present



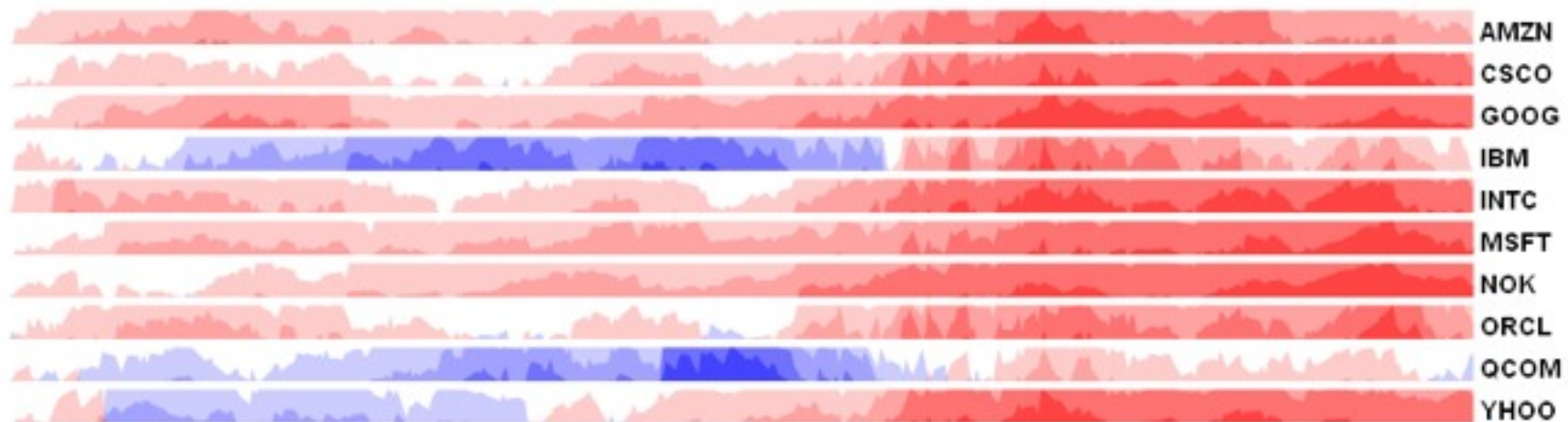
Relative Technology Stock Performance: Jan 2008 - Present



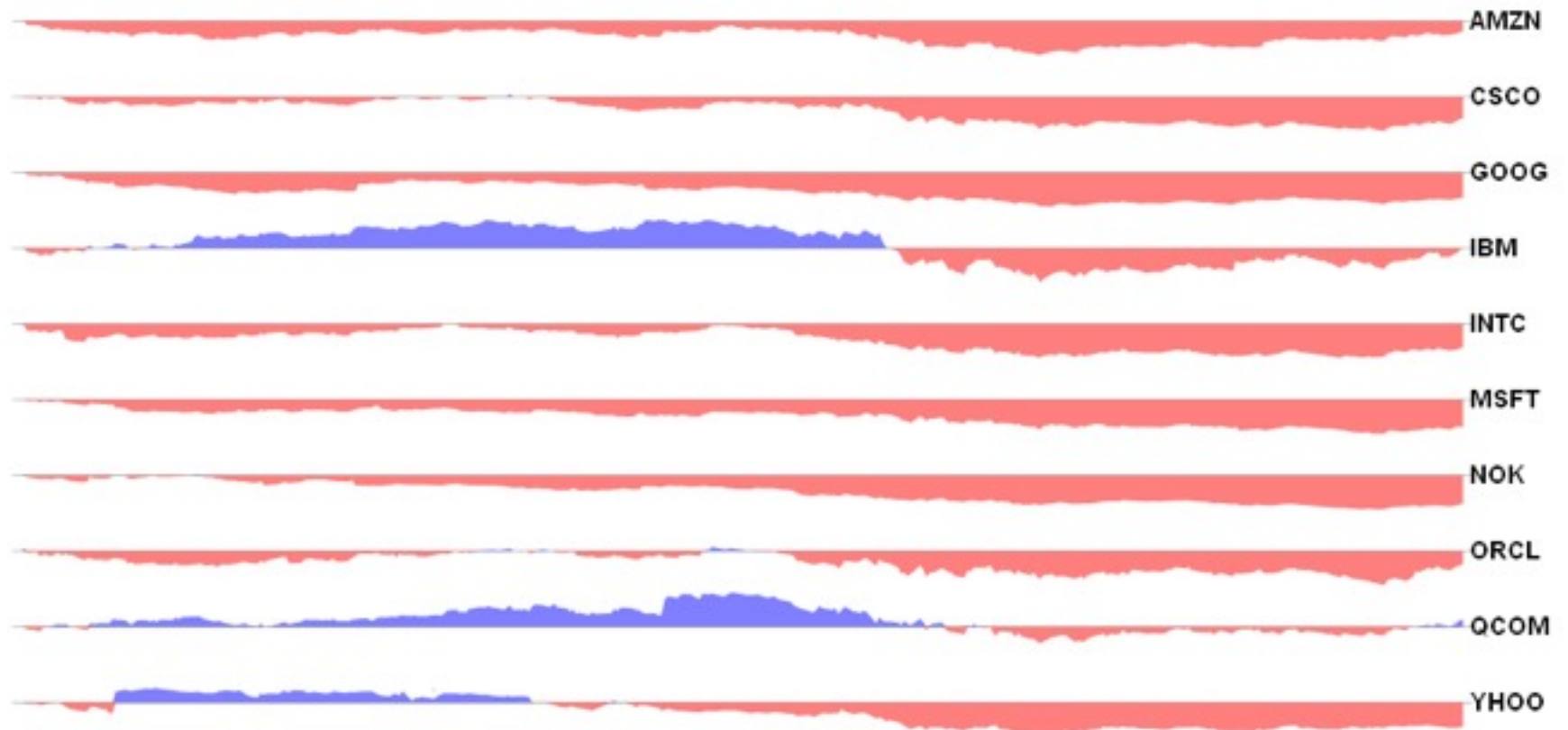
Relative Technology Stock Performance: Jan 2008 - Present



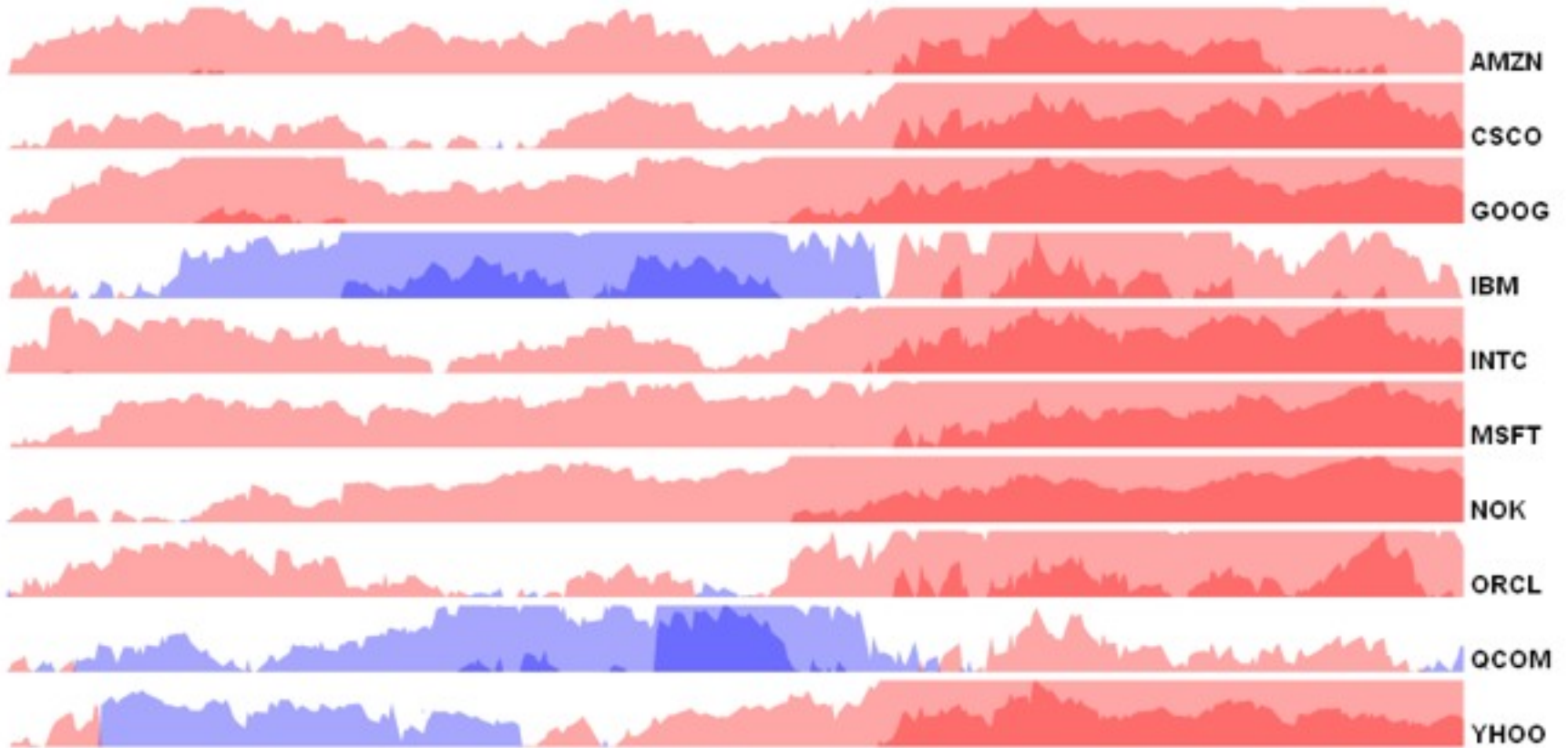
Relative Technology Stock Performance: Jan 2008 - Present



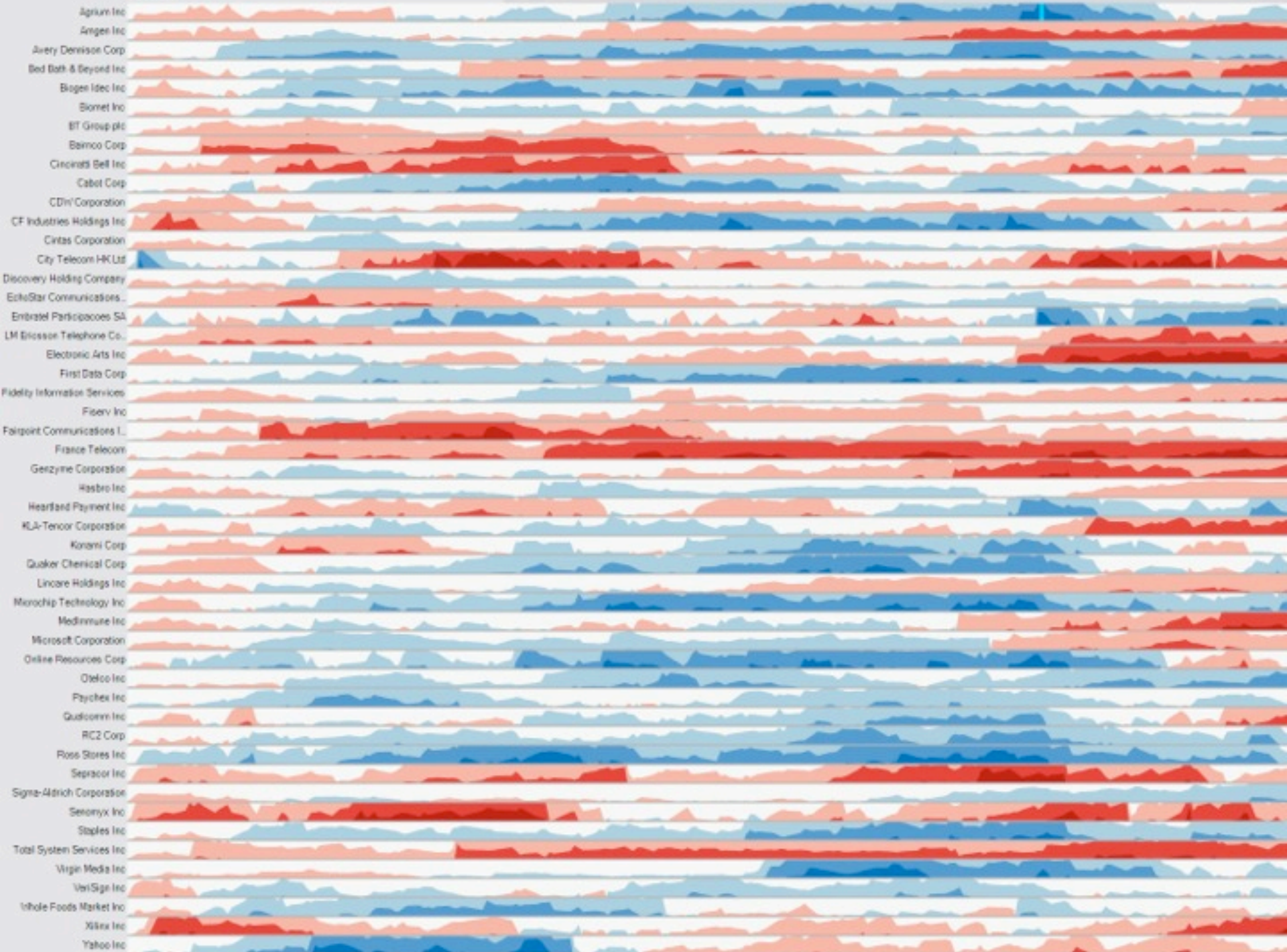
Relative Technology Stock Performance: Jan 2008 - Present



Relative Technology Stock Performance: Jan 2008 - Present



10/03/2005



Experiment 1

Horizon Graph Variants

Exp 1: Mirror/Offset & Banding

Q1: How does the choice of mirrored or offset horizon graph affect estimation time and accuracy?



