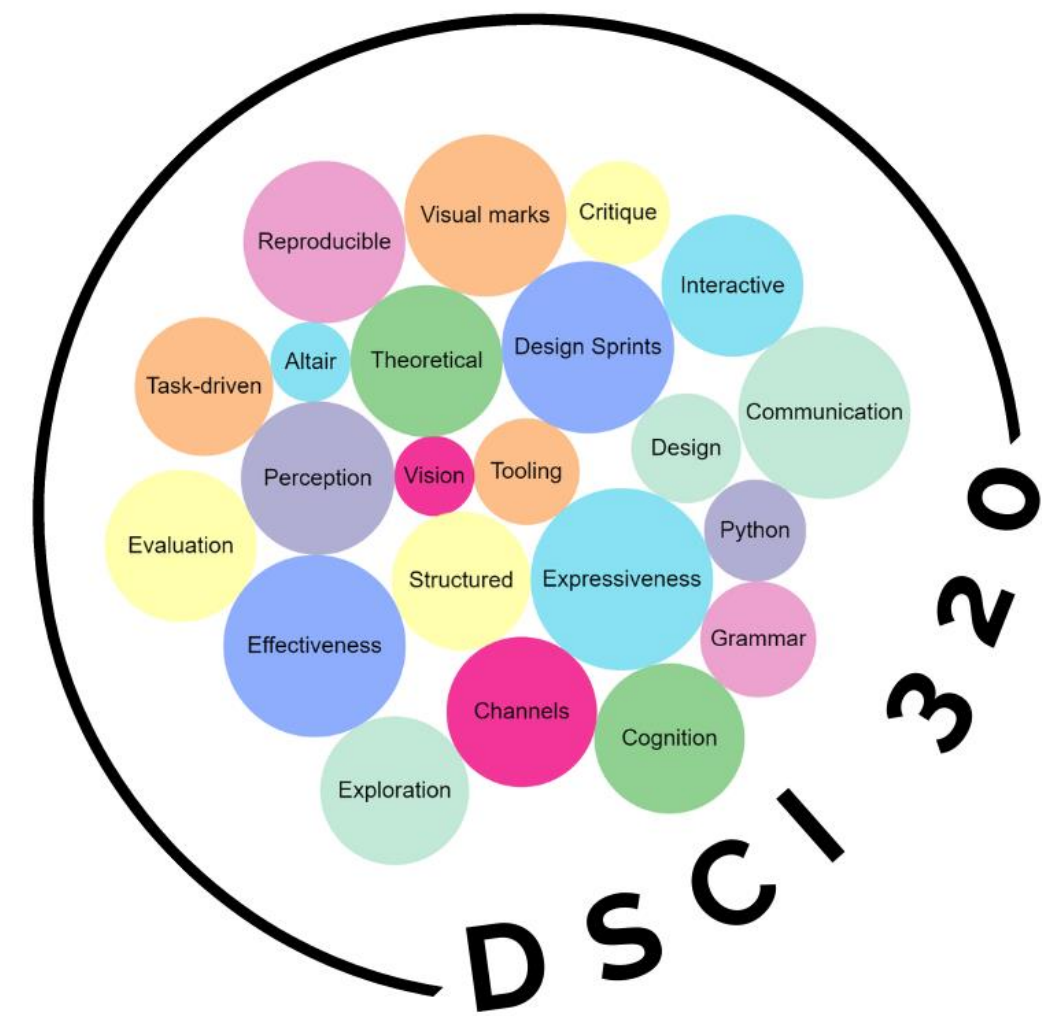
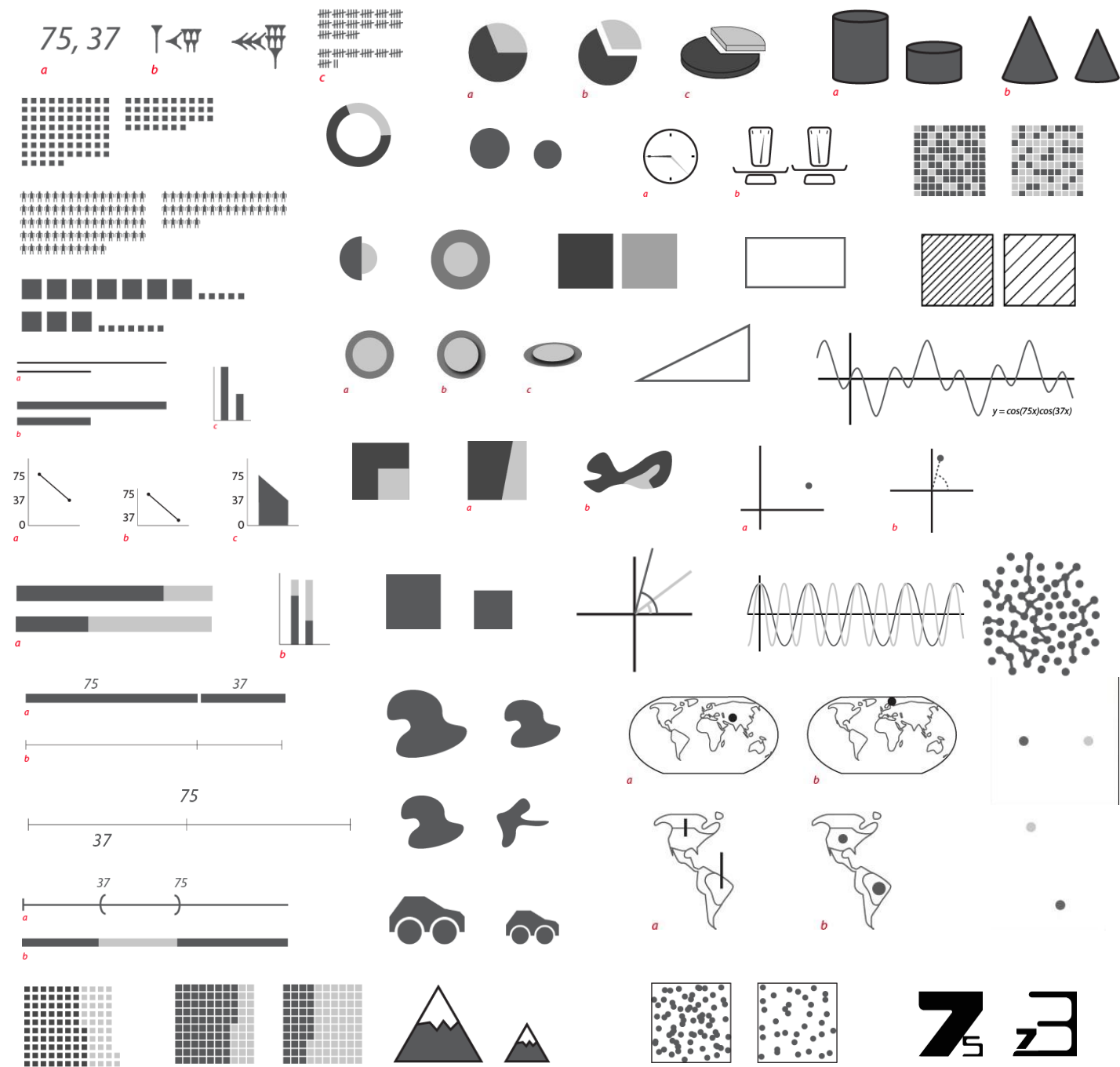


Visualization for Data Science

Visual Marks & Channels II

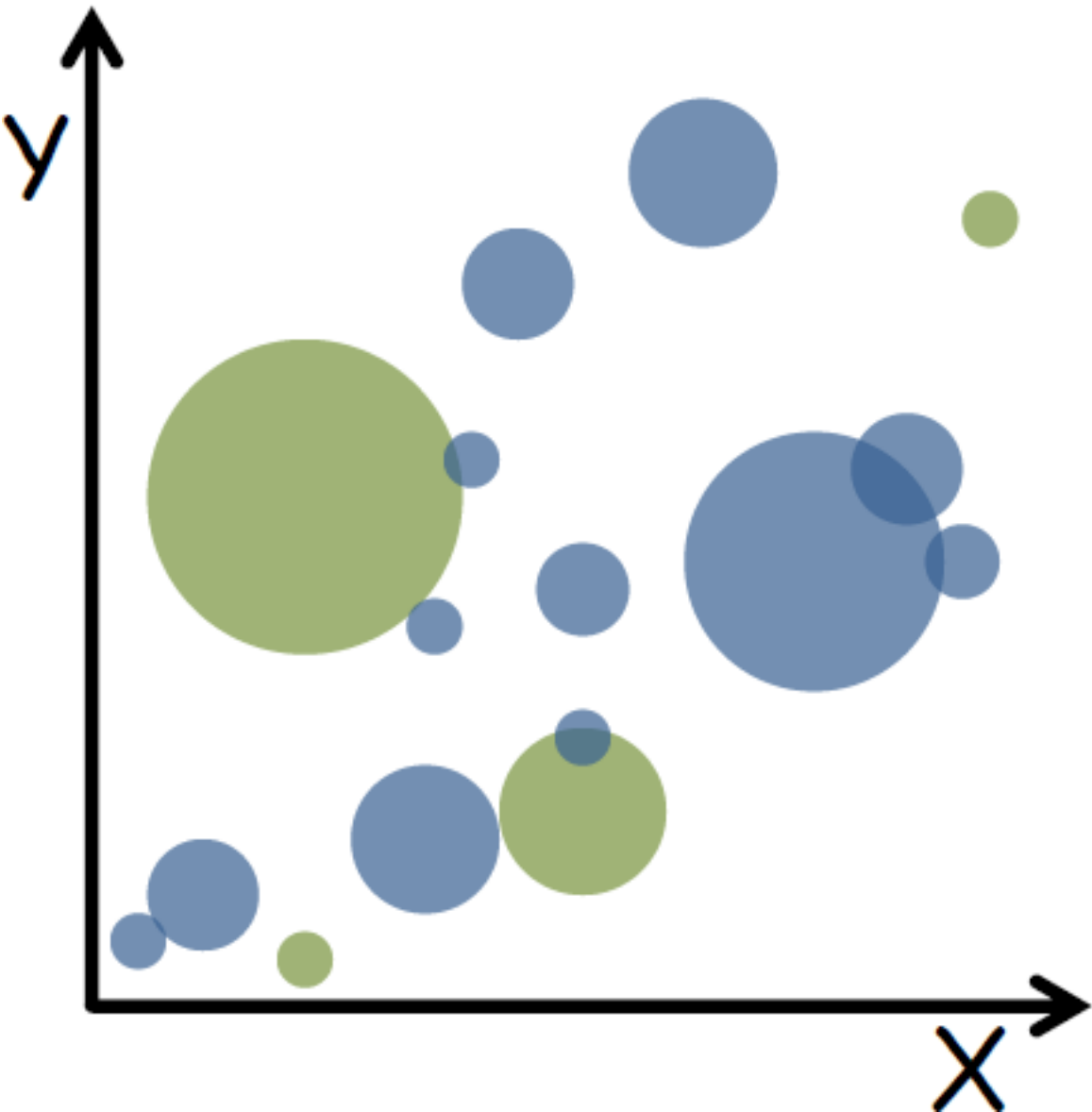


75 and 37 revisited



This resource shows 45 ways <https://rockcontent.com/blog/45-ways-to-communicate-two-quantities/>

Recap



MARK:

➔ Points



➔ Lines



➔ Areas



CHANNEL:

