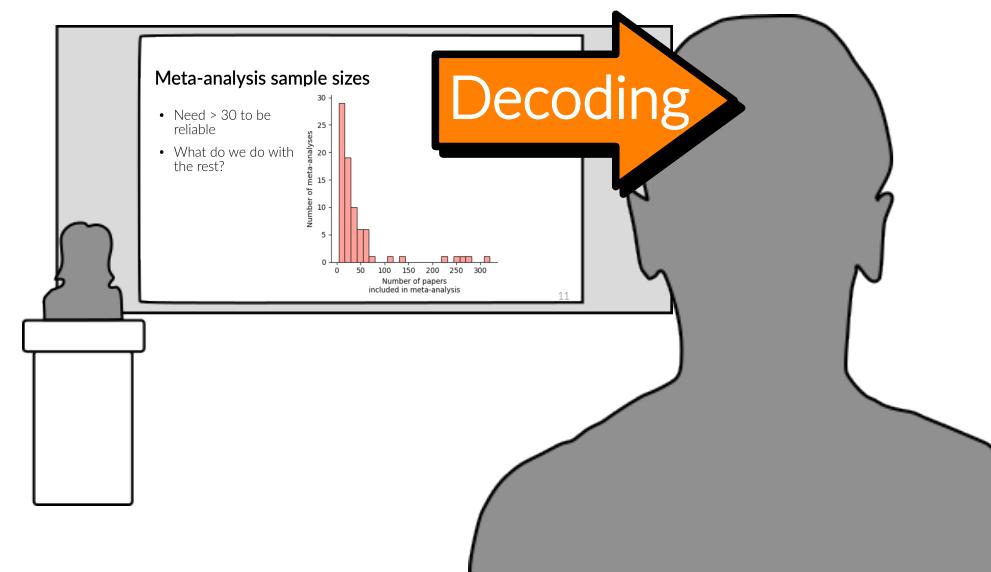


Introduction to Data Visualization

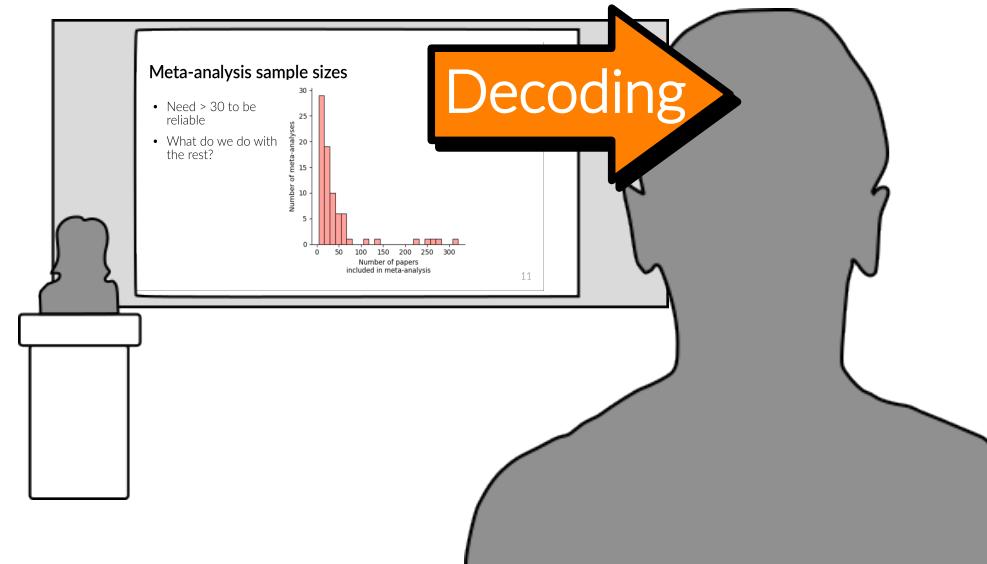
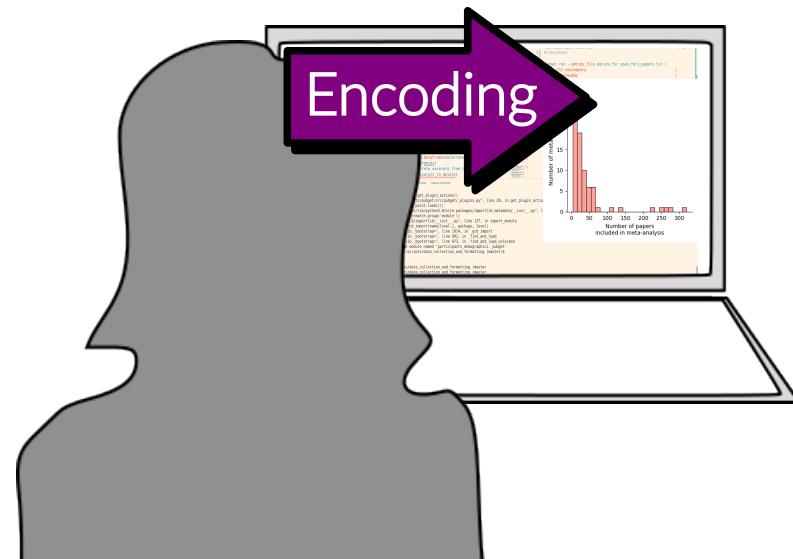
Part 1: Decoding