Mirror Horizon Graph

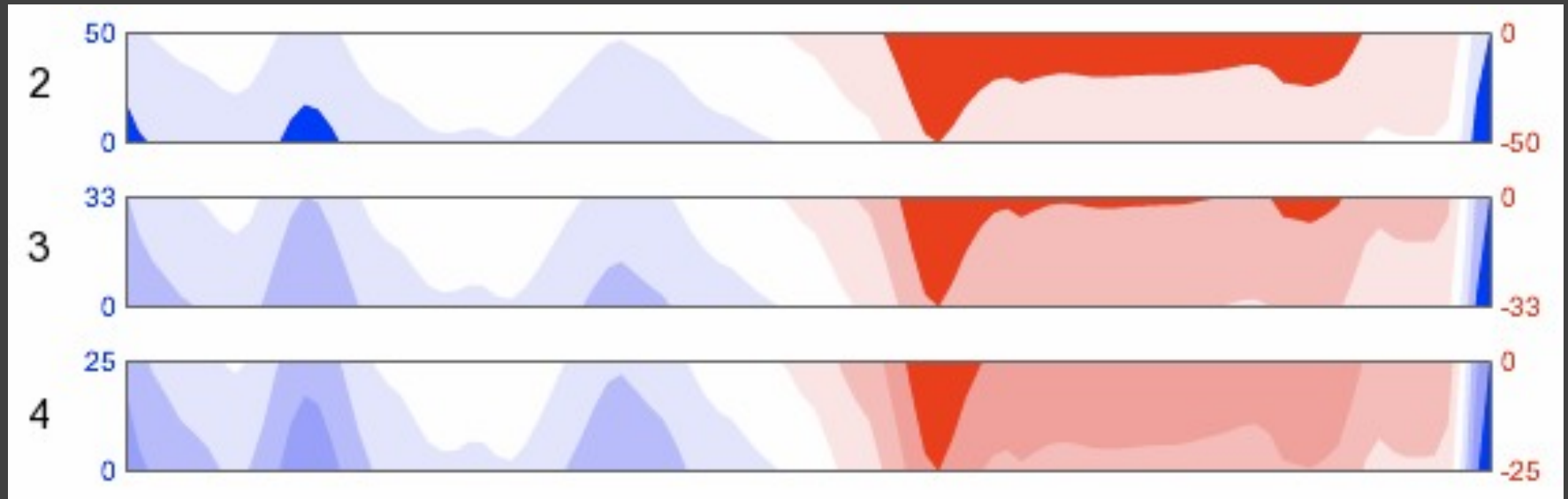


Offset Horizon Graph

Exp 1: Mirror/Offset & Banding

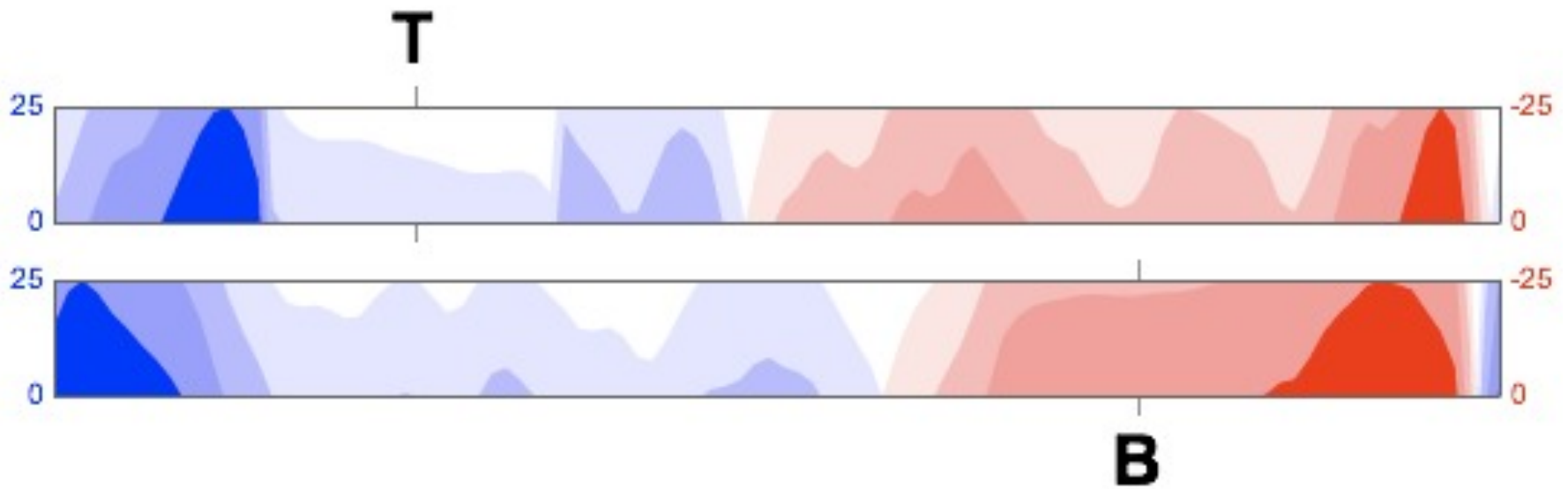
- Q1:** How does the choice of mirrored or offset horizon graph affect estimation time and accuracy?
- Q2:** How does the number of bands in a horizon chart affect estimation time and accuracy?

Experiment 1 Design

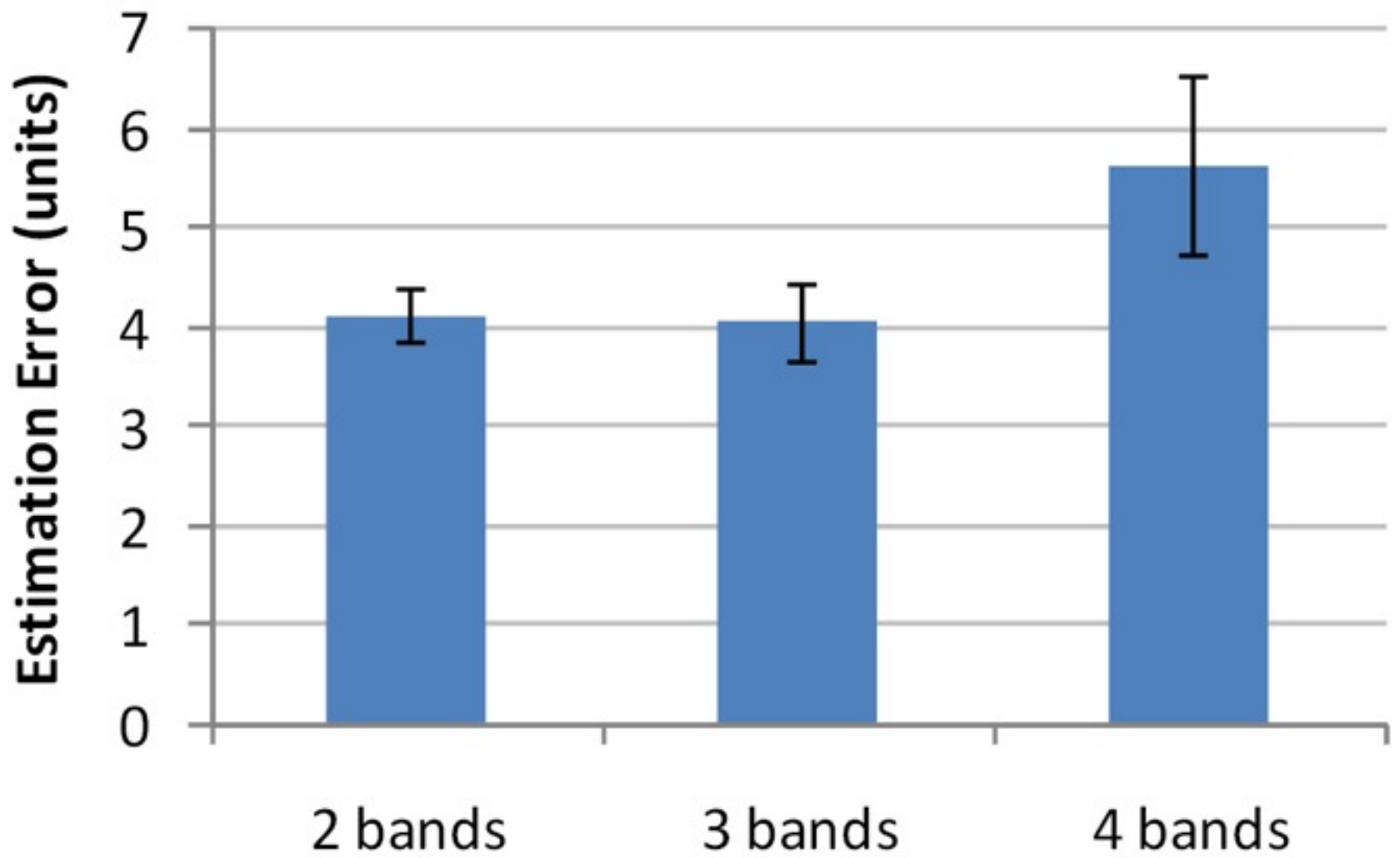


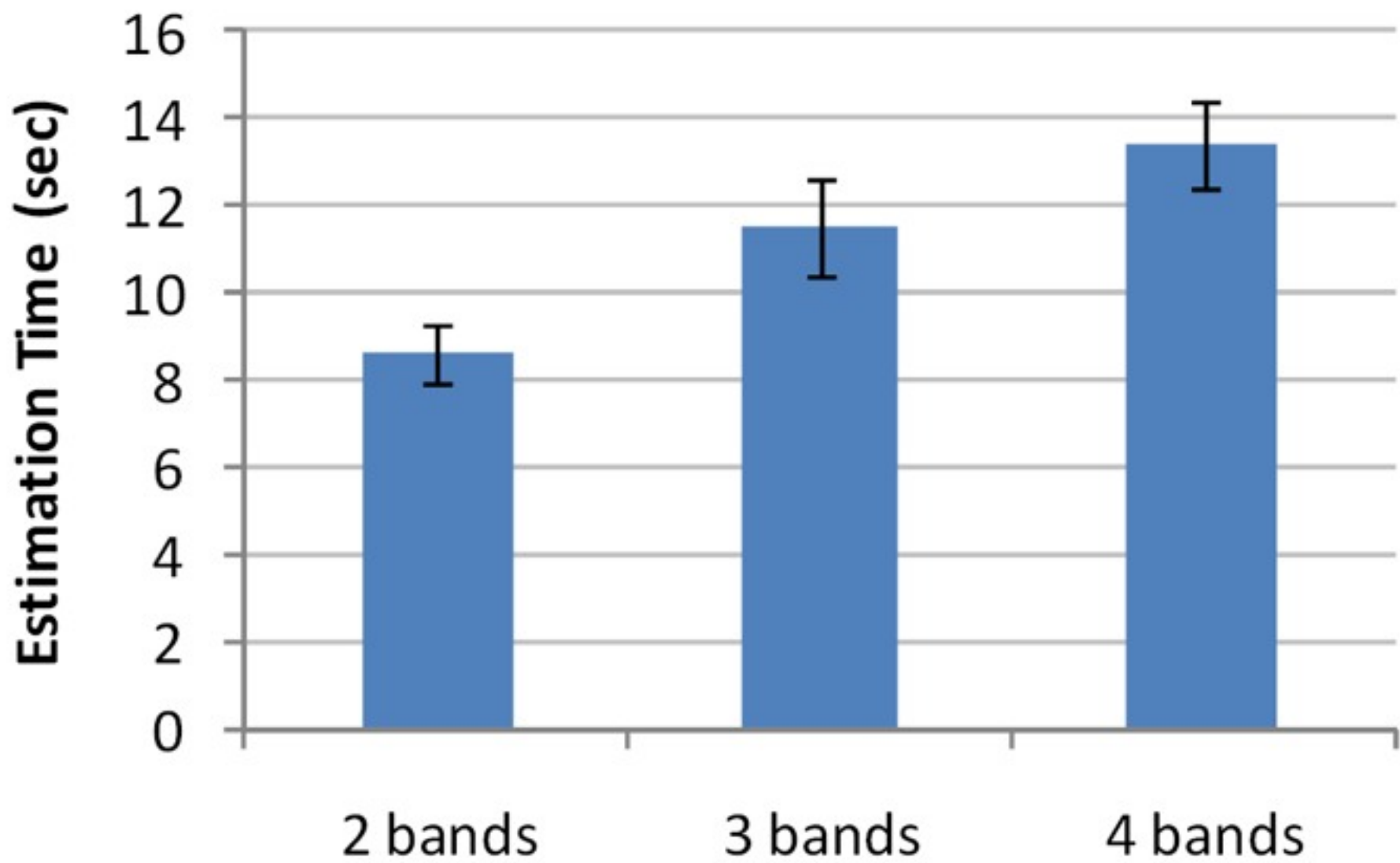
2 (type) x 3 (band count) within-subjects design

- N = 18 (13 male, 5 female), UCB students
- Deployed on the web as a Flash applet



Estimate the difference between T and B (0-200) to within 5 values.





Experiment 1 Results

Q1: No significant difference between mirrored and offset horizon graphs.

Q2: Both estimation time and error increased with more bands.

→ Higher band counts led to **difficulty identifying bands** and **fatigue from mental mathematics**.

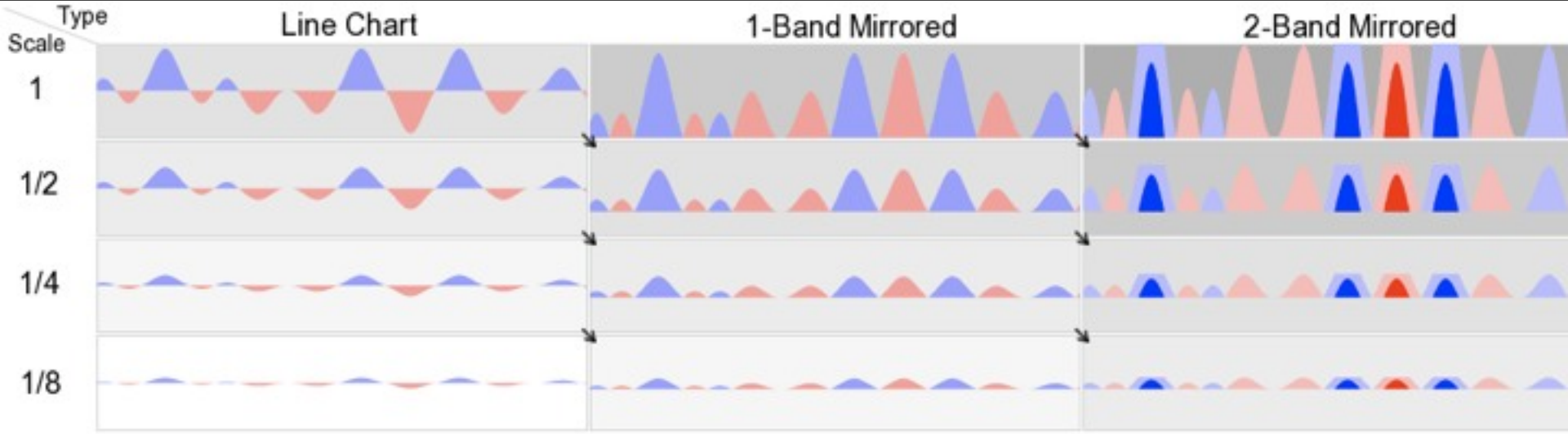
Experiment 2

Chart Type and Size

Exp 2: Chart Type & Size

- Q1:** How do mirroring and layering affect estimation time and accuracy compared to line charts?
- Q2:** How does chart size affect estimation time and accuracy?