➔ Position

➔ Horizontal



➔ Vertical



➔ Both



➔ Color



➔ Shape



➔ Tilt



➔ Size

➔ Length



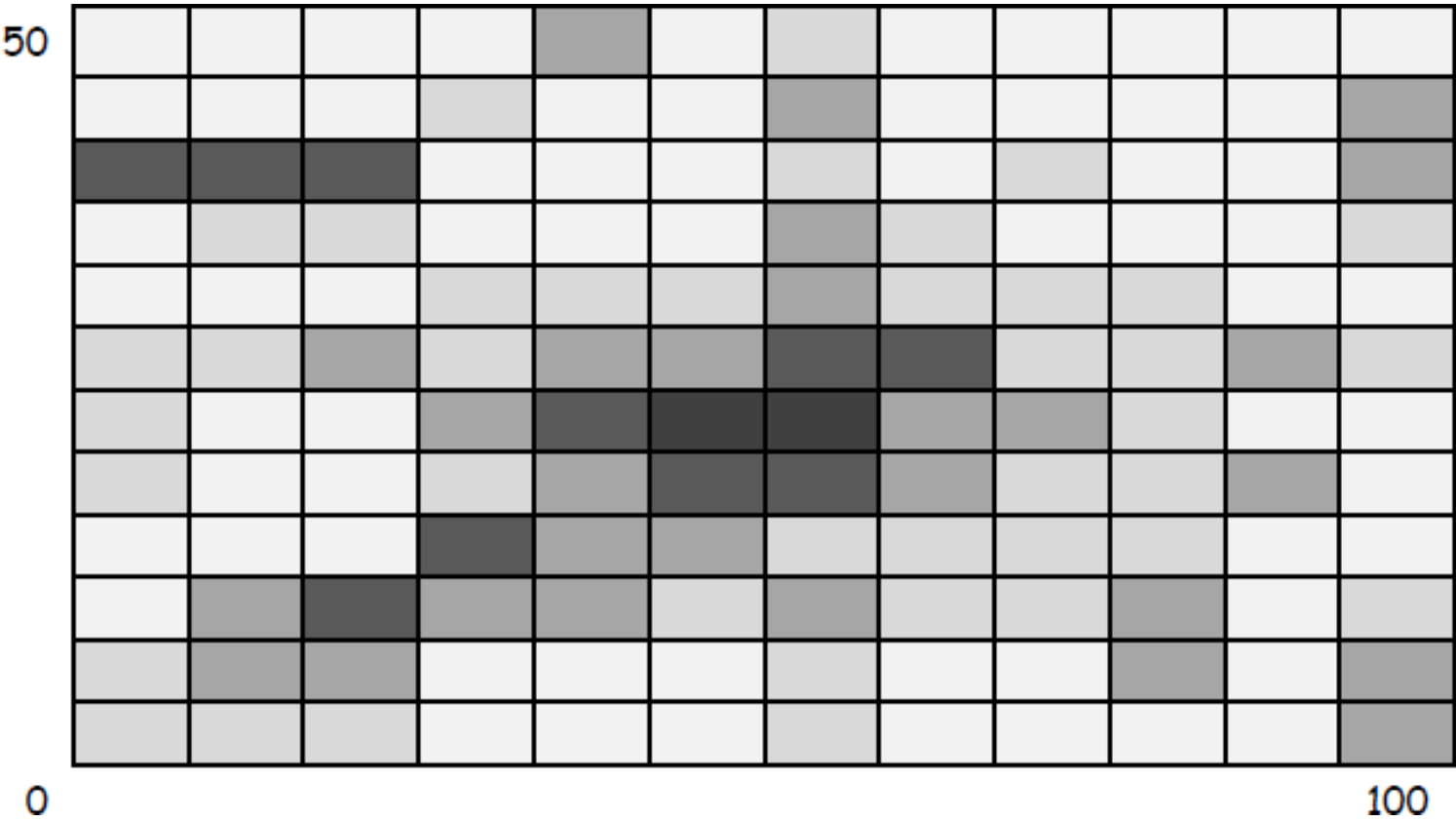
➔ Area



➔ Volume



Recap



MARK:

→ Points

→ Lines

→ Areas

CHANNEL :

→ Position

→ Horizontal

→ Vertical

→ Both

→ Color

→ Shape

→ Tilt

→ Size

→ Length

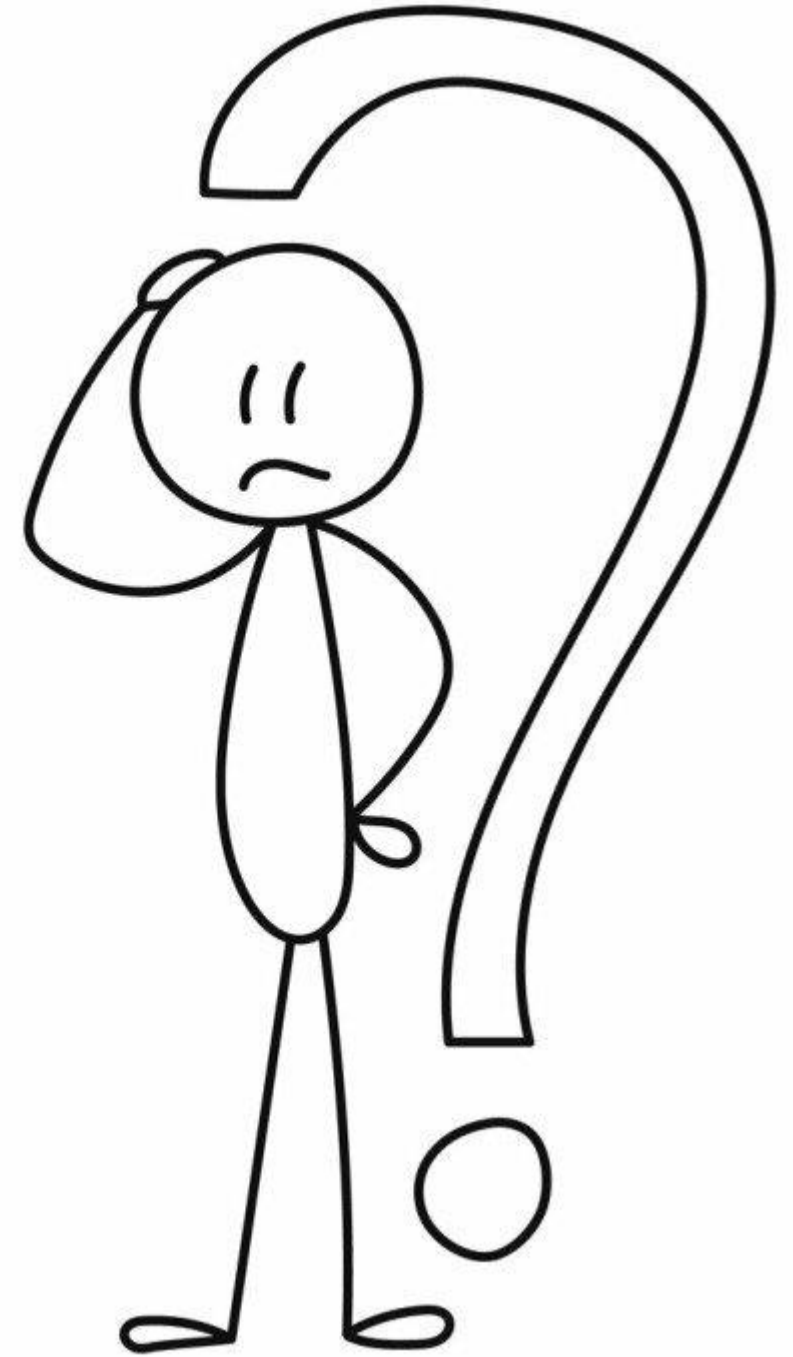
→ Area

→ Volume

Learning Outcomes

- Learn which marks and channels are most effective for a given task (“perceptual ordering”)
- Describe a visualization using appropriate viz grammar
- Differentiate between effectiveness and expressiveness
- Critique a visualization based on the use of its effectiveness and expressiveness
- Compare and contrast the benefits and limitations of channels for a given task
- List and describe the various characteristics of channels

How do I pick
which marks or
channels to use?



Channel Selection

Before we can select channels we must understand how its various properties influence their use.

We must consider both

- **Expressiveness:** match the channel type to data characteristics: the visual encoding should express all of, and only, the information in the dataset attributes.
- **Effectiveness:** for a given task, some channels are better than others so it is important to select the most effective channel for the data

Channel Characteristics

- Discriminability: how many unique steps can we perceive?
- Separability: is our ability to use this channel affected by another one?
- Popout: can things jump out using this channel?
- Grouping: can a channel show perceptual grouping of items?
- Accuracy: how precisely can we tell the difference between encoded items?

Discriminability: How many usable steps?

- How many usable steps?
- How easily can differences between attribute levels be perceived?

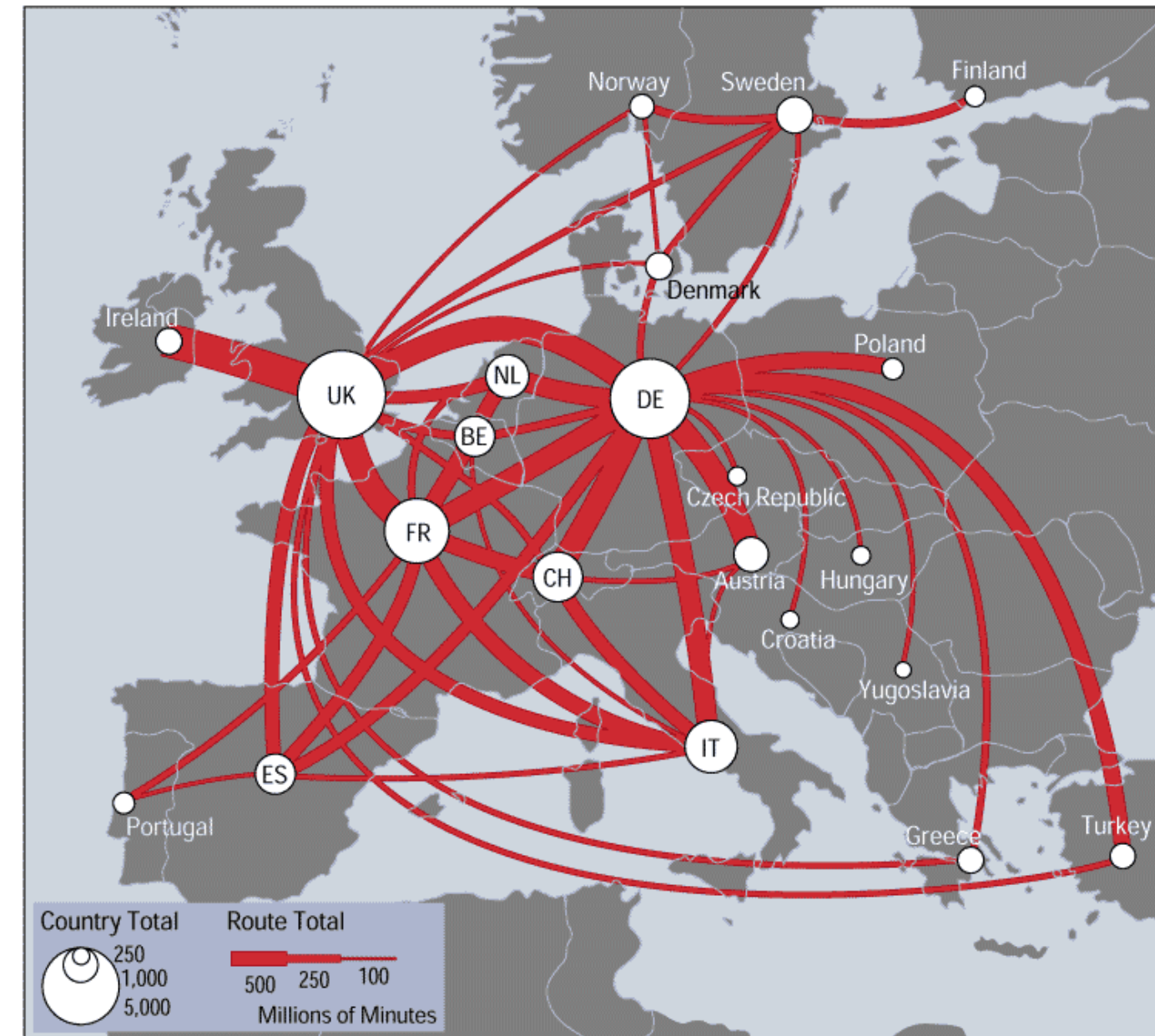
Tips

The channel must be sufficient for number of attribute levels to show

- Linewidth – limited number of steps (maybe 4 at most)
- Color hue – max 5, using more isn't recommended
- Shapes – max 5, using more can create difficulties