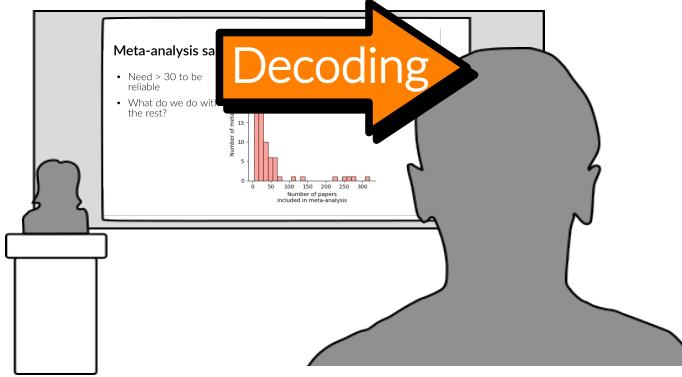
Kendra Oudyk



Goal

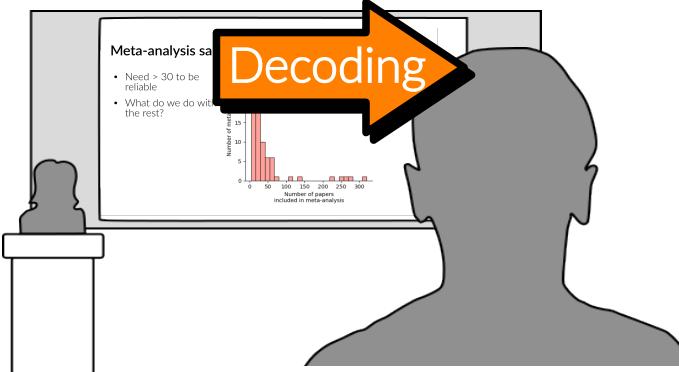
Use principles of visual **encoding** and **decoding**
to **efficiently** create visualizations
that are **effective** and **reproducible**





To plan an **effective visualization**, we need to think about

- **Message**
 - **What** we want to communicate
- **Perception**
 - **How best** to communicate it
- **Conventions**
 - **How** it's usually communicated
- **Context**
 - **Where** it will be seen

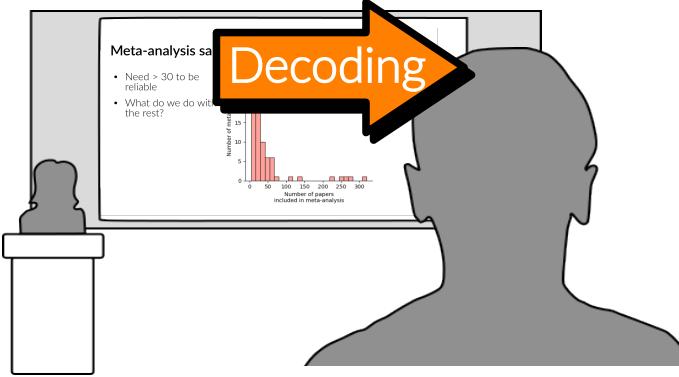


To plan an **effective visualization**, we need to think about

- **Message**
 - What we want to communicate
- **Perception**
 - How best to communicate it
- **Conventions**
 - How it's usually communicated
- **Context**
 - Where it will be seen

Message

- Raw data has no message
- Abstract it
 - “There are more males than females in science”
--> A difference between magnitudes
 - “These brain areas activate together”
--> A grouping / pattern