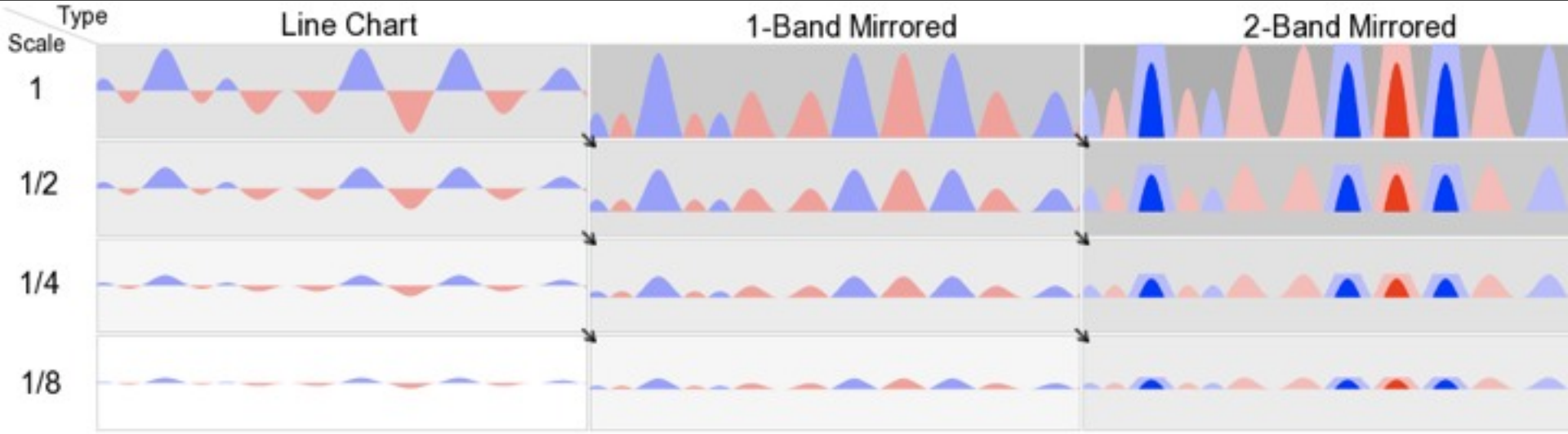
Experiment 2 Design



3 (chart type) x 4 (size) within-subjects design

- N = 30 (17 male, 13 female), UCB undergrads
- 14.1 inch LCD display at 1024 x 768 pixel resolution
- At scale = 1, chart was 13.9 x 1.35 cm (48 pixels)

Experiment 2 Design

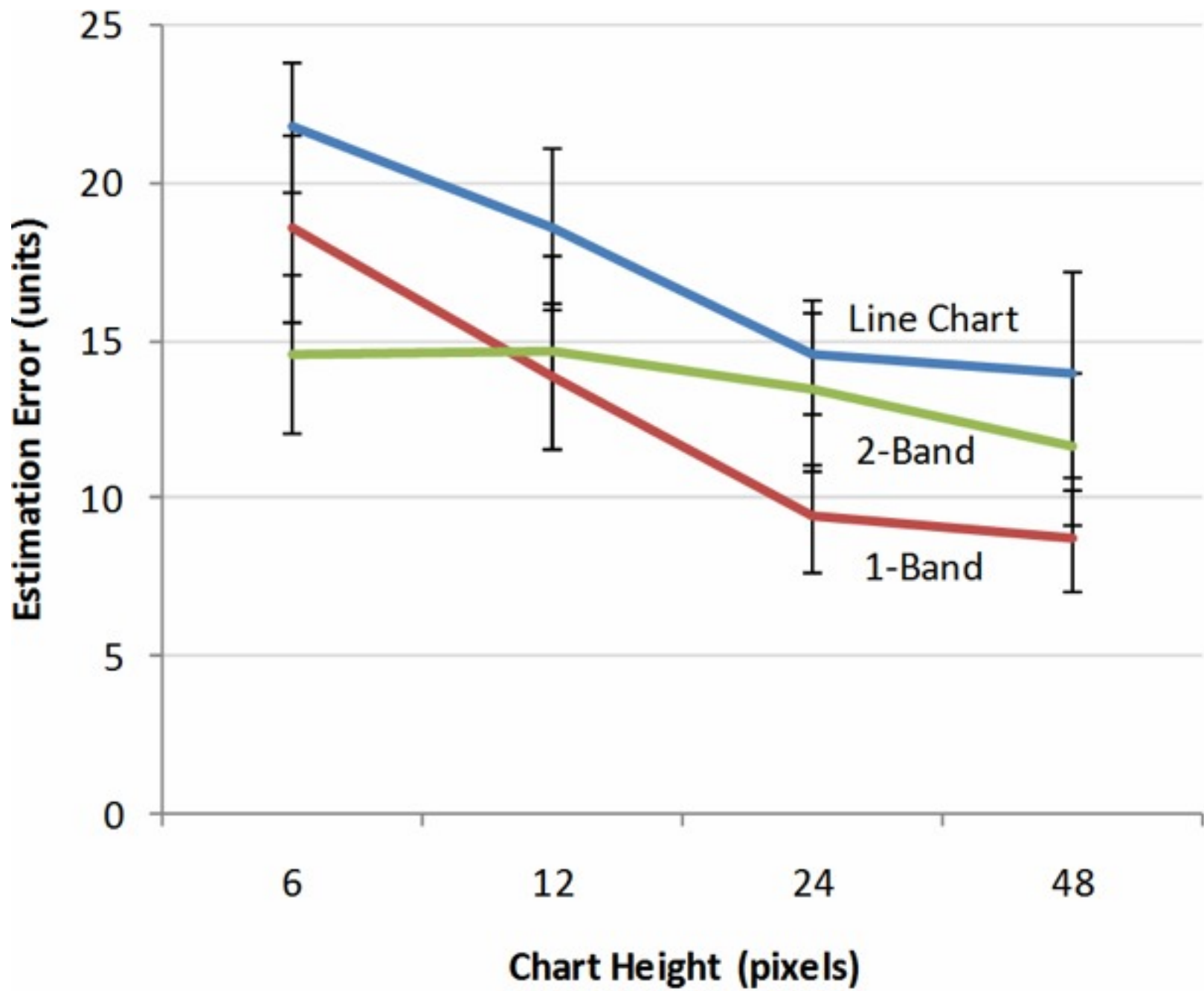


3 (chart type) x 4 (size) within-subjects design

- N = 30 (17 male, 13 female), UCB undergrads

2 (chart type) x 3 (size: 1/8, 1/12, 1/24) follow-up study

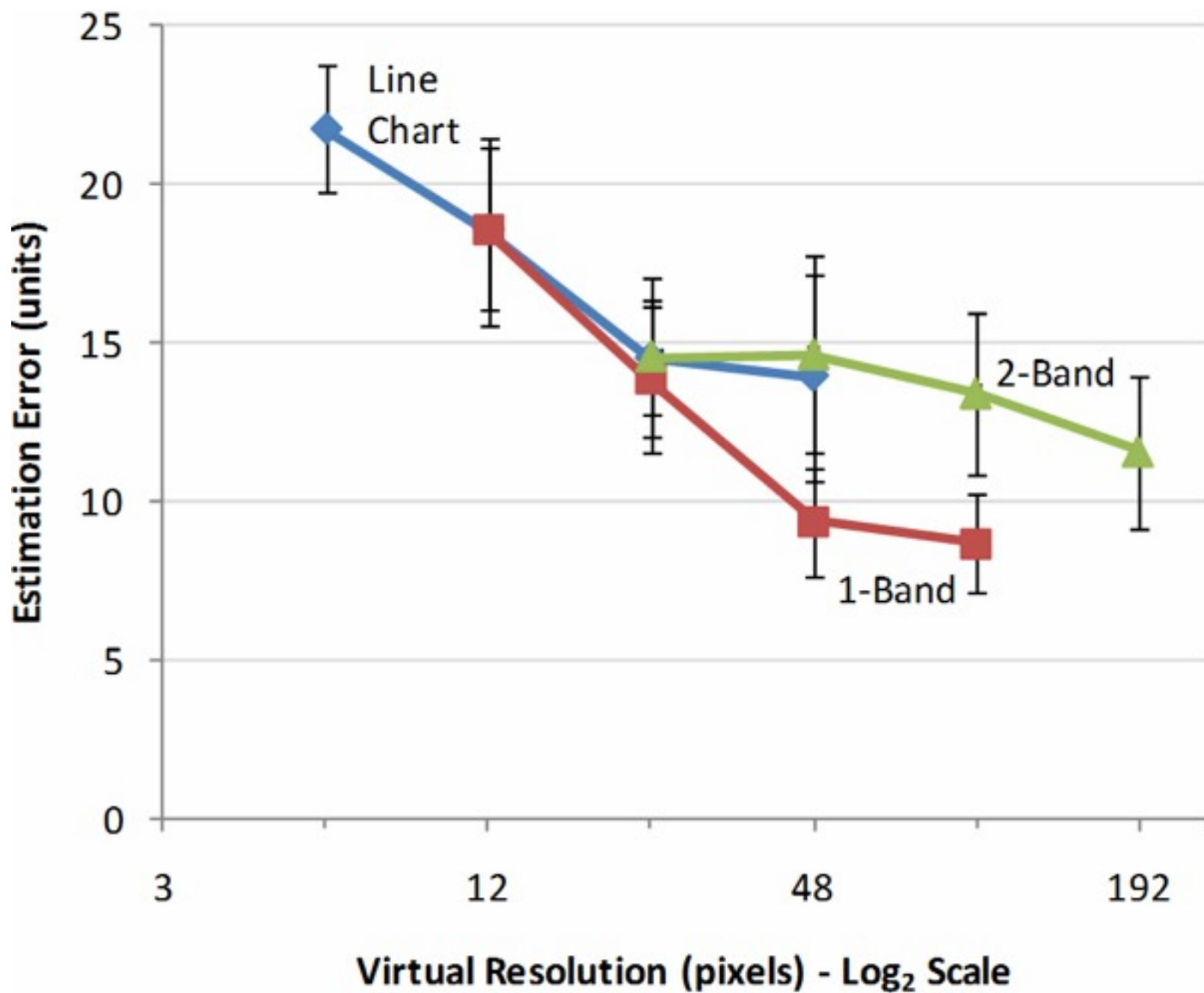
- N = 8 (6 male, 2 female), UCB engineering grads