ColorBrewer demo:

<https://colorbrewer2.org/#type=qualitative&scheme=Paired&n=3>



READING, WRITING, AND EARNING MONEY

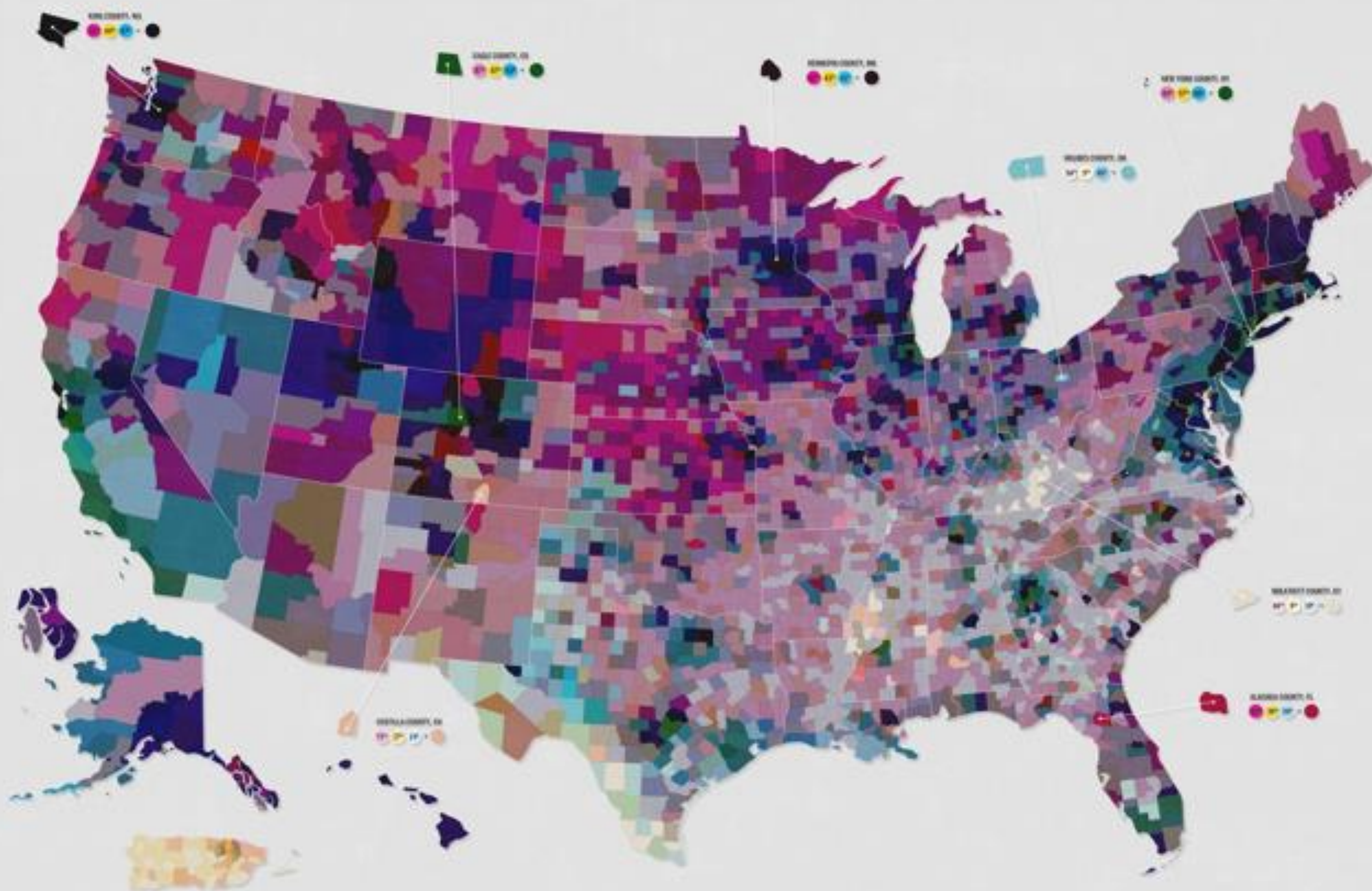
The 2000 Census shows that 15% of the country's population is currently living in poverty. This is a significant increase from the 12% reported in the 1990 Census. The map shows the distribution of poverty across the country, with the highest concentrations in the South and the West.



The map at right is a product of overlaying the three sets of data. The variables are not yet fully integrated, but the data shows a clear correlation between poverty and the other two variables. The map shows that poverty is highest in the South and the West, and lowest in the Northeast and the Midwest.



A correlation between poverty and the other two variables.



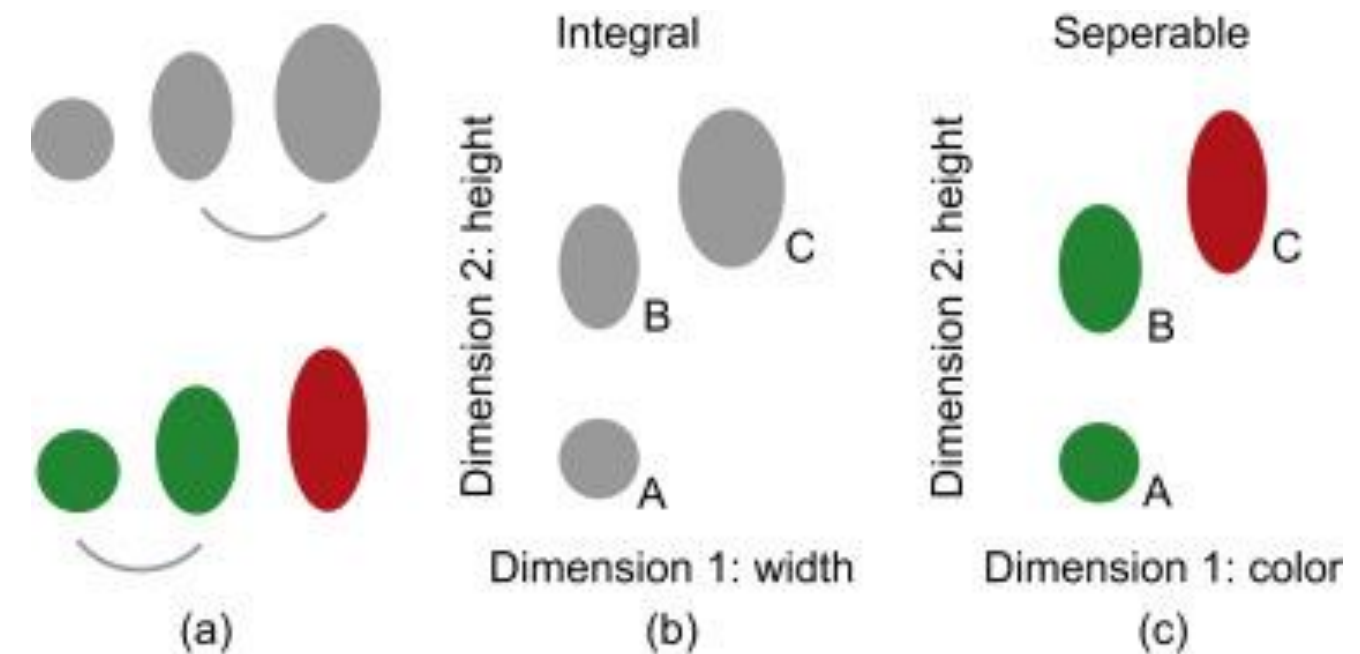
Separability vs. Integrality

Integral dimensions: two or more attributes are perceived holistically (not independently)

Separable dimensions: people tend to make separate judgements about each dimension (i.e., attribute)

Separability is our ability to use this channel affected by another one?

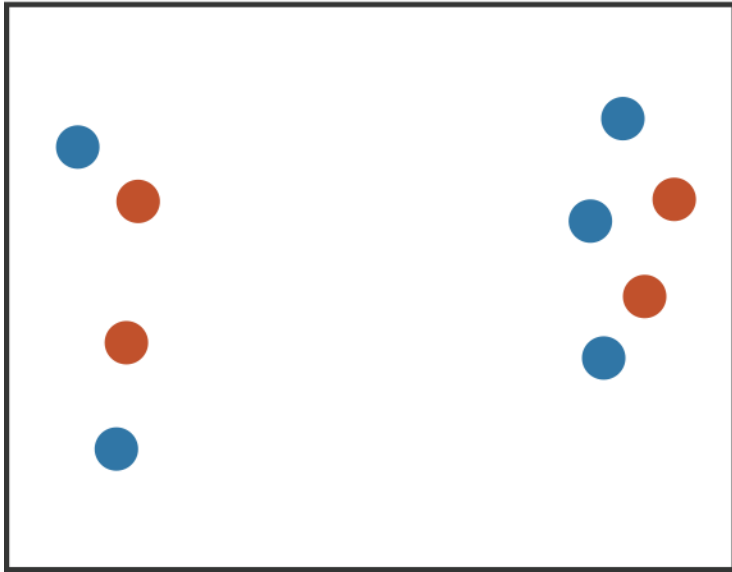
Can channels be used independently or is there interference from other channels being used?



(a) The width and height of an ellipse are perceived integrally, so the ellipses are seen as more similar to each other (because they have the same shape) than the pair having the same width. The color and height of a shape are perceived separably, so the two green shapes are seen as most similar. (b, c) Space plots of the two examples.

Separability vs. Integrality

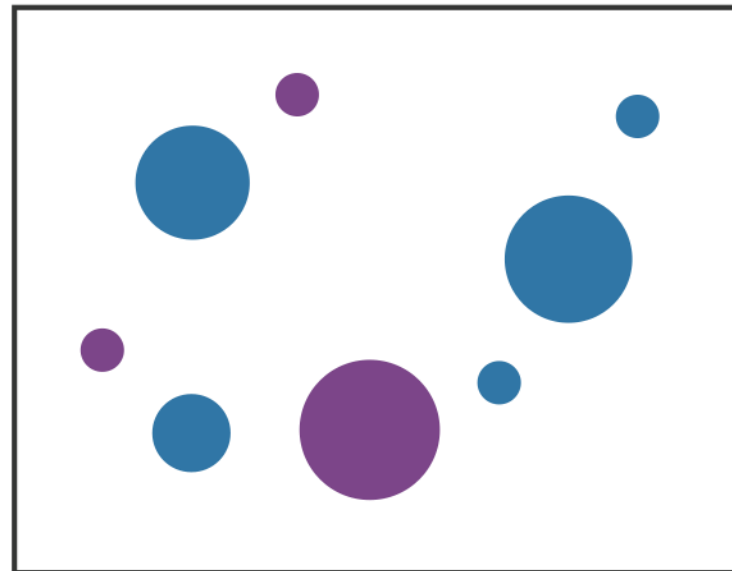
Position
+ Hue (Color)