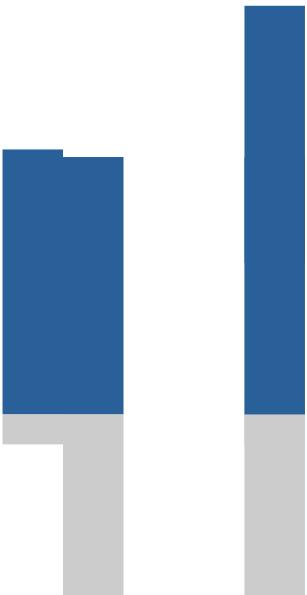


Decoding

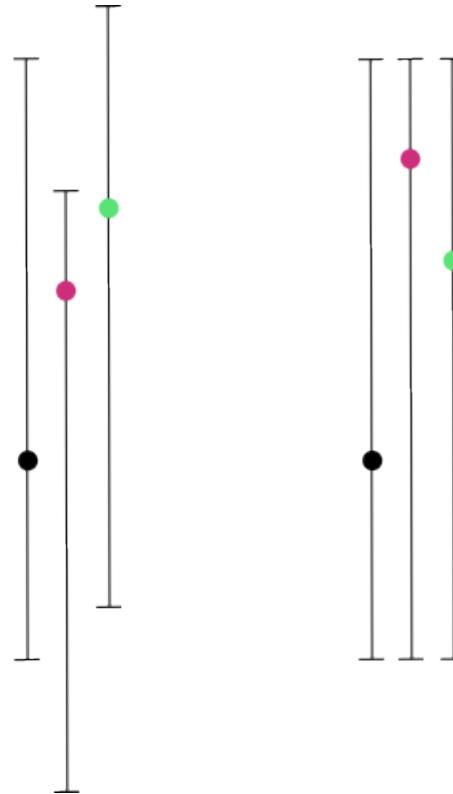
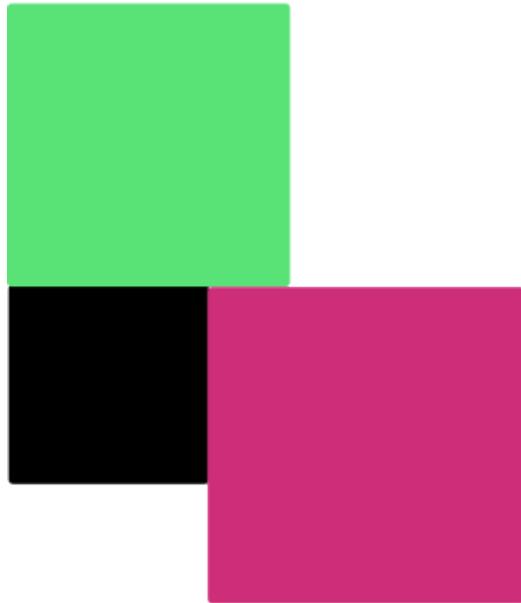
To plan an **effective visualization**, we need to think about

- **Message**
 - What we want to communicate
- **Perception**
 - How best to communicate it
- **Conventions**
 - How it's usually communicated
- **Context**
 - Where it will be seen

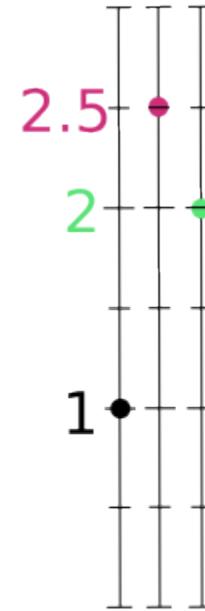
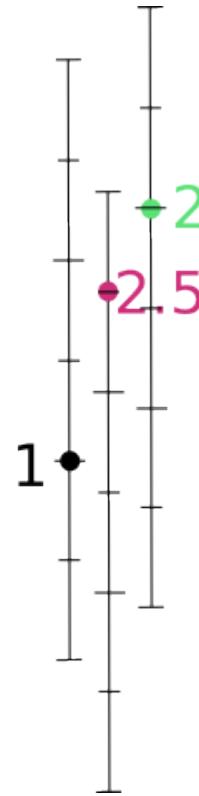
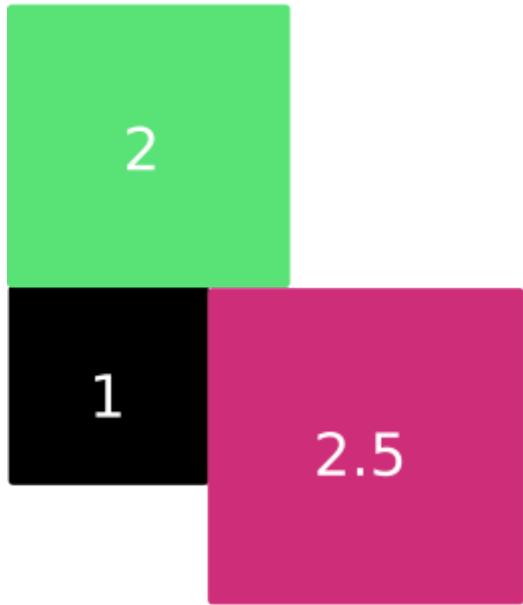
Which blue bar is bigger?