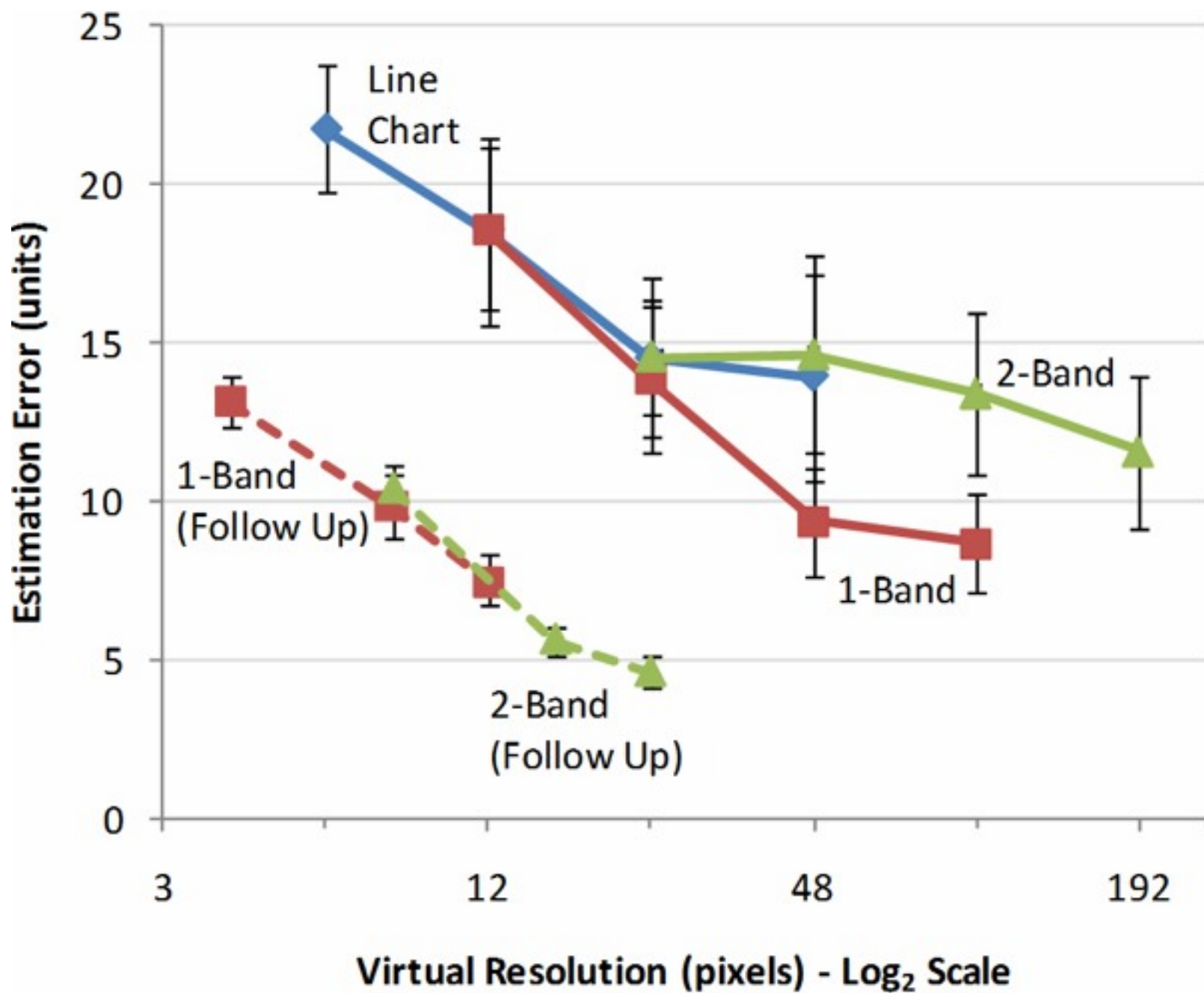


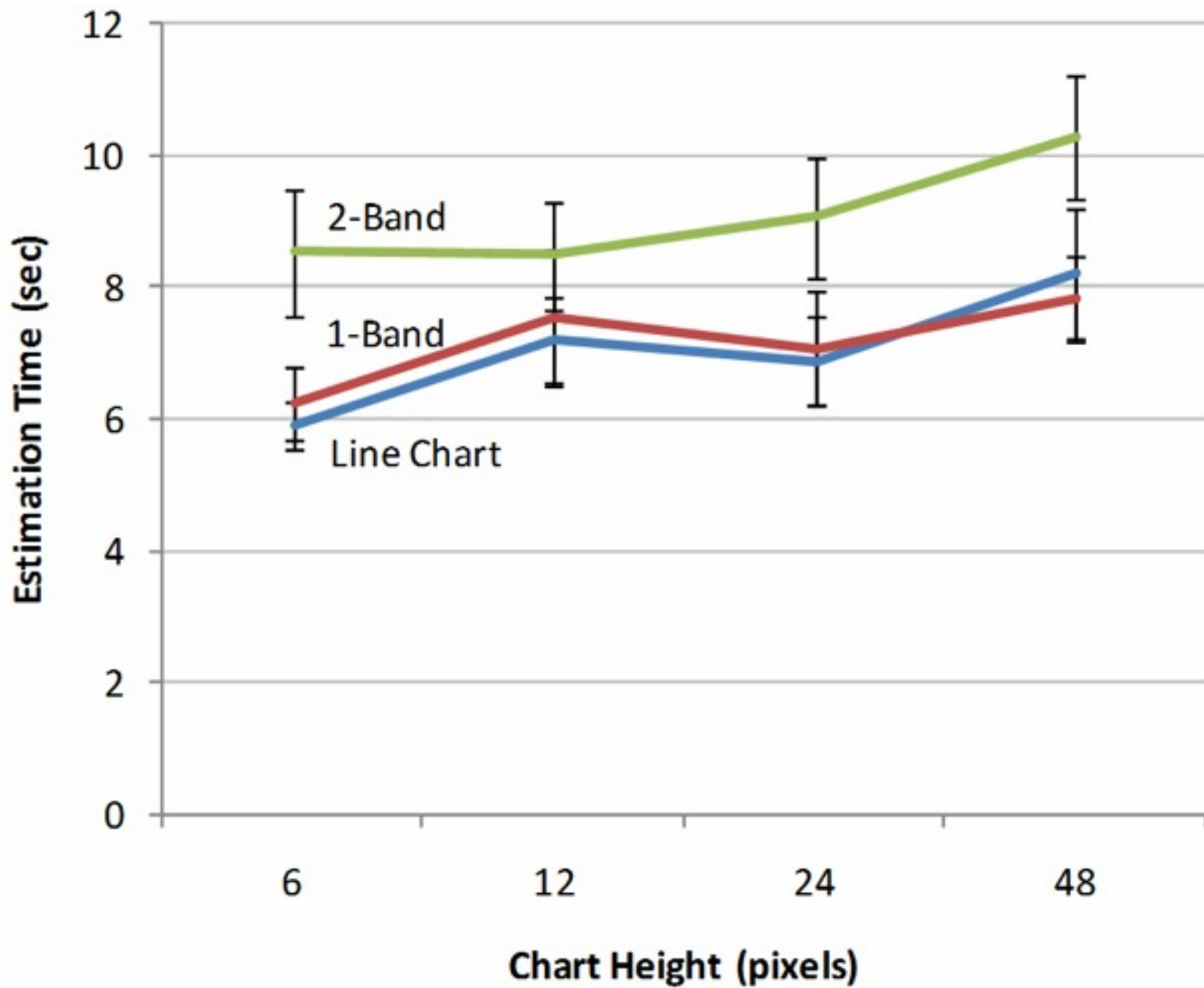
Virtual Resolution (VR)

The un-mirrored, un-layered height of a chart









Experiment 2 Results

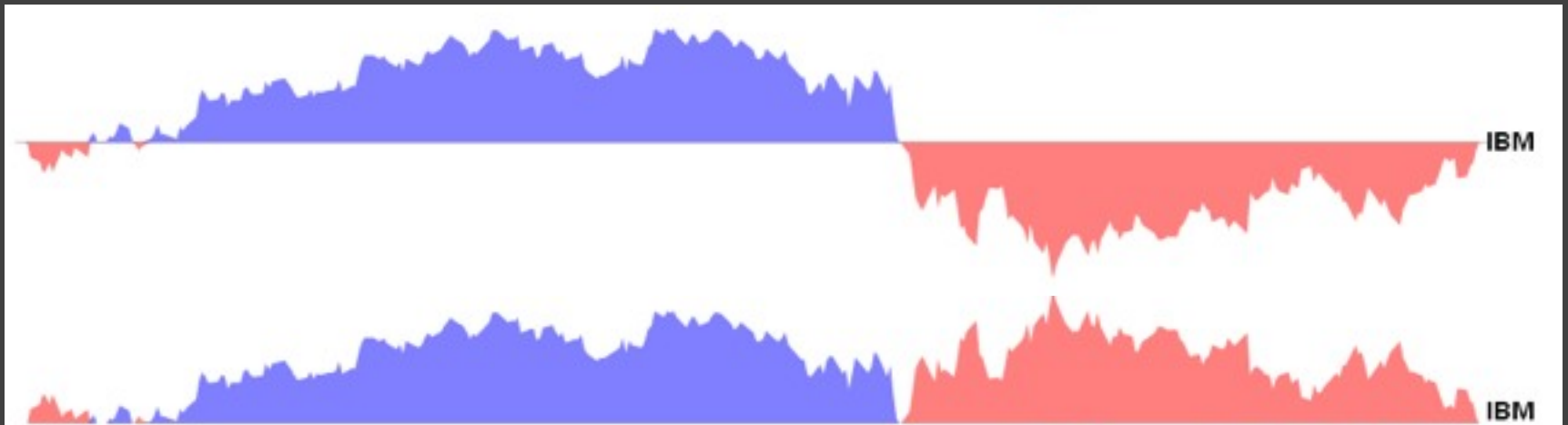
Q1: 2-band horizon graph (but not mirrored graph) had higher baseline estimation time and error.

Q2: Estimation error increased as the *virtual resolution* decreased.

Estimation time decreased as the *physical height* decreased.

Design Implications

Mirroring does not hamper perception



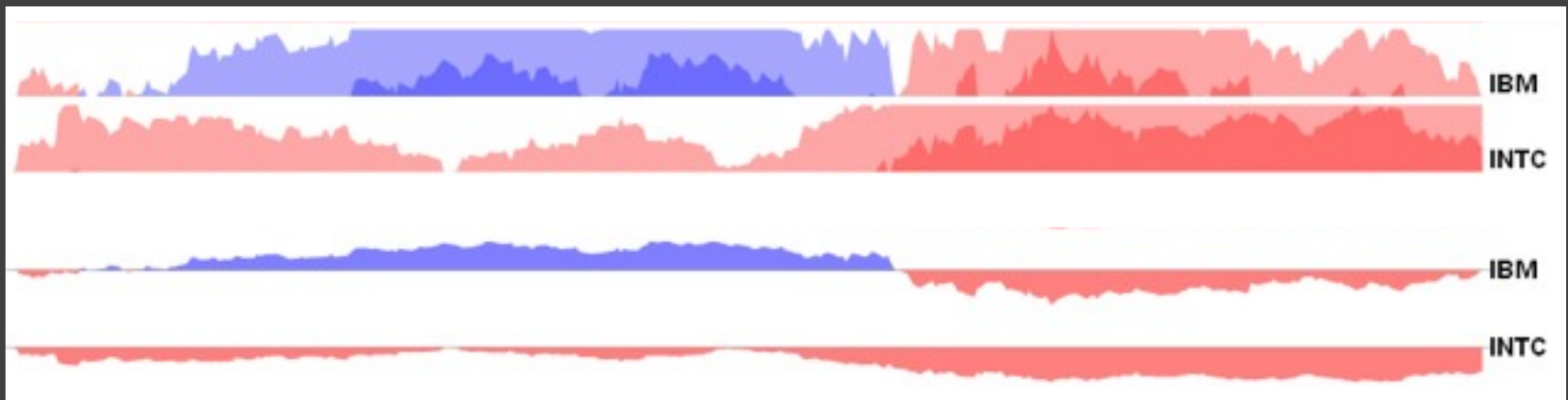
Design Implications

Mirroring does not hamper perception

Layered bands beneficial for smaller charts

2-band mirror charts more accurate for heights under 6.8mm (24 pixels @ 1024x768)

Predict benefits for 3 bands under 1.7mm (6 px)



Design Implications

Mirroring does not hamper perception

Layered bands beneficial for smaller charts

Optimal chart sizing

Sweet spots in time/error curves

6.8mm (24 px) for line chart & mirrored chart

3.4mm (12 px) for 2-band horizon graph

Administrivia

Final Project

Poster Presentations

Session is Thu Mar 13 5-8pm in CSE Atrium
Bring **Poster + Laptop/Device** for demos
Arrive *early* to setup!

Post Webpage on GitHub Pages

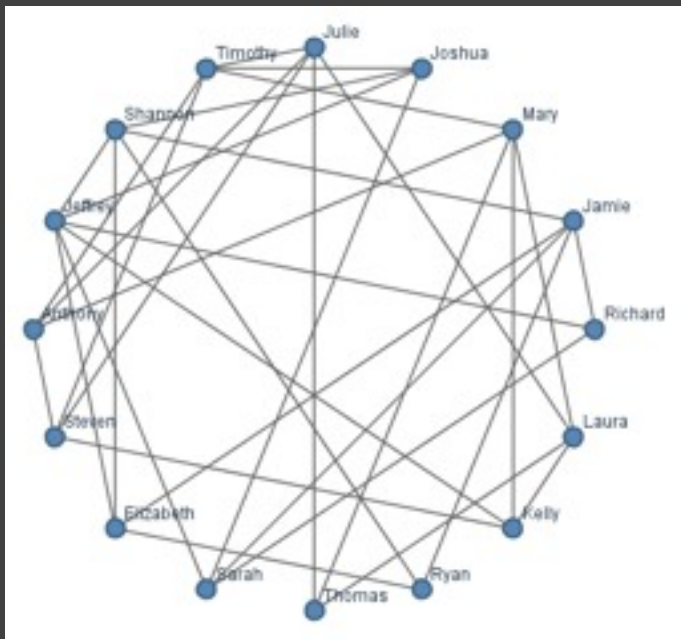
List team members, title, abstract, link to paper
Include summary image for project!

Final Project Reports

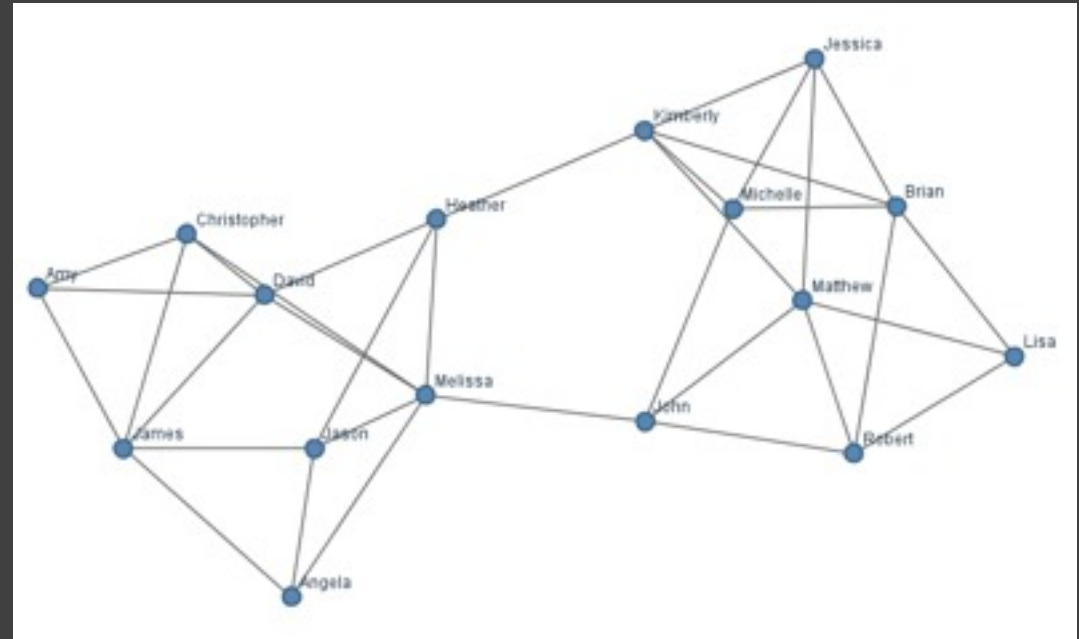
Due Thu Mar 20, by 7am, posted to GitHub
4-6 pages in ACM or IEEE TVCG format

Perceptual Organization of Graphs

Perceptual organization of graphs



Circular



Force-Directed

Experiment Design

Factors

Circular or Force-Directed Seed Layout

of Between-Cluster Edges (“masking”)