Fully separable

2 groups each

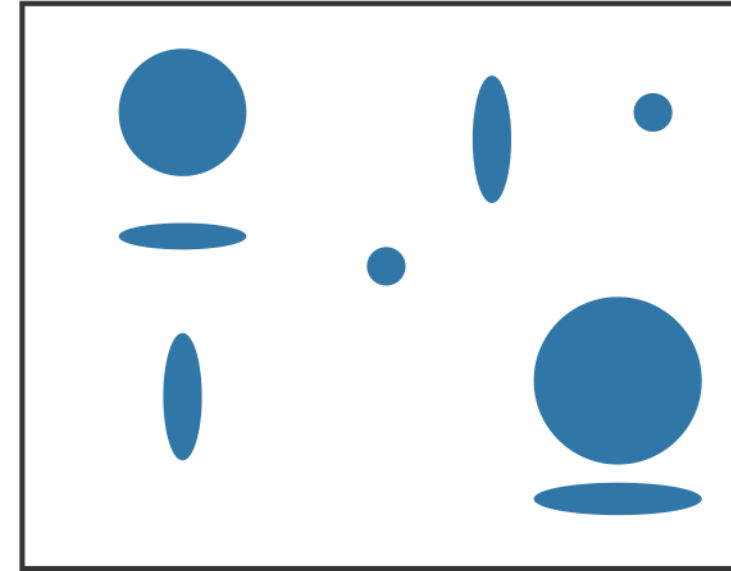
Size
+ Hue (Color)



Some interference

2 groups each

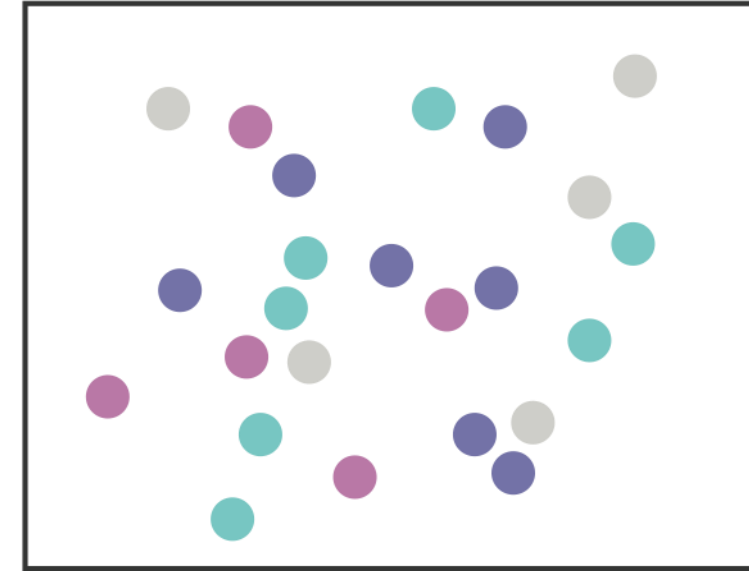
Width
+ Height



Some/significant
interference

3 groups total:
integral area

Red
+ Green



Major interference

4 groups total:
integral hue

All Spending

Types of Spending

Changes

Department Totals

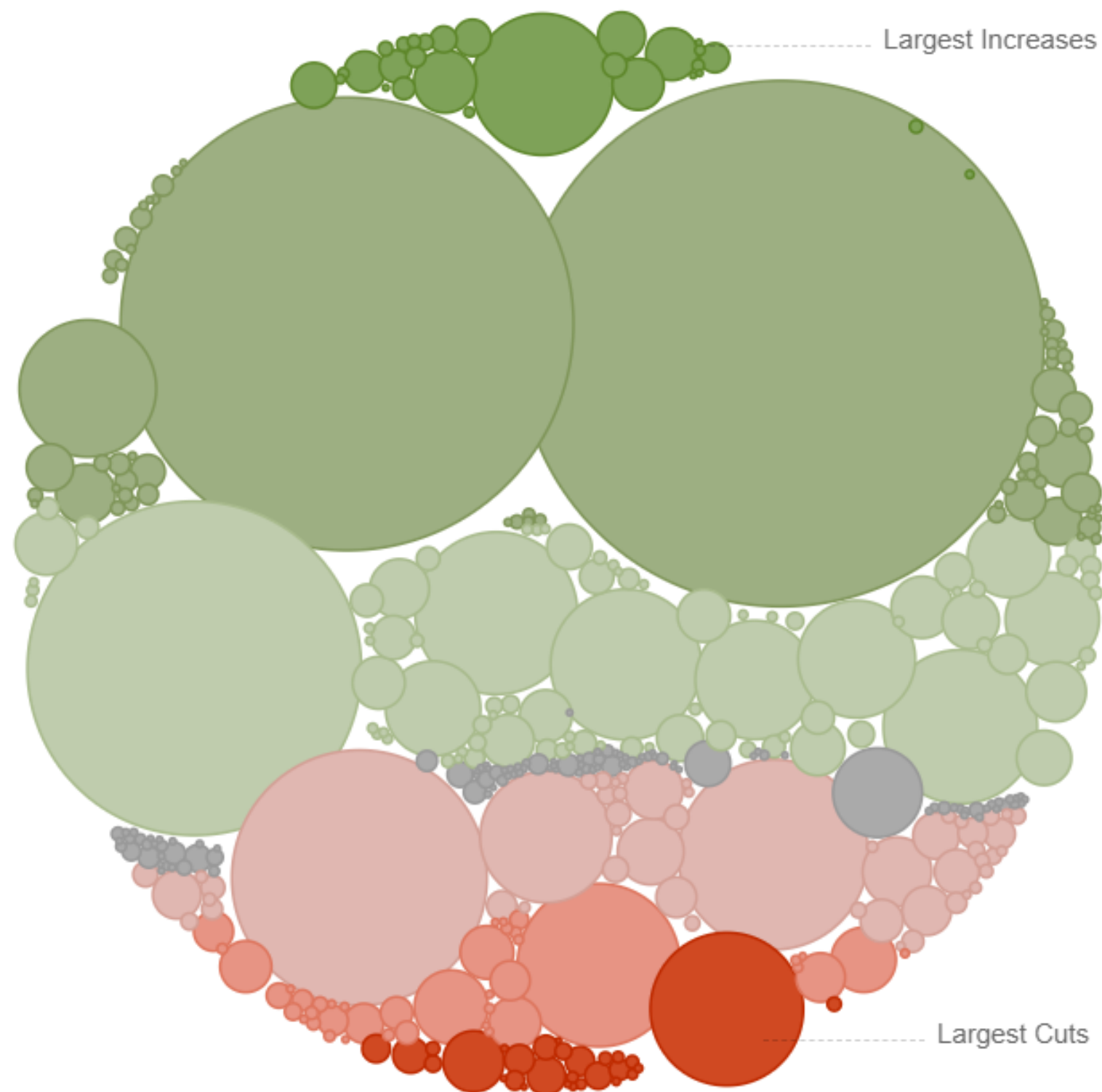
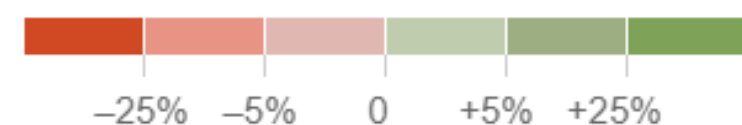
How \$3.7 Trillion Is Spent

Mr. Obama's budget proposal includes \$3.7 trillion in spending in 2013, and forecasts a \$901 billion deficit.

Circles are sized according to the proposed spending.



Color shows amount of cut or increase from 2012.



Popout/Ease of recognition

Can things jump out using this channel?

Can a channel provide popout where a difference is perceived preattentively?

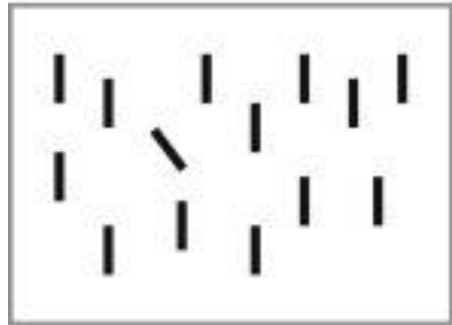
- Properties detected by the low-level visual system
- very rapid - 200-250 ms
- very accurate
- processed in parallel
- happens before focused attention -> **preattentive**

Popout

- find the red dot
 - how long does it take?
- parallel processing on many individual channels
 - speed independent of distractor count
 - speed depends on channel and amount of difference from distractors
- serial search for (almost all) combinations
 - speed depends on number of distractors

Popout

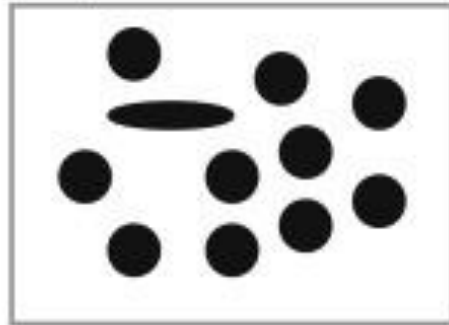
Orientation



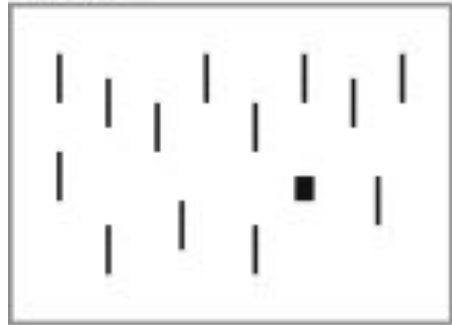
Curved straight



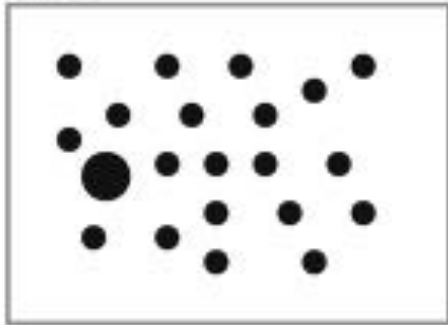
Shape



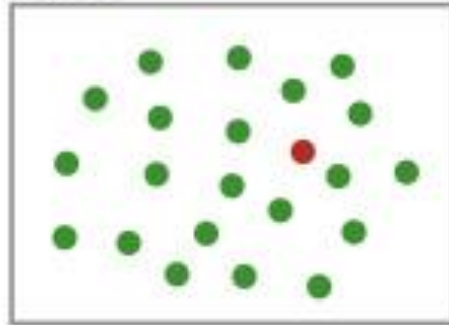
Shape



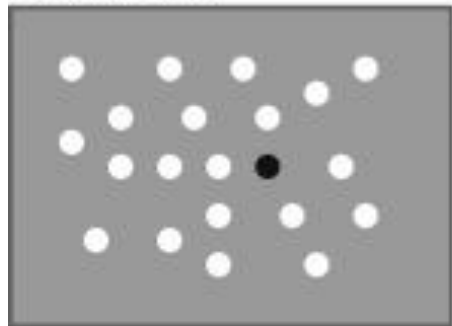
Size



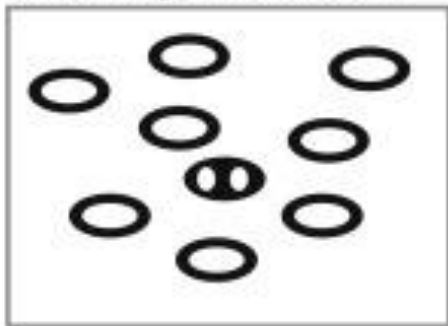
Color



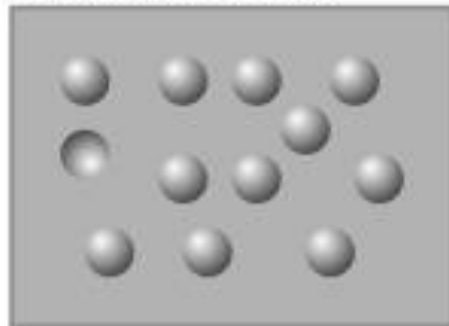
Light/dark



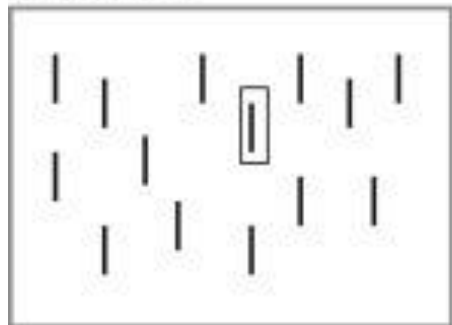
Topology (or count)