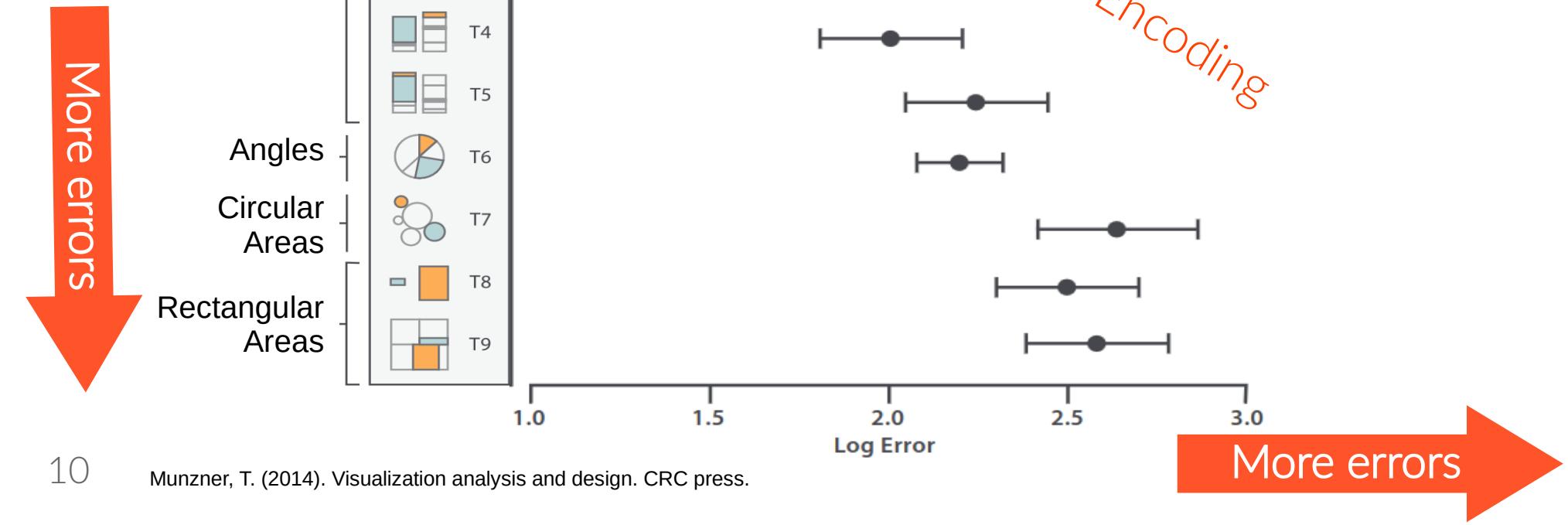
Which is 2x black, green or pink?



Which is 2x black, green or pink?



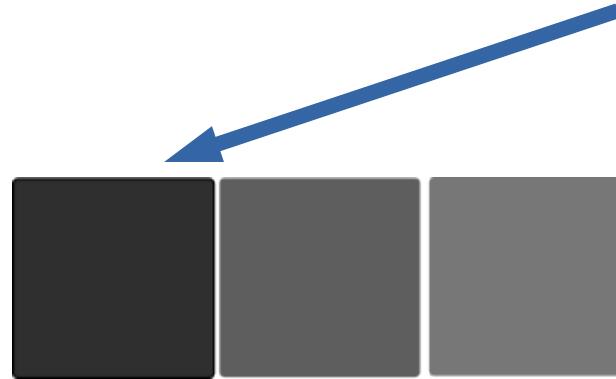
We're better at judging aligned positions



Decoding values vs. patterns

- “There are more males than females in science”
--> A difference between **magnitudes**
- “These brain areas activate together”
--> A grouping / **pattern**

Which is 2x lighter than the leftmost box?