All graphs had two primary clusters

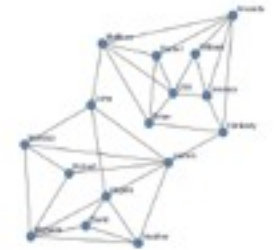
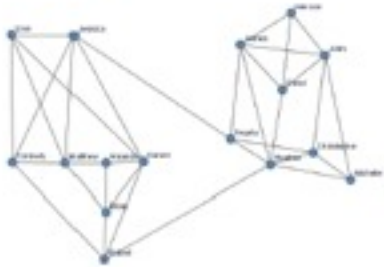
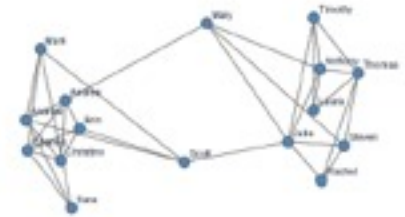
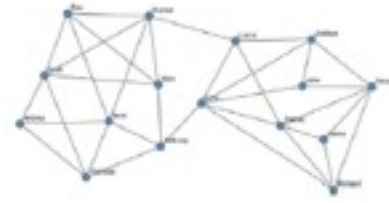
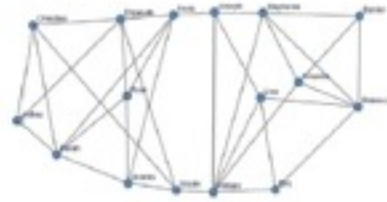
Measures

of Edge Crossings

Average Edge Length

Average Node Distance

within or between clusters



Circular



Force-Directed

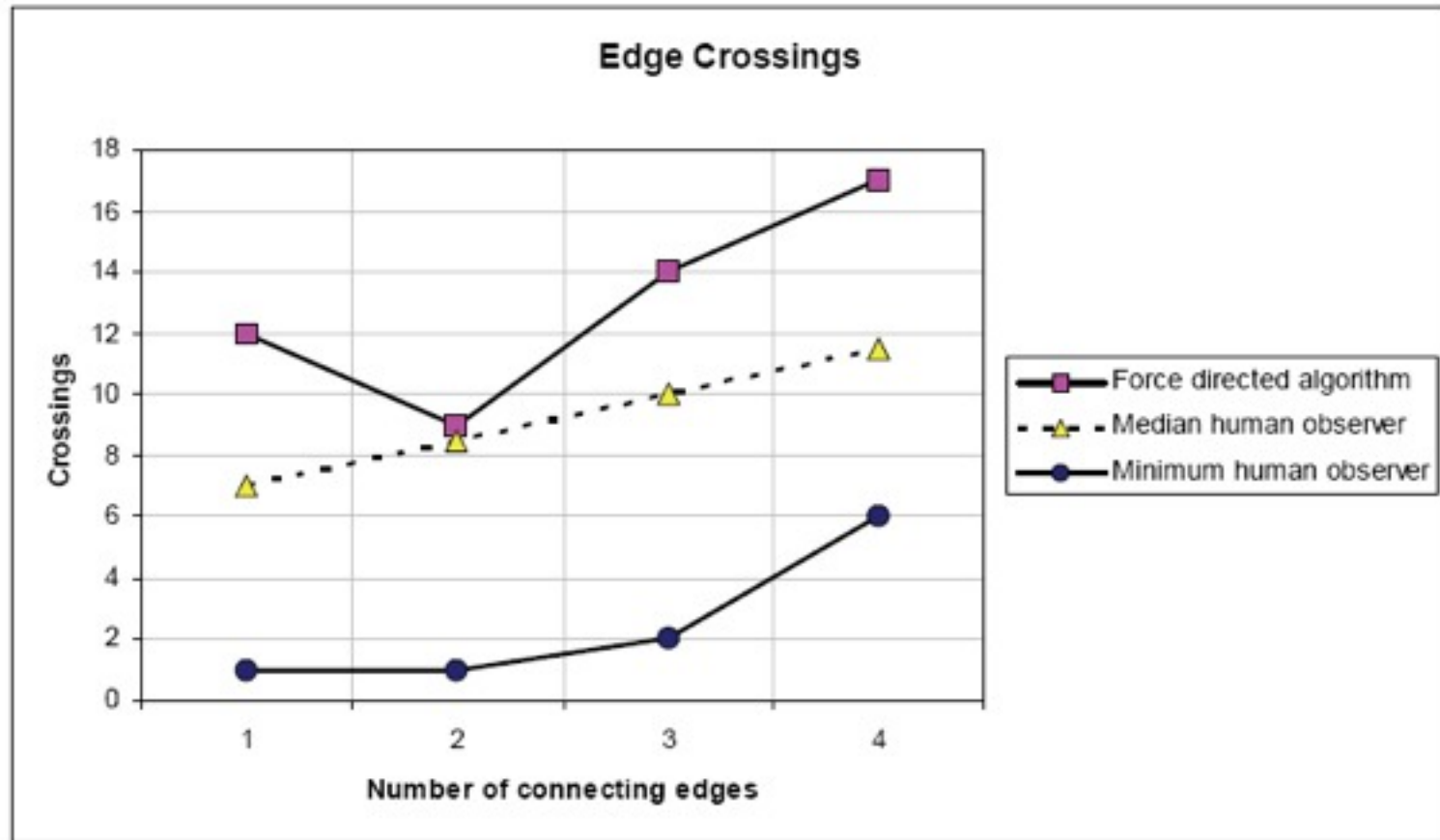


Figure 4. Edge Crossings. Human observers produced graph layouts with fewer edge crossings than the force-directed graph algorithm.

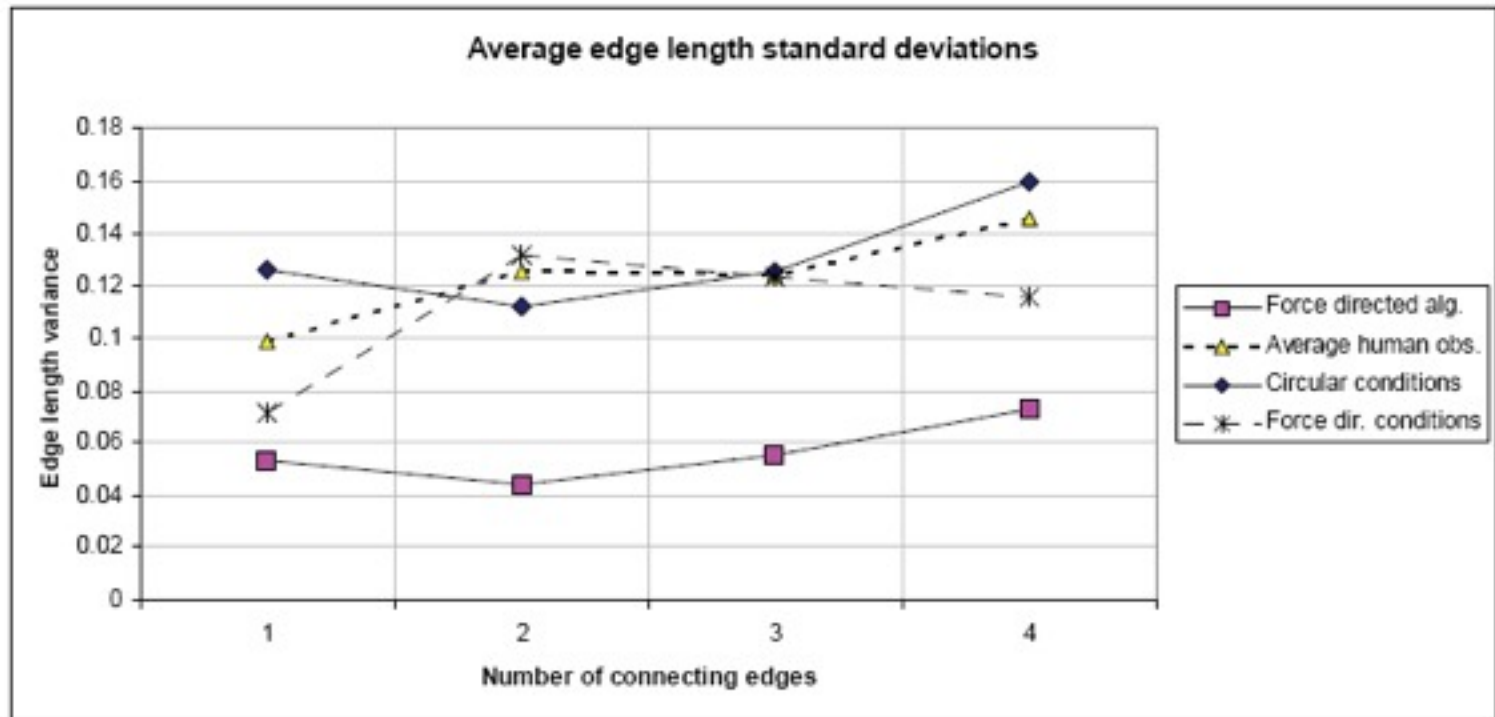


Figure 5. Edge Length Distribution. Human observers did not focus on maintaining equal edge length as much as the force directed algorithm.

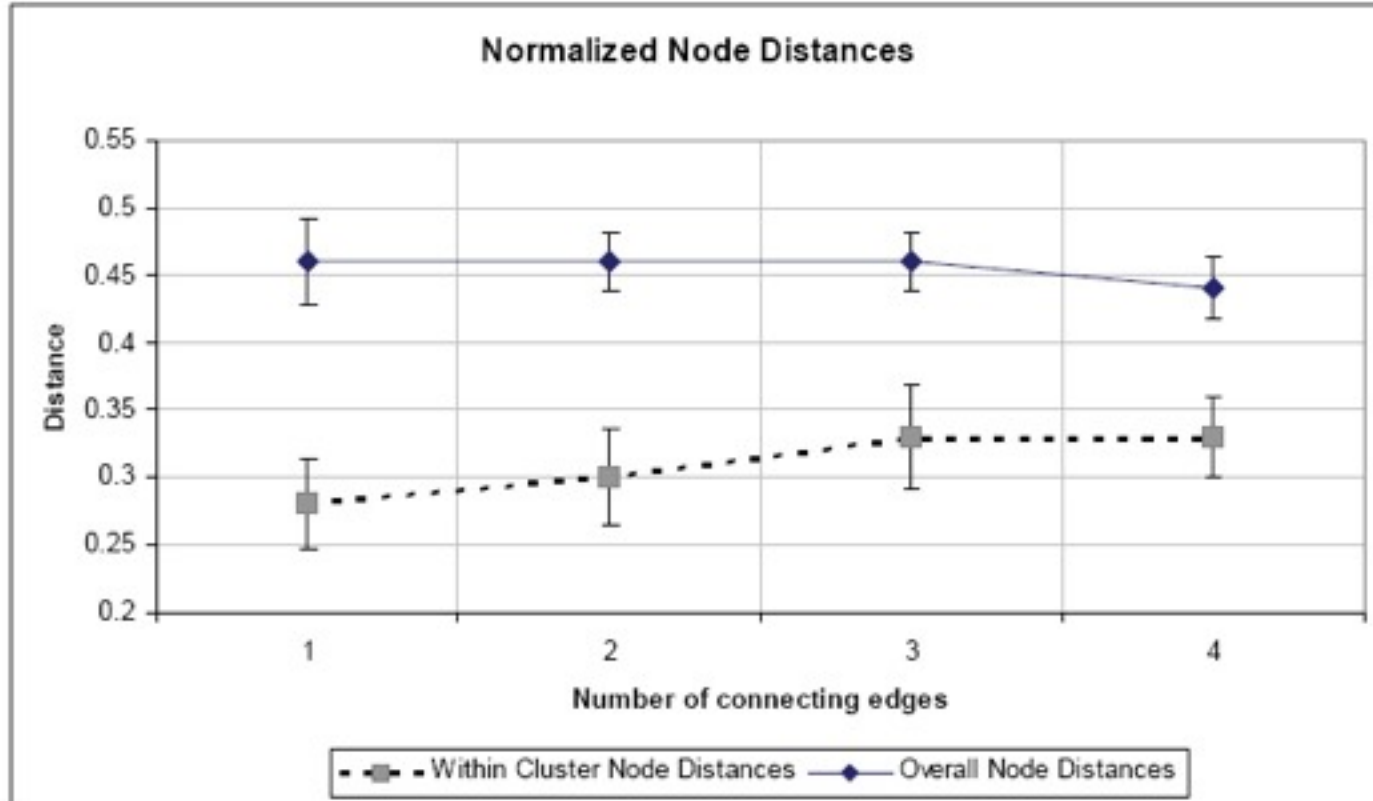


Figure 7. Cluster Extraction. For all levels of masking, the distance between nodes within a cluster is significantly smaller than the overall inter-node distance, demonstrating perceptual grouping. Error bars show 95% confidence intervals

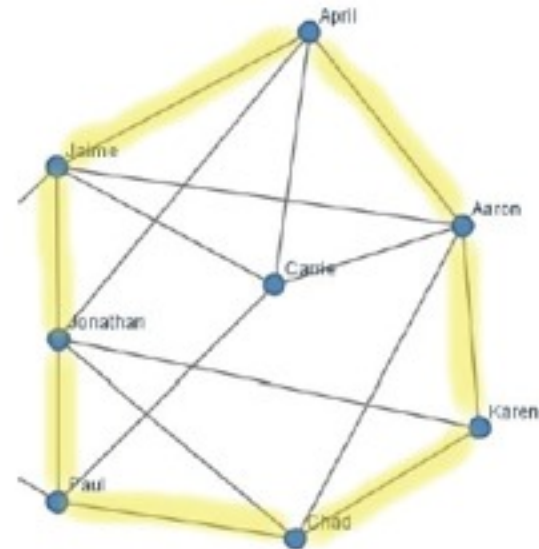
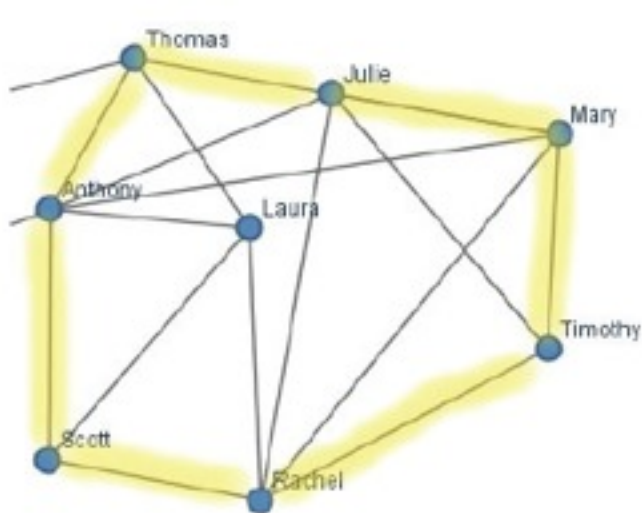
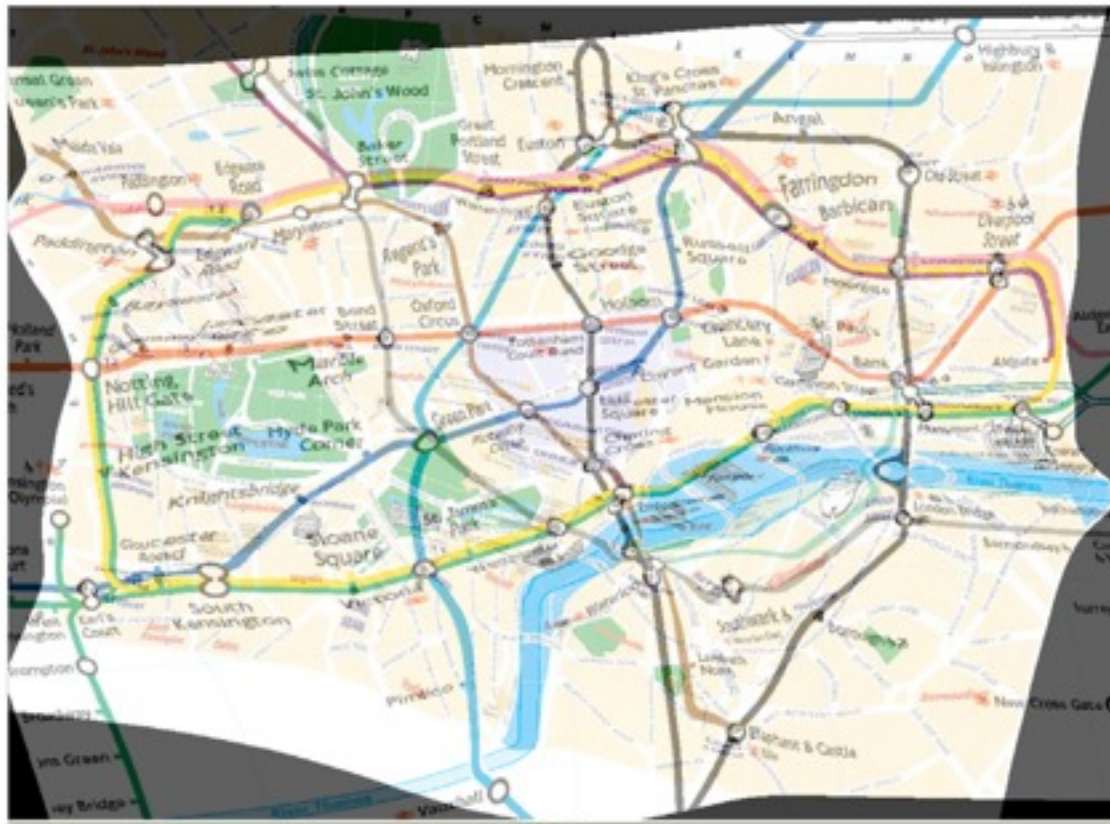


Figure 9. Cluster Hulls. Two examples of user-generated layouts where cluster edges formed a hull enclosing the cluster, organizing it into a single perceptual group.