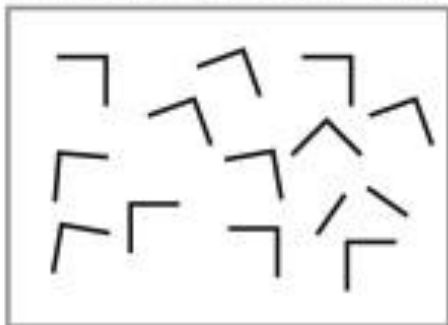
Convex/concave



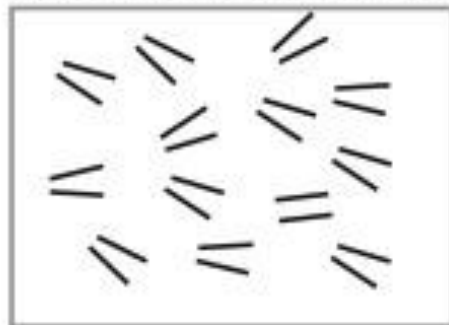
Addition



Juncture (not pre-att)



Parallelism (not pre-att)



Information Visualization by Colin Ware. Ch. 5. Figure 12 [Visual Saliency: Finding and Reading Data Glyphs](#)

All Spending

Types of Spending

Changes

Department Totals

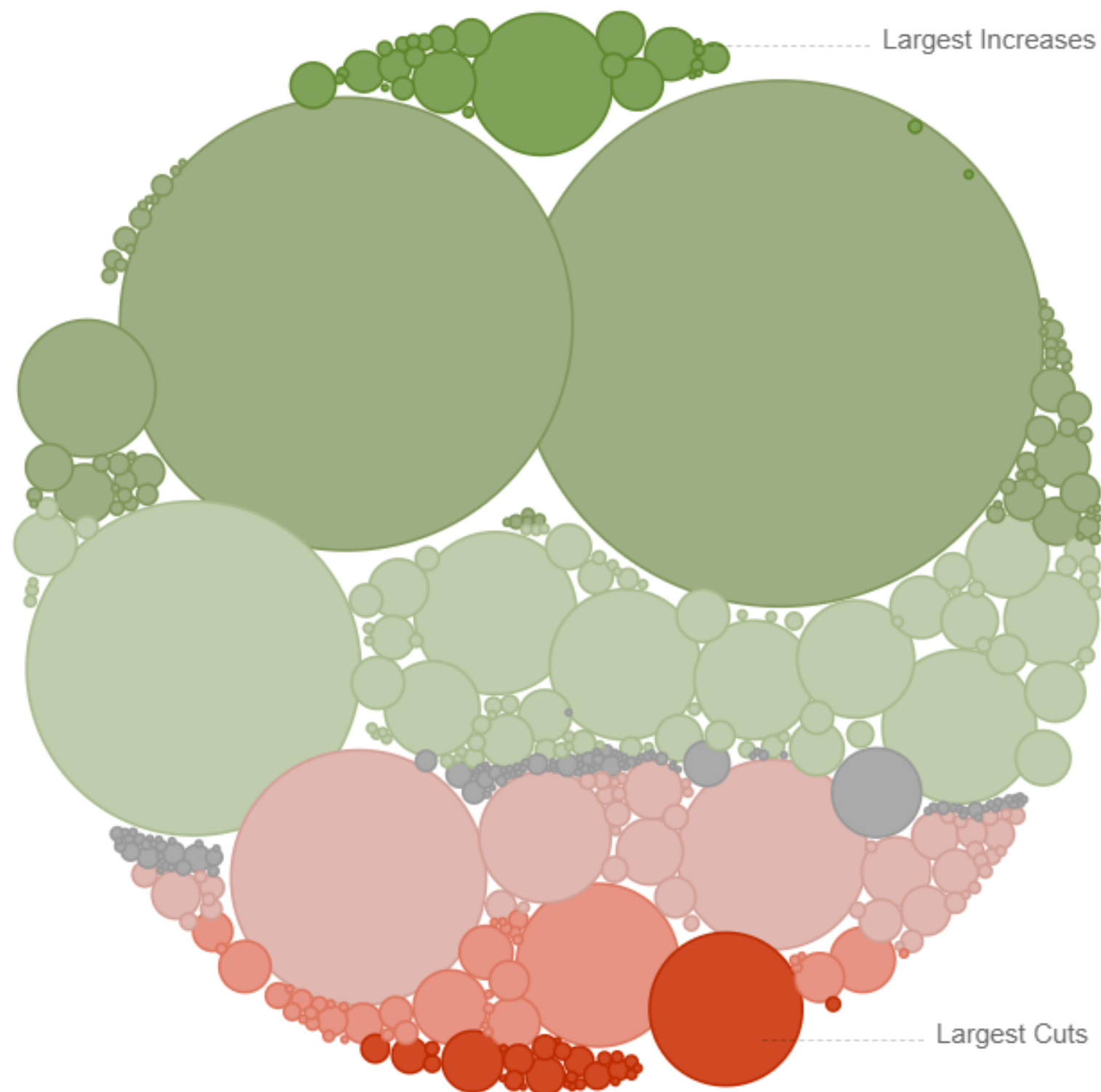
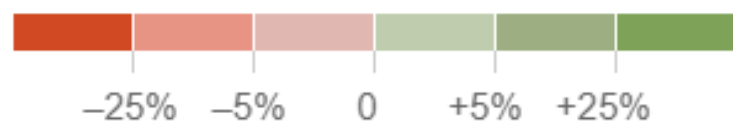
How \$3.7 Trillion Is Spent

Mr. Obama's budget proposal includes \$3.7 trillion in spending in 2013, and forecasts a \$901 billion deficit.

Circles are sized according to the proposed spending.



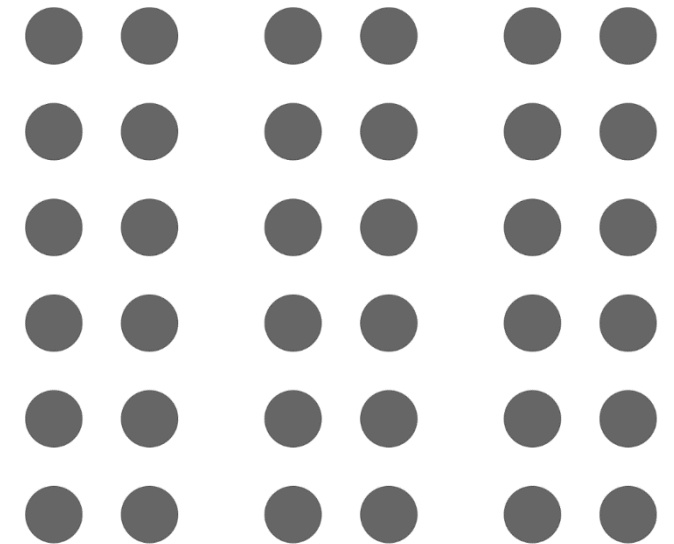
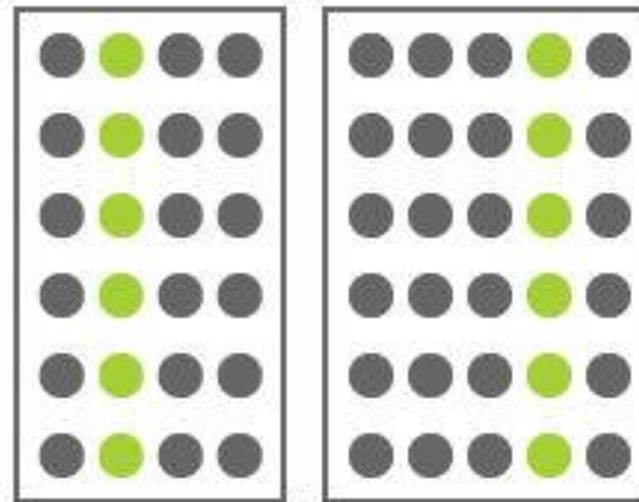
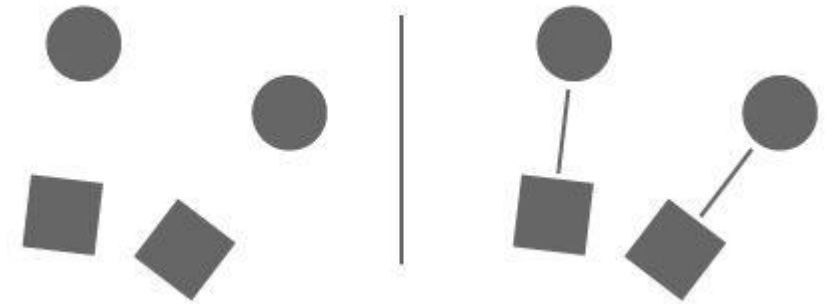
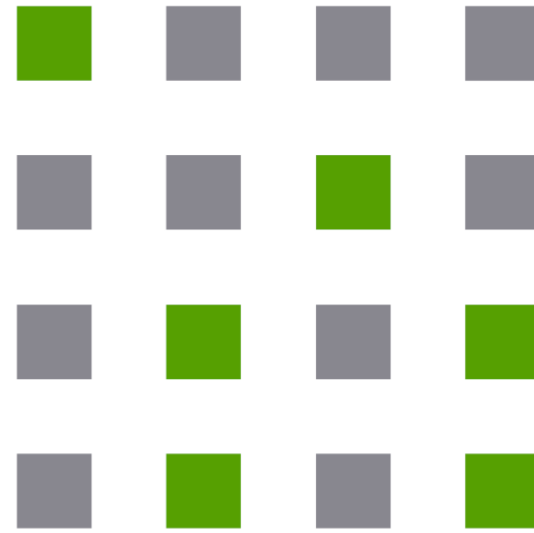
Color shows amount of cut or increase from 2012.



Grouping

can channel show
perceptual grouping of
items?

- A. containment
- B. connection
- C. proximity
- D. similarity



Grouping

can channel show
perceptual grouping of
items?

- containment
- connection
- proximity
- similarity

➔ Containment



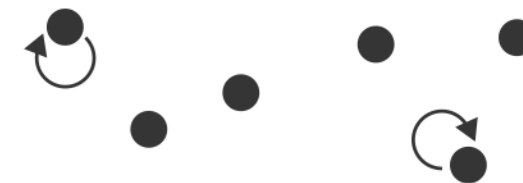
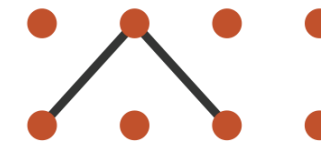
Spatial region

Color hue

Motion

Shape

➔ Connection



All Spending

Types of Spending

Changes

Department Totals

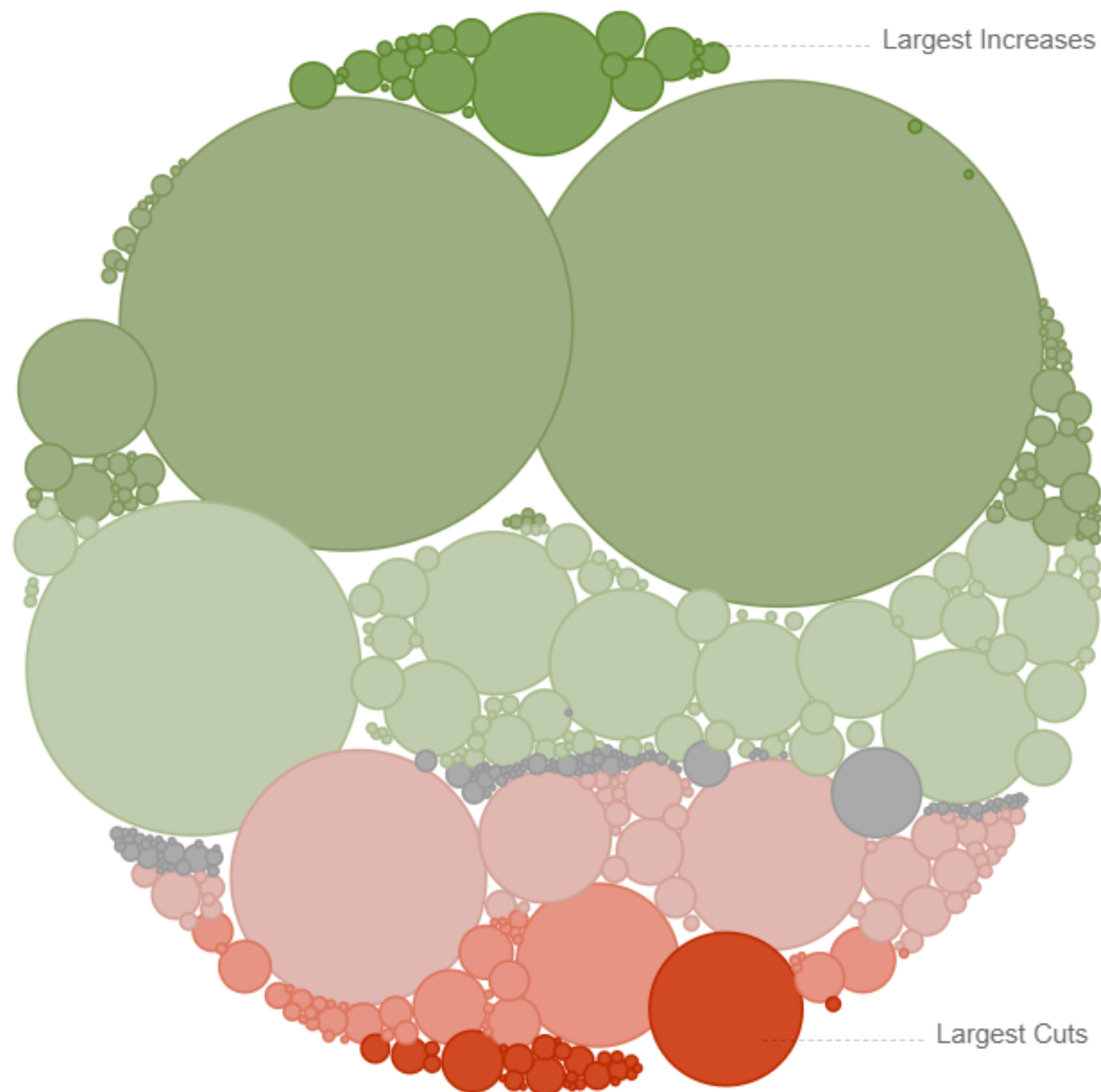
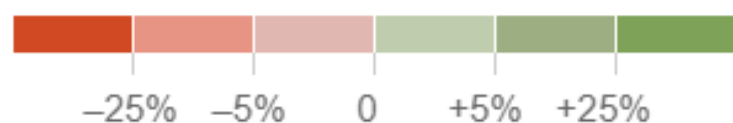
How \$3.7 Trillion Is Spent

Mr. Obama's budget proposal includes \$3.7 trillion in spending in 2013, and forecasts a \$901 billion deficit.

Circles are sized according to the proposed spending.



Color shows amount of cut or increase from 2012.



Accuracy: Fundamental theory

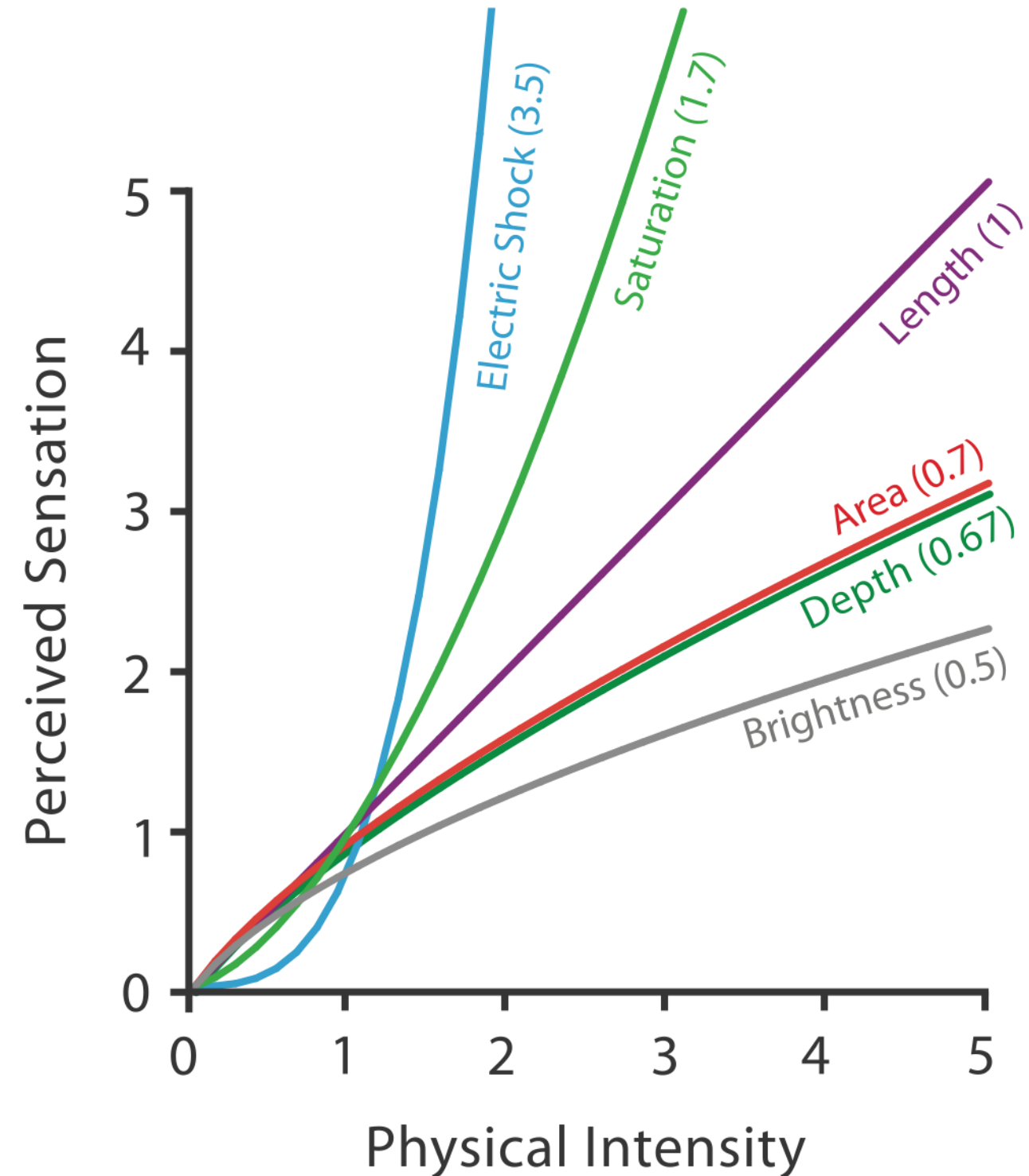
How precisely can we tell the difference between encoded items?

length is accurate:
linear

others magnified or compressed

–exponent characterizes

Steven's Psychophysical Power Law: $S = I^N$



S = sensation

I = intensity

Accuracy: User studies

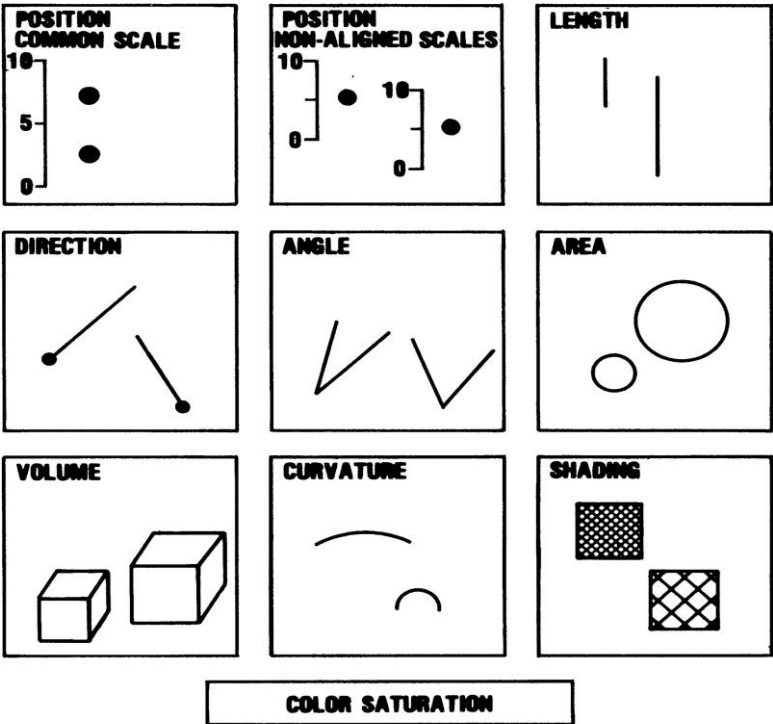


Figure 1. Elementary perceptual tasks.

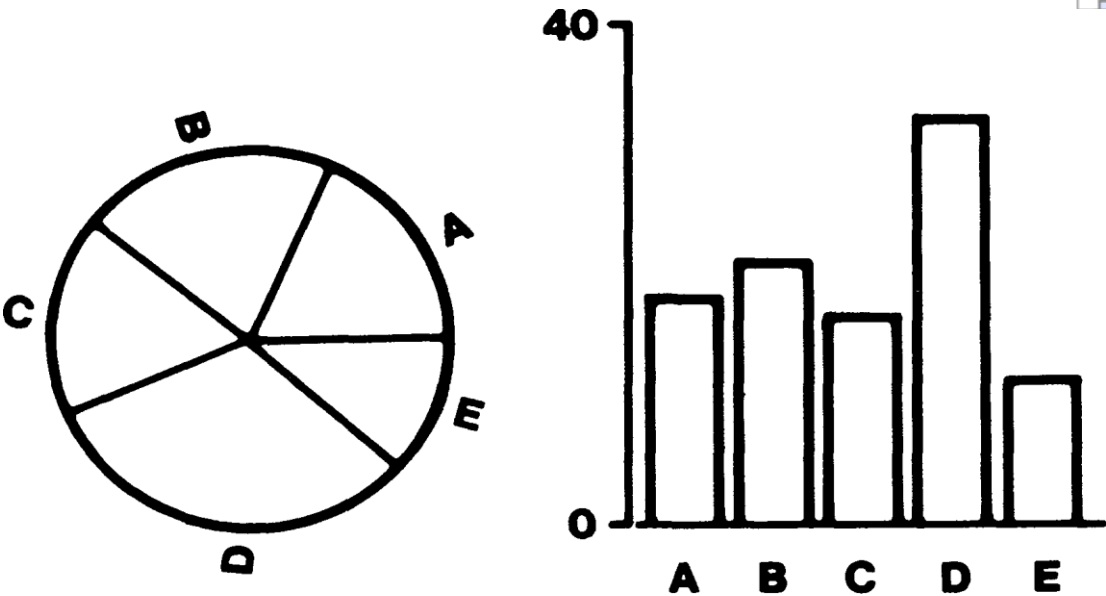


Figure 3. Graphs from position–angle experiment.