The role of environment [Reilly 07]

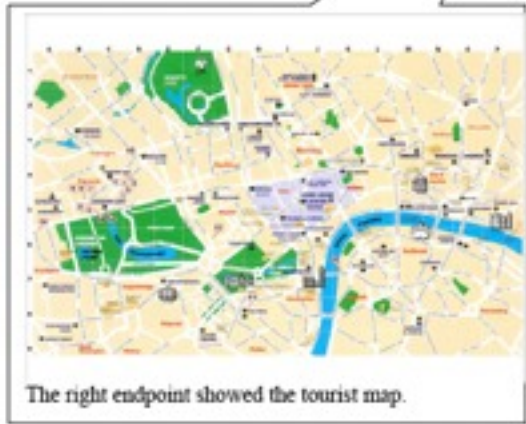




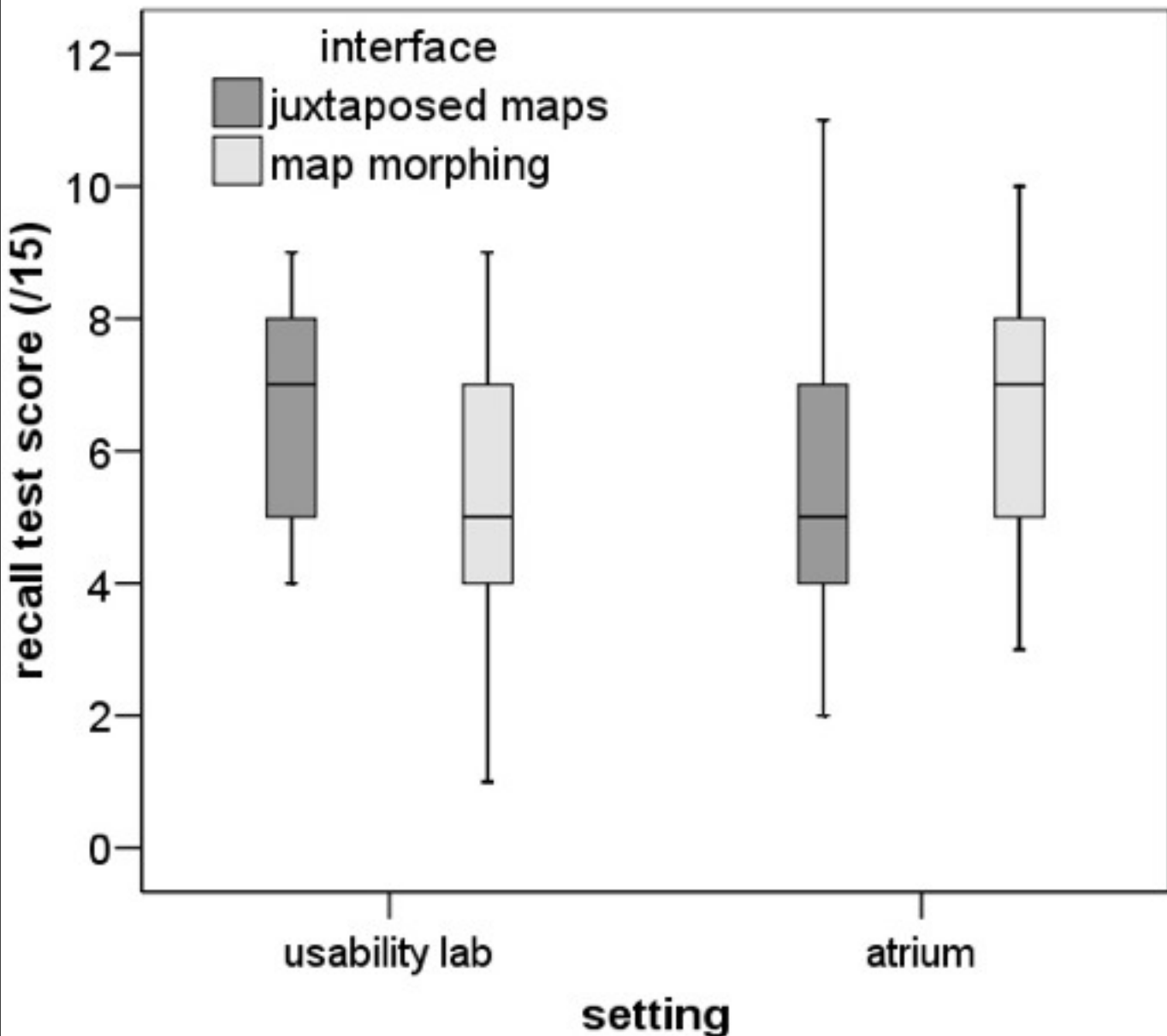
The slider was used to control the bidirectional morph. Shape blending (distortion) and alpha blending occurred simultaneously, in proportion to the distance of the slider from each endpoint.



The Tube map was at the left endpoint of the morph.



The right endpoint showed the tourist map.



Summary

Design and analyze visualization techniques in context of real-world use

Time/error analyses can be insightful, but they don't provide a complete picture.

Performance measures may be more suited to serious analysis than casual use?

Course Summary

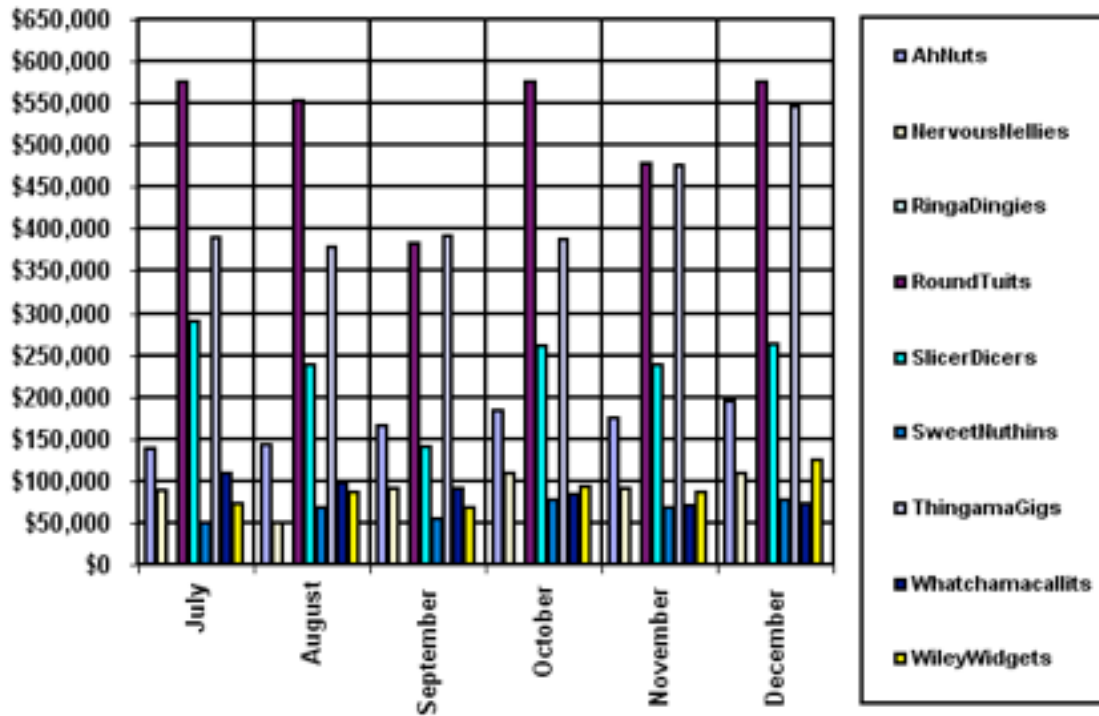
Data and Image Models

LES VARIABLES DE L'IMAGE									
			POINTS	LIGNES	ZONES	12 14			
XY	2 DIMENSIONS DU PLAN								
Z	TAILLE								
	VALEUR								
LES VARIABLES DE SÉPARATION DES IMAGES									
	GRAIN								
	COULEUR								
	ORIENTATION								

Sémiologie Graphique [Bertin 67]

Visualization (Re-)Design

SlicerDicers' Sales Compared to Other Products



Problematic design

Sales of SlicerDicers Compared to Other Products

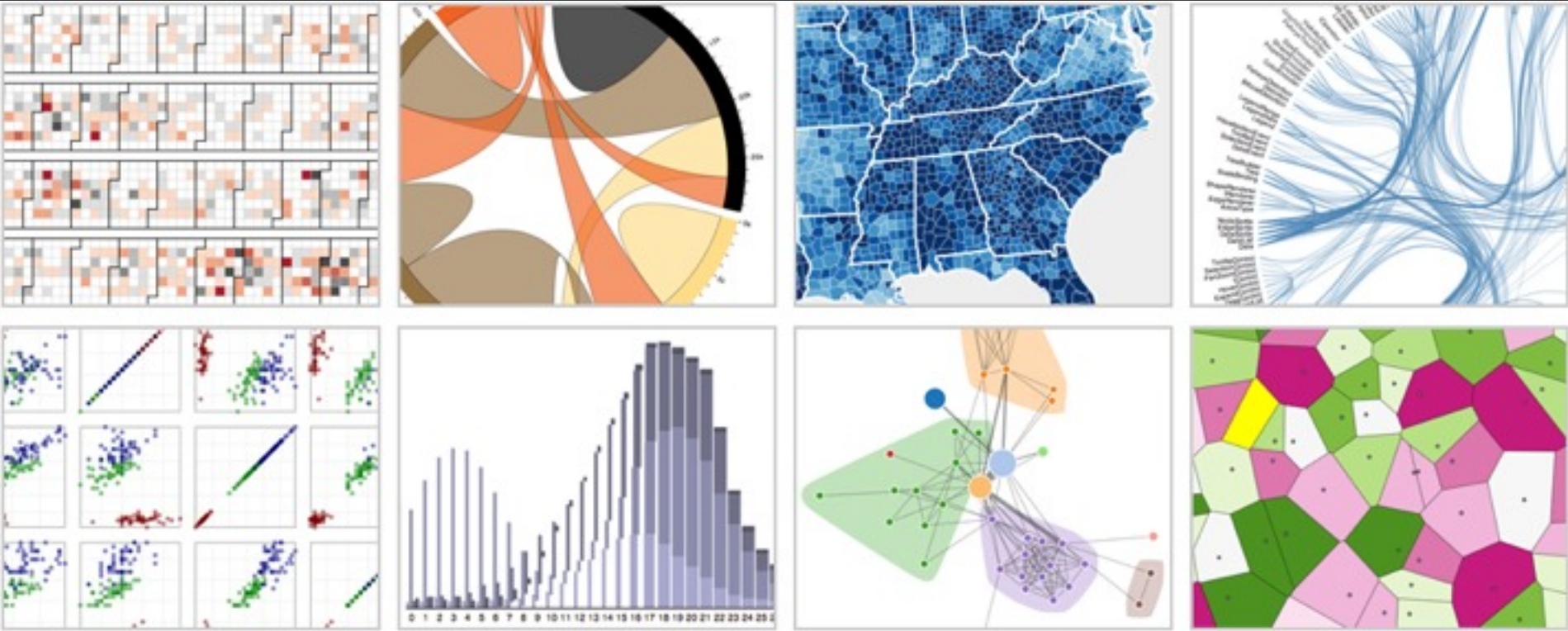
July - December, 2003

(SlicerDicers' sales are displayed as black reference lines of 100%; the red bars represent the average monthly sales percentage for July through December.)