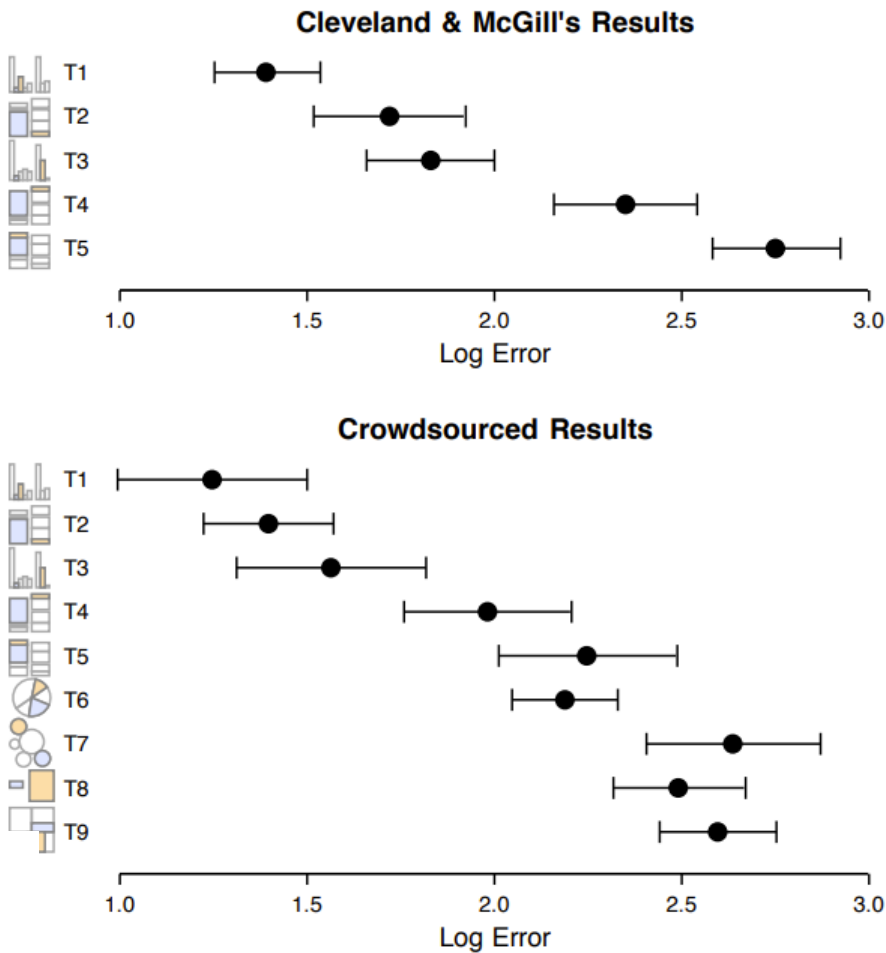
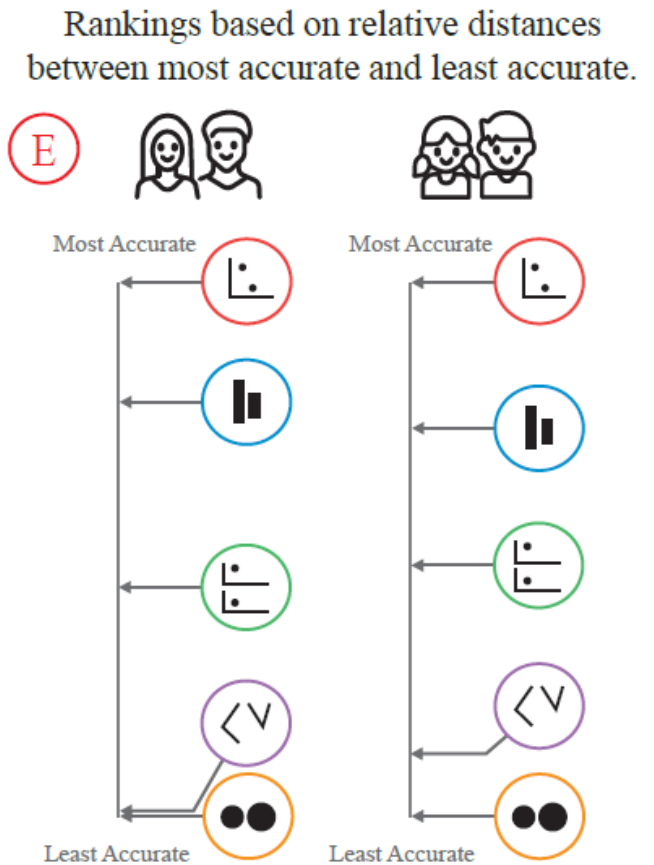
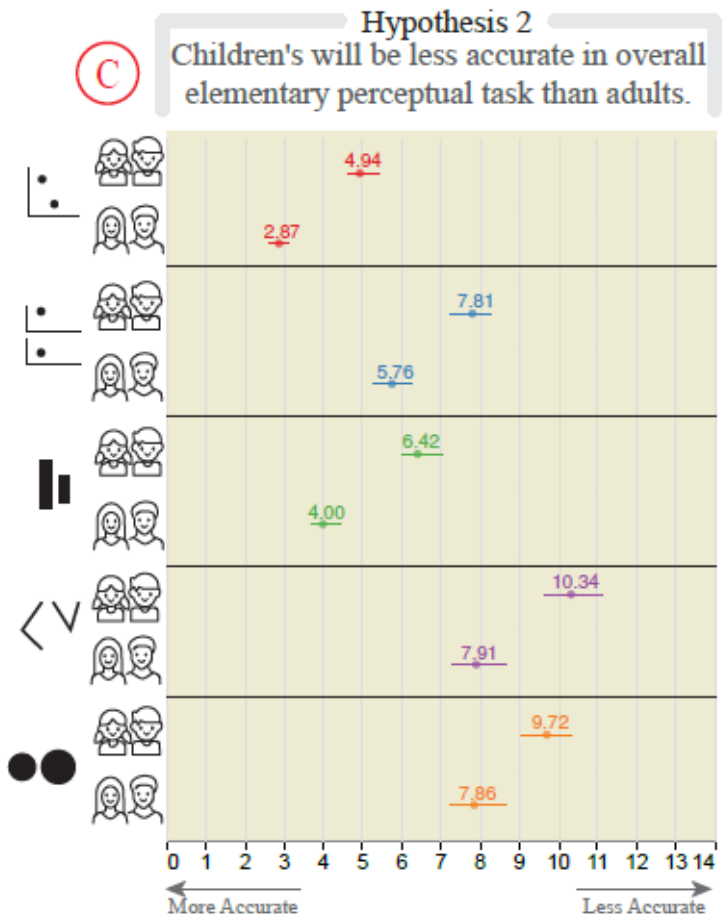


Figure 4: Proportional judgment results (Exp. 1A & B). Top: Cleveland & McGill's [7] lab study. Bottom: MTurk studies. Error bars indicate 95% confidence intervals.



[Cleveland & McGill, 1984](#)
[Mackinlay, 1986](#)
[Heer & Bostock, 2010](#)
[Panavas et al., 2022](#)

Factors affecting accuracy

- alignment
- distractors
- distance
- common scale / alignment

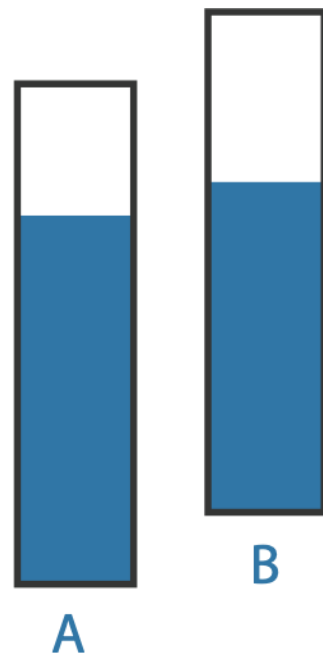


Relative vs. absolute judgements

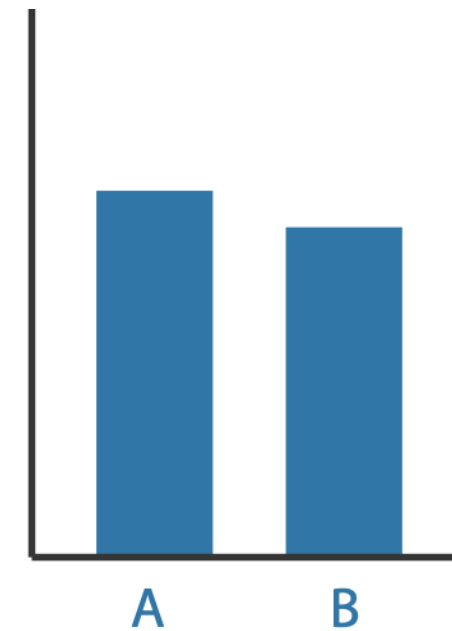
- perceptual system mostly operates with relative judgements, not absolute
 - that's why accuracy increases with common frame/scale and alignment
 - Weber's Law: ratio of increment to background is constant
 - filled rectangles differ in length by 1:9, difficult judgement
 - white rectangles differ in length by 1:2, easy judgement



length



position along
unaligned
common scale



position along
aligned scale

All Spending

Types of Spending

Changes

Department Totals

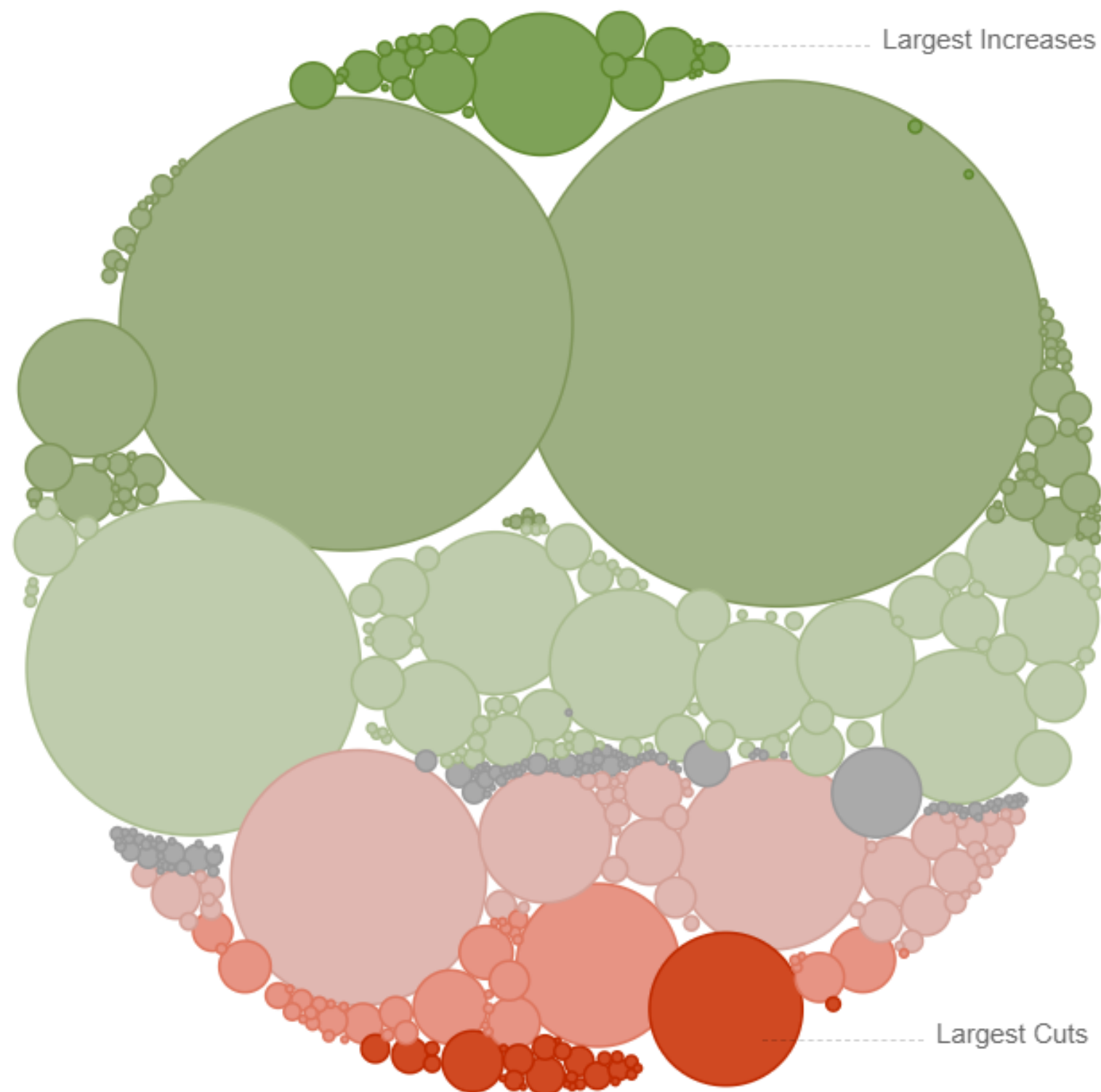
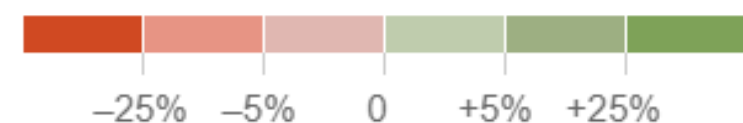
How \$3.7 Trillion Is Spent