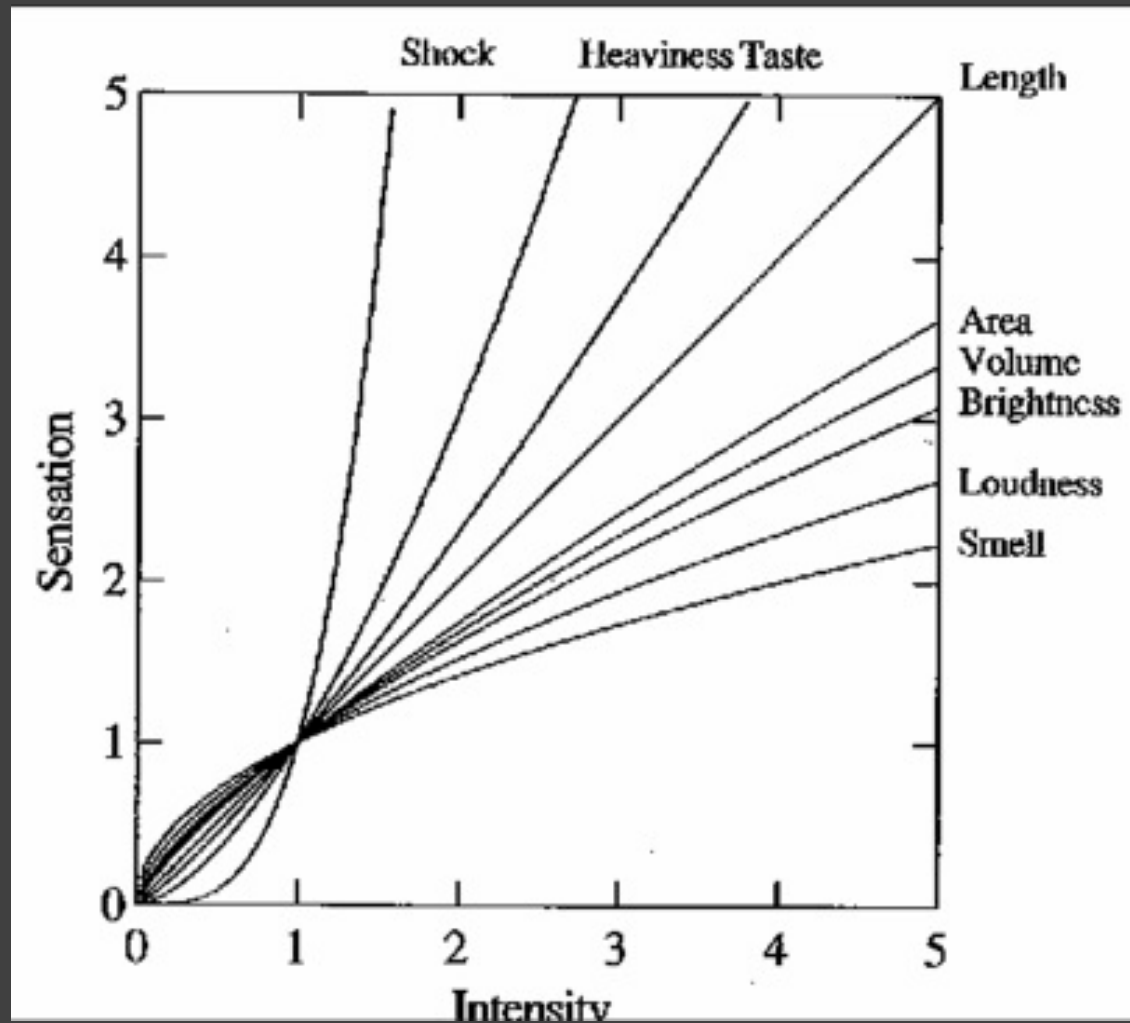
Redesign

Visualization Software



D3: Data-Driven Documents

Graphical Perception

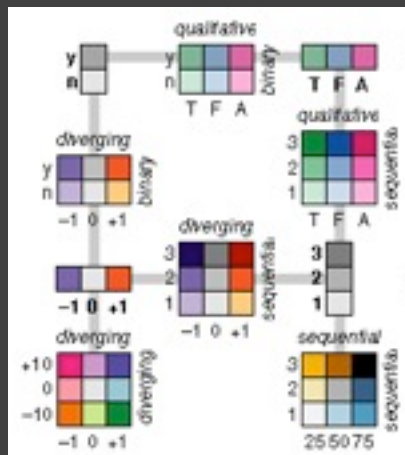
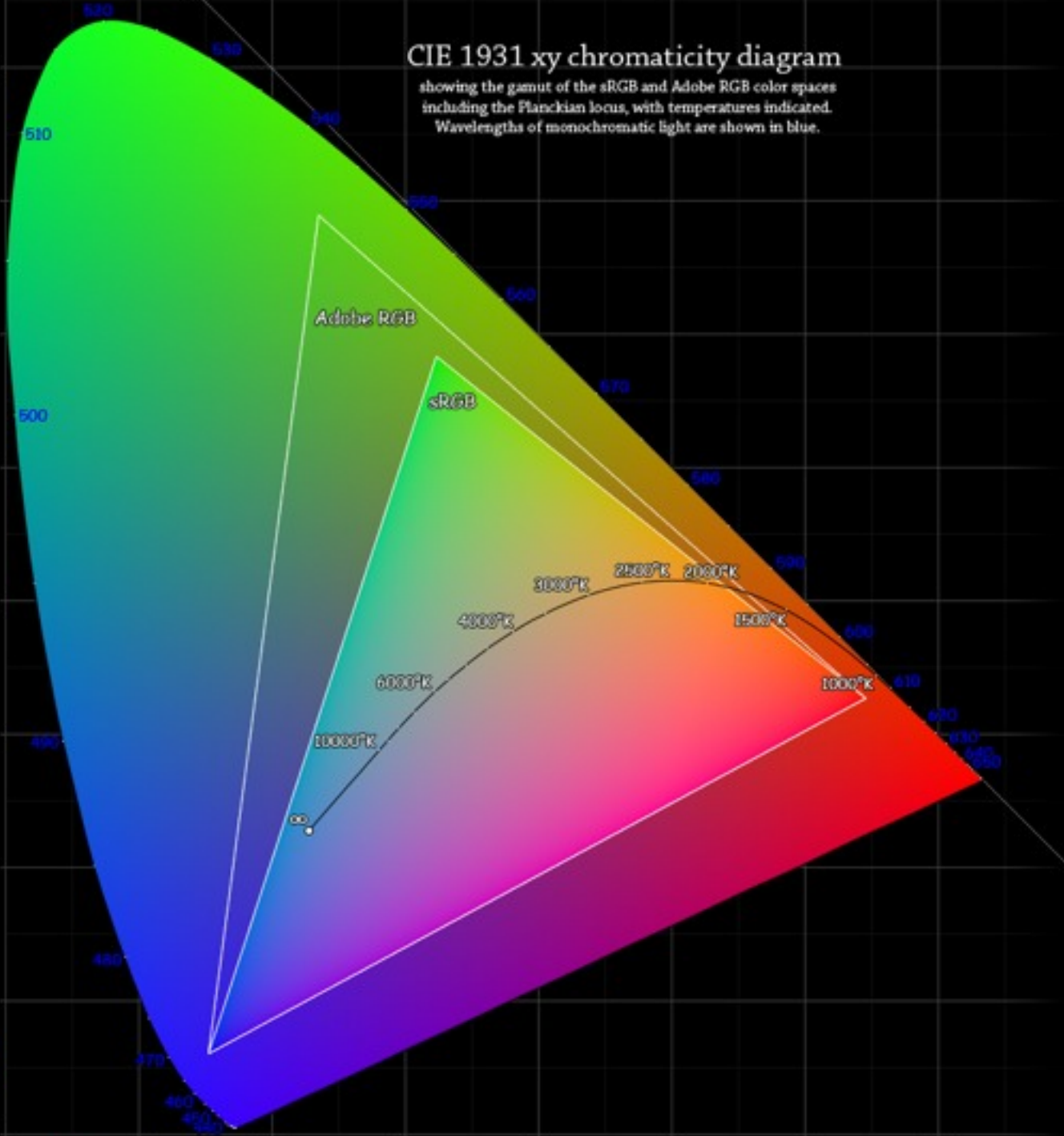


The psychophysics of sensory function [Stevens 61]

Color

CIE 1931 xy chromaticity diagram

showing the gamut of the sRGB and Adobe RGB color spaces including the Planckian locus, with temperatures indicated. Wavelengths of monochromatic light are shown in blue.

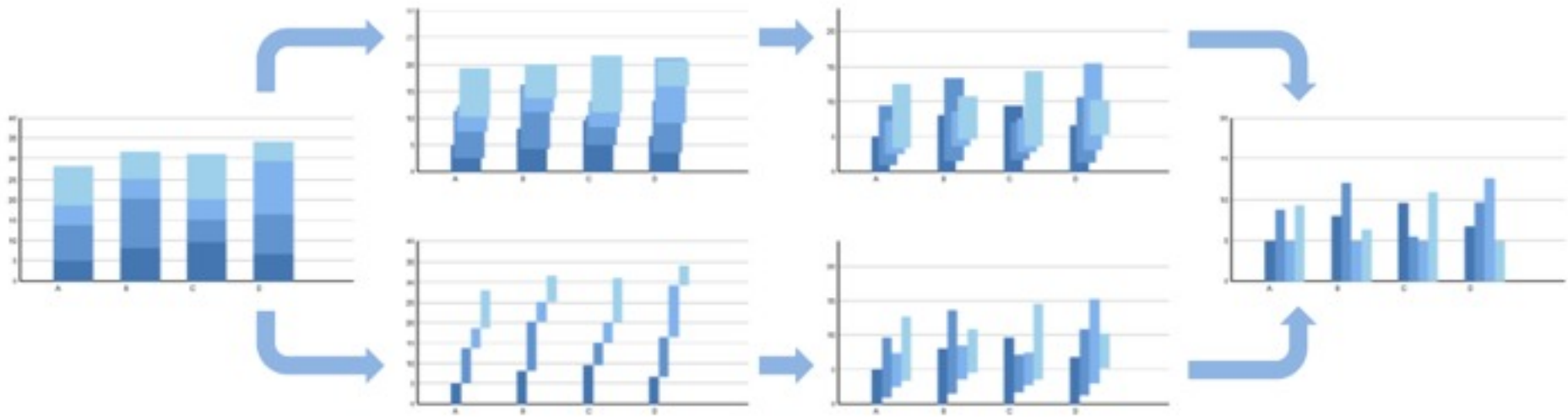


Color Brewer

Interaction

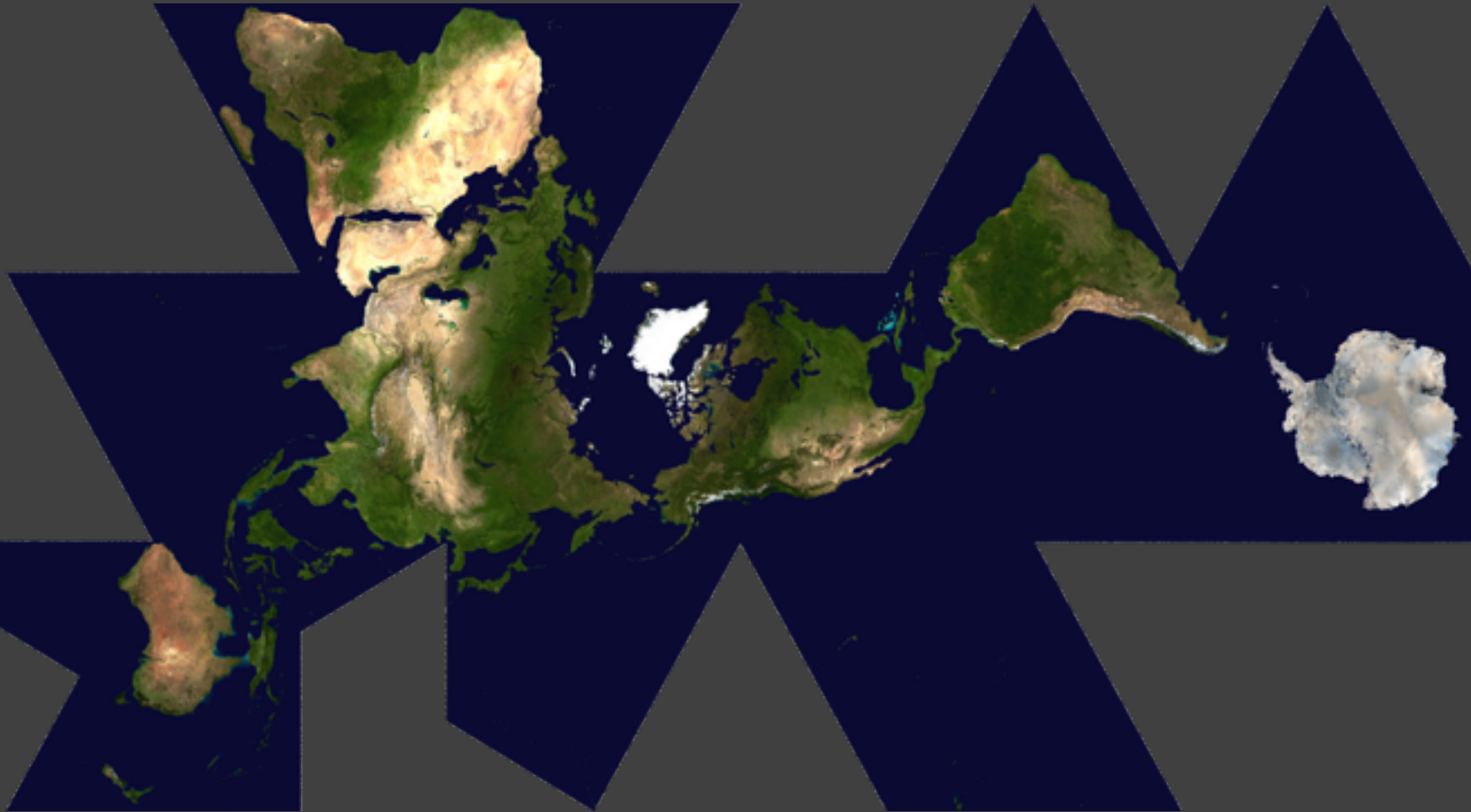


Animation



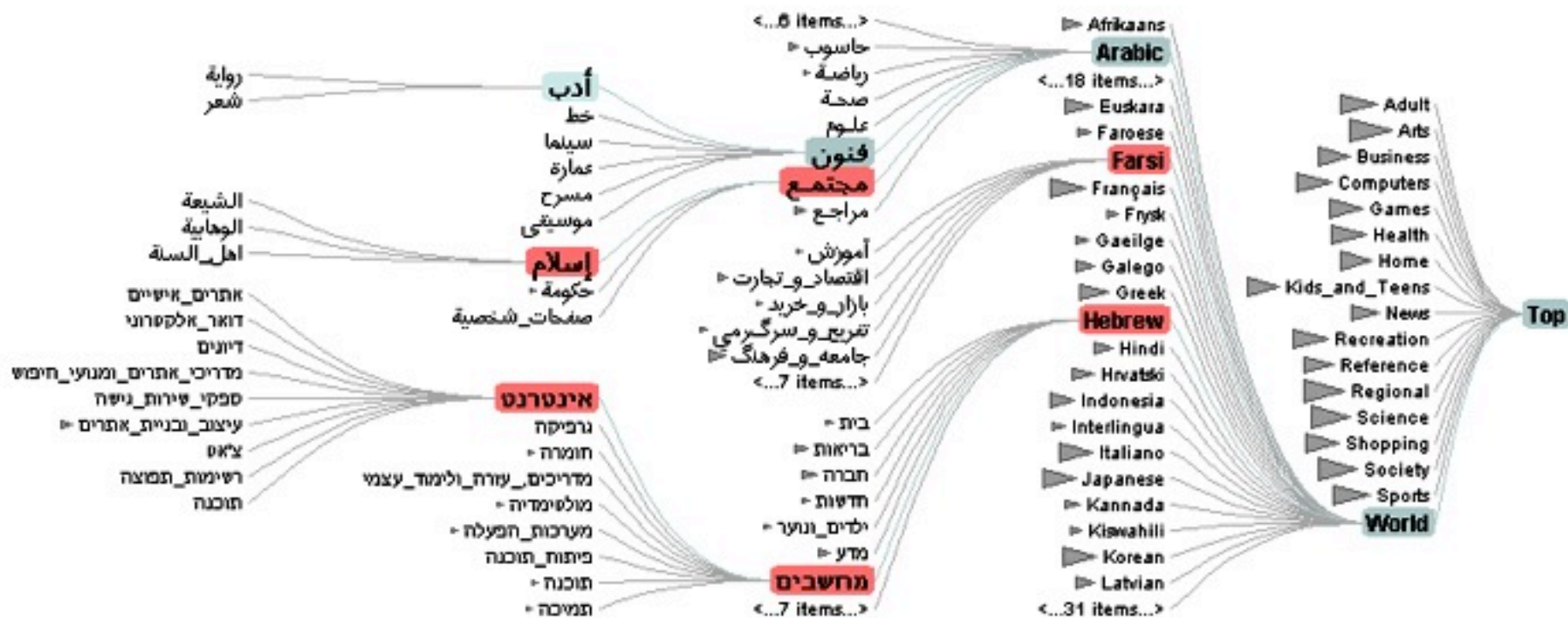
Animated transitions in statistical data graphics [Heer & Robertson 07]