Mr. Obama's budget proposal includes \$3.7 trillion in spending in 2013, and forecasts a \$901 billion deficit.

Circles are sized according to the proposed spending.



Color shows amount of cut or increase from 2012.



Channels: Rankings

➔ **Magnitude** Channels: **Ordered** Attributes

Position on common scale 

Position on unaligned scale 

Length (1D size) 

Tilt/angle 

Area (2D size) 

Depth (3D position) 

Color luminance 

Color saturation 

Curvature 

Volume (3D size) 

Same
Same

➔ **Identity** Channels: **Categorical** Attributes

Spatial region 

Color hue 

Motion 

Shape 

➔ **Attribute Types**

➔ Categorical

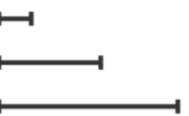


➔ Ordered

➔ Ordinal



➔ Quantitative



- expressiveness
 - match channel and data characteristics
 - magnitude for ordered
 - how much? which rank?
 - identity for categorical
 - what?

Channels: Rankings

➔ Magnitude Channels: Ordered Attributes

Position on common scale



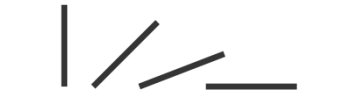
Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



Same

Same

➔ Identity Channels: Categorical Attributes

Spatial region



Color hue



Motion



Shape

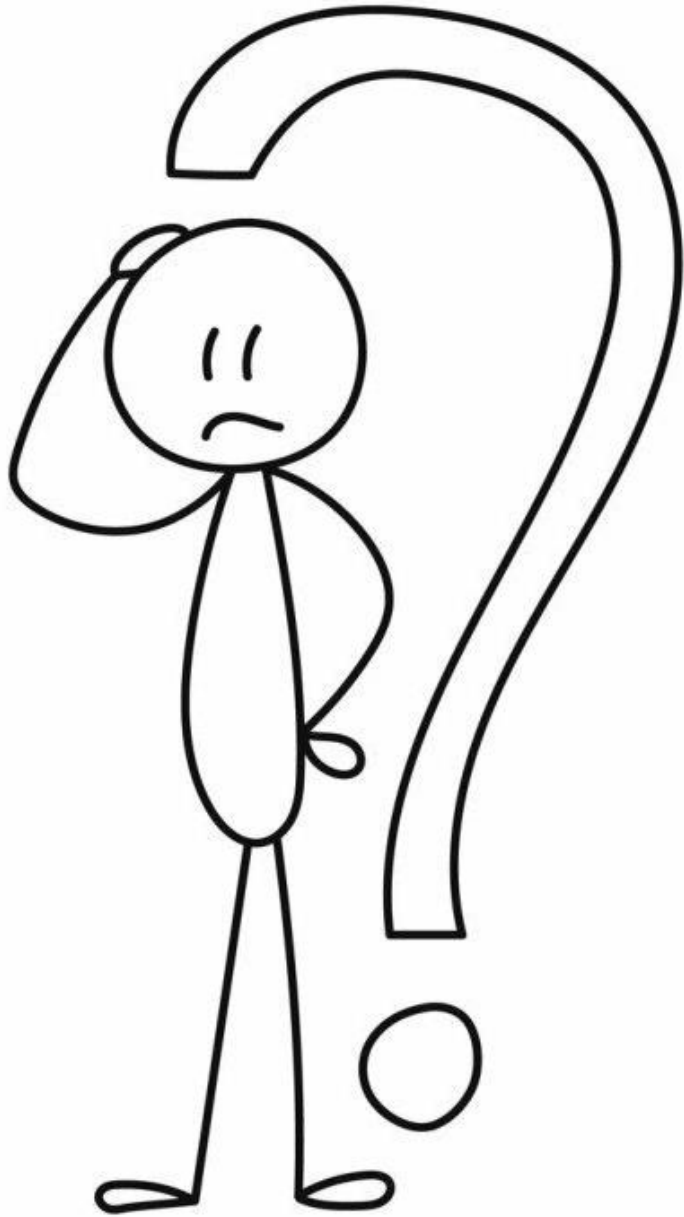


Best

Effectiveness

Least

- expressiveness
 - match channel and data characteristics
- effectiveness
 - channels differ in accuracy of perception
 - spatial position ranks high for both



Prioritize choosing the
most appropriate
channel for each
attribute

How do I pick *which* marks
or channels to use?

Channel Selection

Before we can select channels we must understand how its various properties influence their use.

We must consider both

- **Expressiveness:** match the channel type to data characteristics: the visual encoding should express all of, and only, the information in the dataset attributes.
- **Effectiveness:** for a given task, some channels are better than others so it is important to select the most effective channel for the data

Channel Characteristics

- Discriminability: how many unique steps can we perceive?
- Separability: is our ability to use this channel affected by another one?
- Popout: can things jump out using this channel?
- Grouping: can a channel show perceptual grouping of items?
- Accuracy: how precisely can we tell the difference between encoded items?

All Spending

Types of Spending

Changes

Department Totals

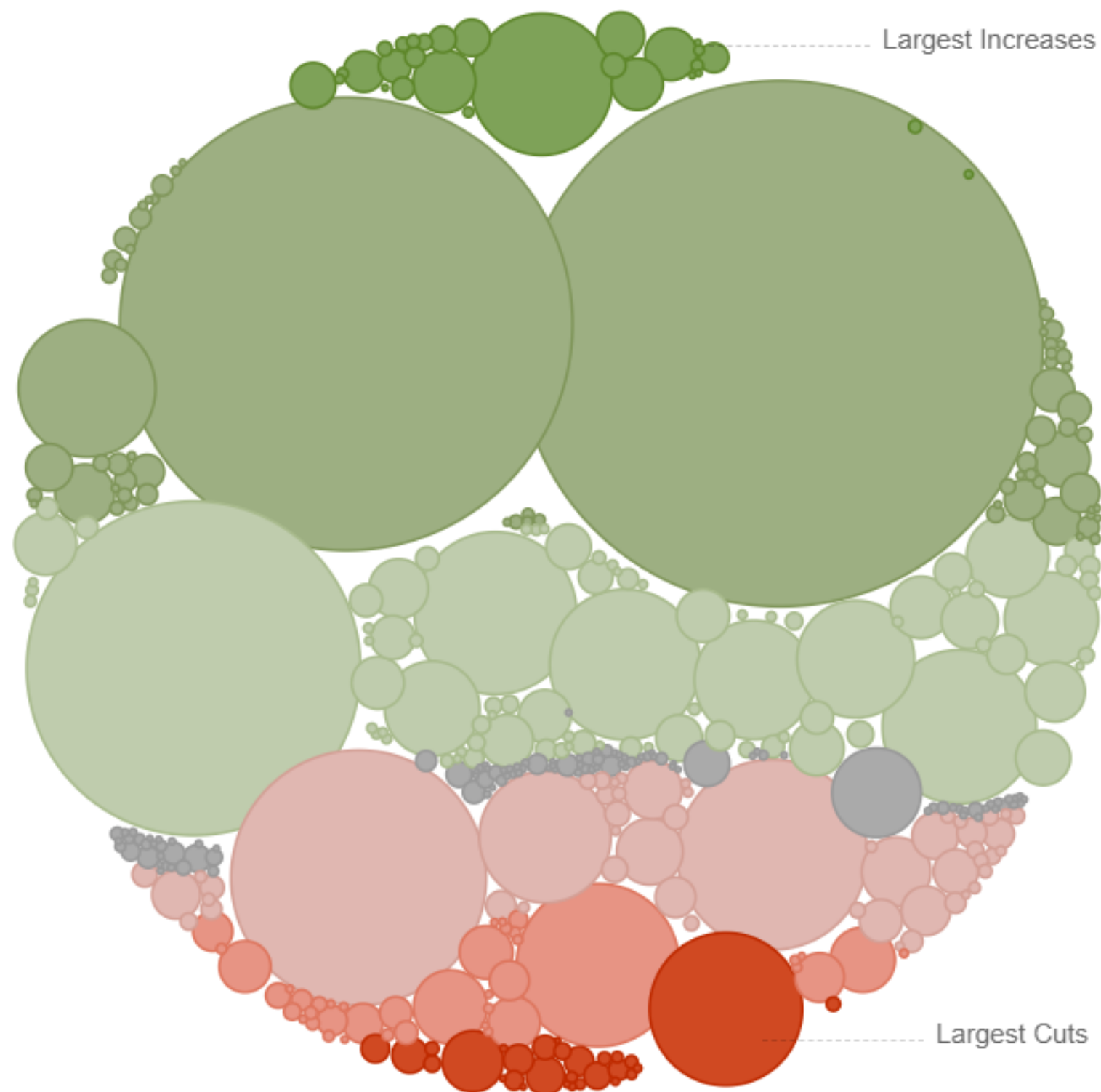
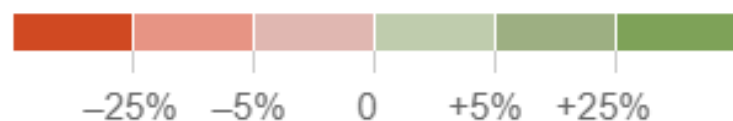
How \$3.7 Trillion Is Spent

Mr. Obama's budget proposal includes \$3.7 trillion in spending in 2013, and forecasts a \$901 billion deficit.

Circles are sized according to the proposed spending.



Color shows amount of cut or increase from 2012.



Lab

- Tomorrow is your first lab
- Make sure that the Setup is complete
- Make sure that you have done T1 – T3 prior to the lab
- The lab will be released at 11am.
- You must complete, get your tests to pass and then submit on Gradescope.
- TAs are there to support you, not to give you the answers
- During each lab add to your crib sheet, this is a list of common altair and pandas methods that you use, (will be helpful during the exam)
- Each lab is intended to be done individually, please do not share your work with anyone.
- There are 6 required labs, the other are just drop in office hours.

Next on Viz

- Lab due in lab.
- Theory Quiz: There will be a short quiz during lecture next week to assess your understanding of everything you have been exposed to thus far. This includes theory, design and tooling.
- If you haven't already, please fill out the doodle poll so we can finalize office hours for the TAs.