Mapping / Cartography



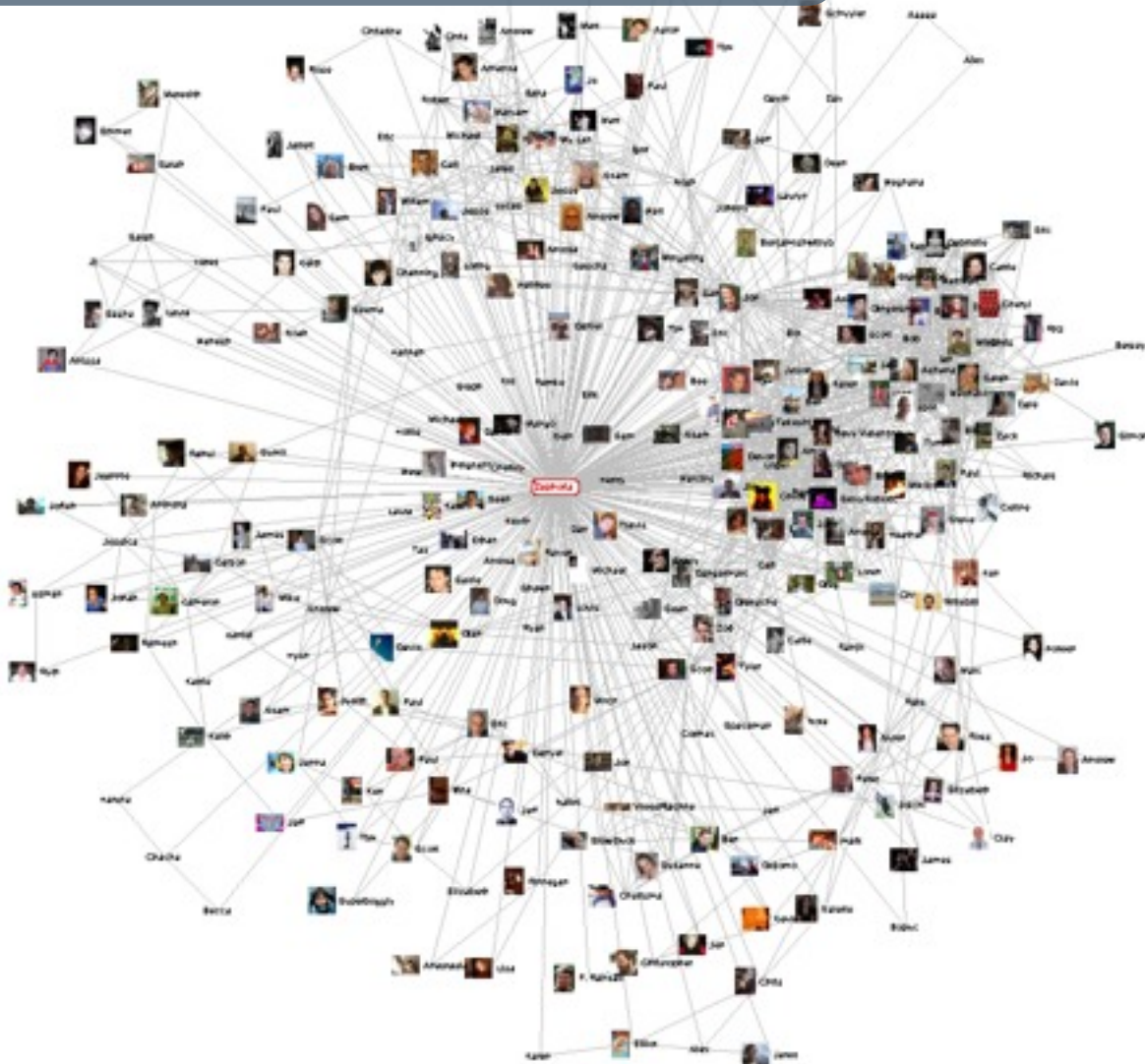
Dymaxion Maps [Fuller 46]

Graphs and Trees



Degree-Of-Interest Trees [Heer & Card 04]

Graphs and Trees



Zephoria

User ID 21721
Friends 288
Age ??
Gender Female
Status Single
Location San Francisco, CA
Hometown Lancaster, PA
Occupation researcher: social networks, identity, context
Interests apophenia, observing people, culture, questioning power, reading, buddhism, ipseity, computer-mediated communication, social networks, technology, anthropology, stamping
Music psytrance/goatrance [Infected Mushroom, Son K9e... Iboga/Digital Structures], Ani Difranco, downtempo, Thievery Corporation, Beth Orton, Morcheeba, Ween, White Stripes
Books Authors: Erving Goffman, Stanley Milgram, Jeanette Winterson, Eric Schlosser, Leslie Feinberg, Dorothy Allison, Italo Calvino, Hermann Hesse
TV Shows ??
Movies Koyaanisqatsi, Amelie, Waking Life, Tank Girl, The Matrix, Clockwork Orange, American Beauty, Fight Club, Boys Don't Cry
Member Since ??
Last Login 2003-10-21
Last Updated 2003-10-21
About [Some know me as danah...]

I'm a geek, an activist and an academic, fascinated by people and society. I see life as a very large playground and enjoy exploring its intricacies. I revel in life's chaos, while simultaneously providing my own insane element.

My musings:
<http://www.zephoria.org/thoughts/>

Want to Meet Someone who makes life's complexities seem simply elegant.

community >>

Enable

search >>

Text Visualization

Visualization of Text by Alberto Gonzales

Creator: Martin Wattenberg
Tags:

explore

- visualizations
- data sets
- comments
- topic hubs

participate

- create visualization
- upload data set
- create topic hub
- register

learn more

- quick start
- visualization types
- data format & style
- about Many Eyes
- FAQ
- blog

contact Us

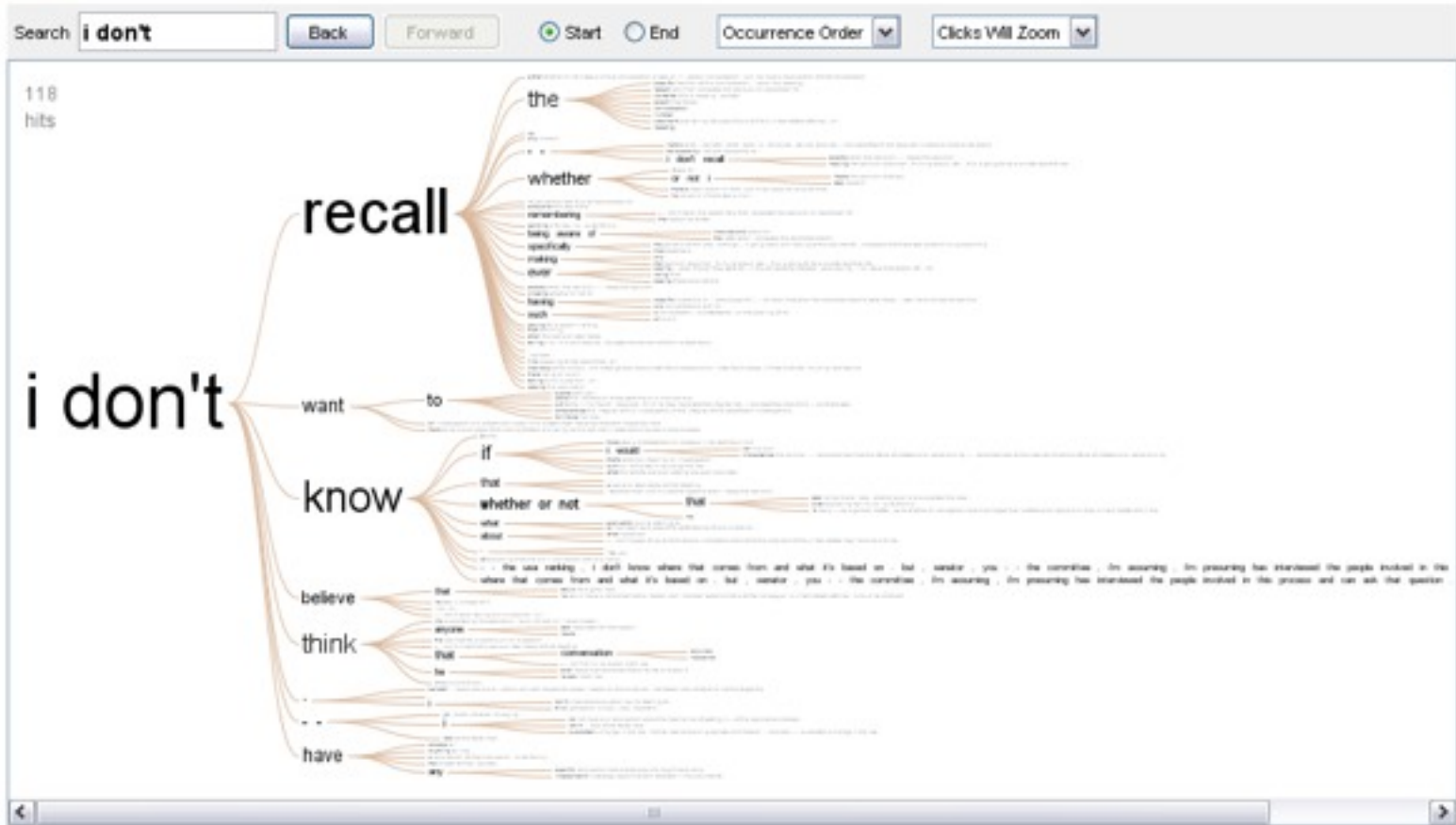
- contact
- report a bug

legal

- terms of use

Popular Dataset Tags

- 2007 2008 bible blog
- books census crime
- education eharmony
- election energy food
- health inauguration
- internet ireland literature
- lyrics media music
- network obama
- people politics
- population
- president prices religion



Data file: [Word in testimony from Gonzales, 4/19/2007](#) Data source: CQ Transcript Wire via the Washington Post

This data set has not yet been rated

- Full image
- share this
- watch this
- add to topic hub
- rate this

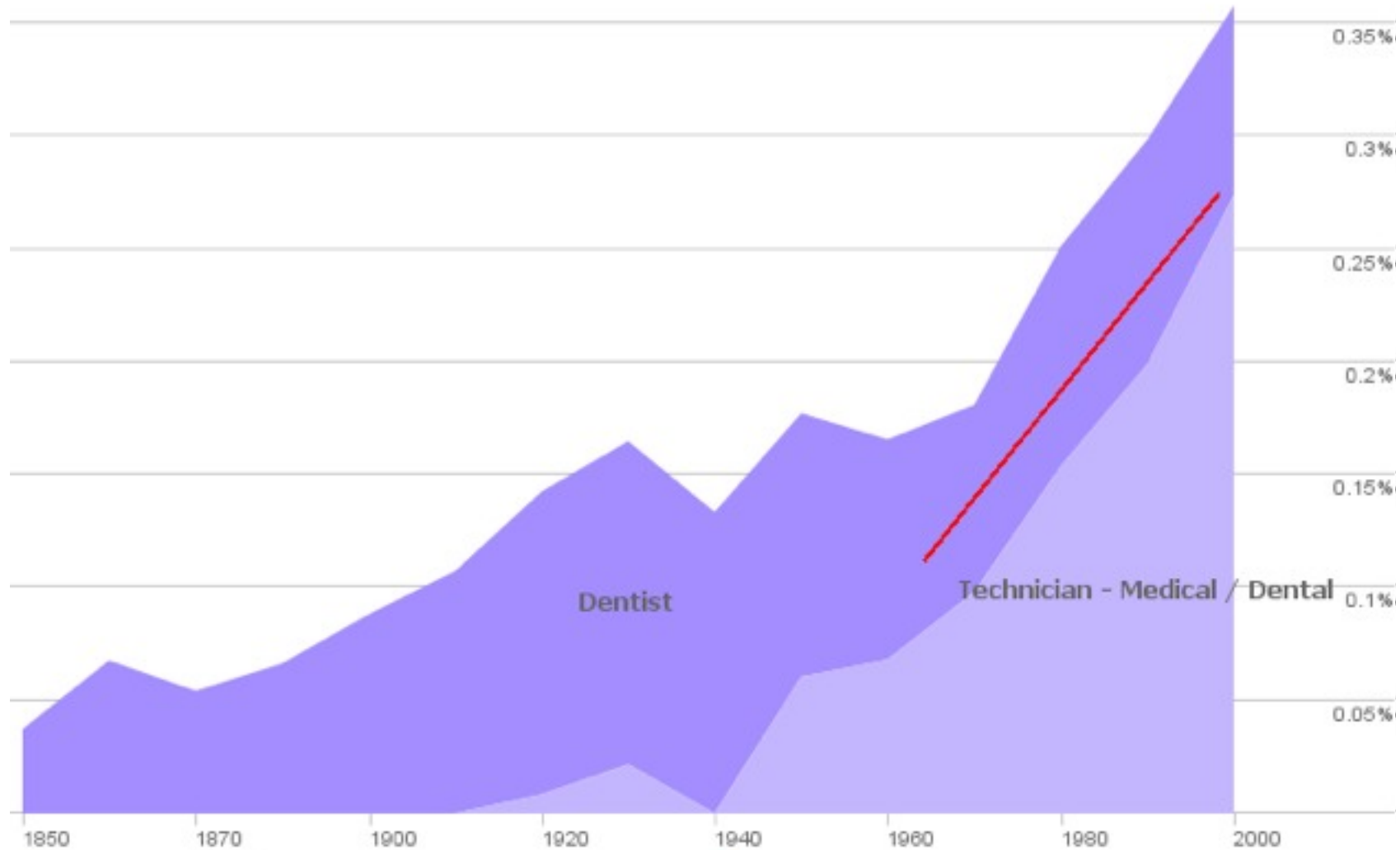
Comments (4)

currently showing

This visualization has 4 positive and 0 negative

Collaboration and History

Where have all the dentists gone?



The Future of Visualization

Where is more work required?

What emerging technologies and societal trends will impact visualization design?

What did you find most difficult in creating visualizations and designing techniques?

Thank You!



Course Evaluation

Course Evaluation

Official campus course evaluation

Complete in class, then give to Ham.

Your opinion is valued!