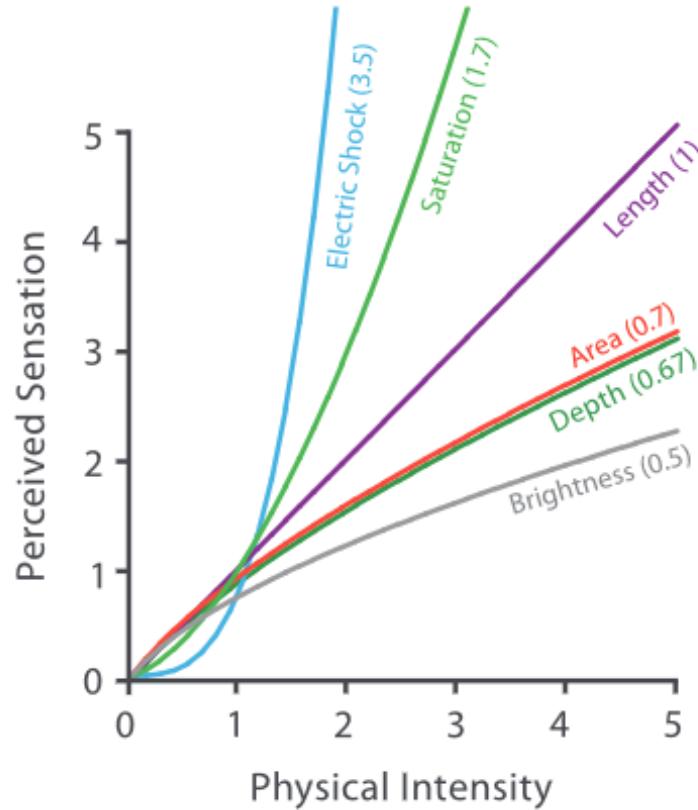


Which is 2x lighter than the leftmost box?



Physical intensity vs perceived intensity

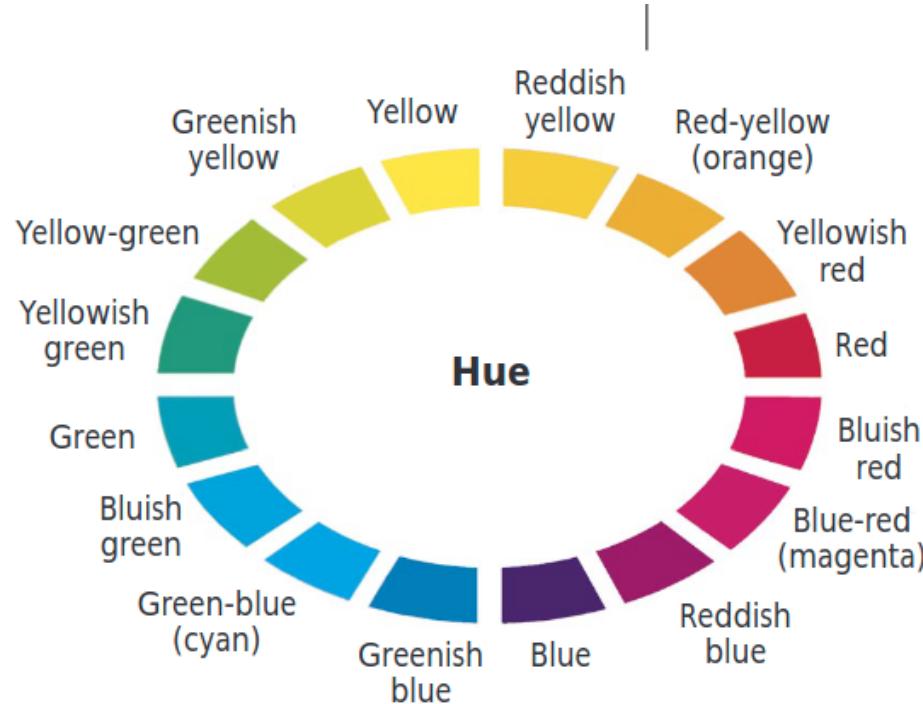
Steven's Psychophysical Power Law: $S = I^N$



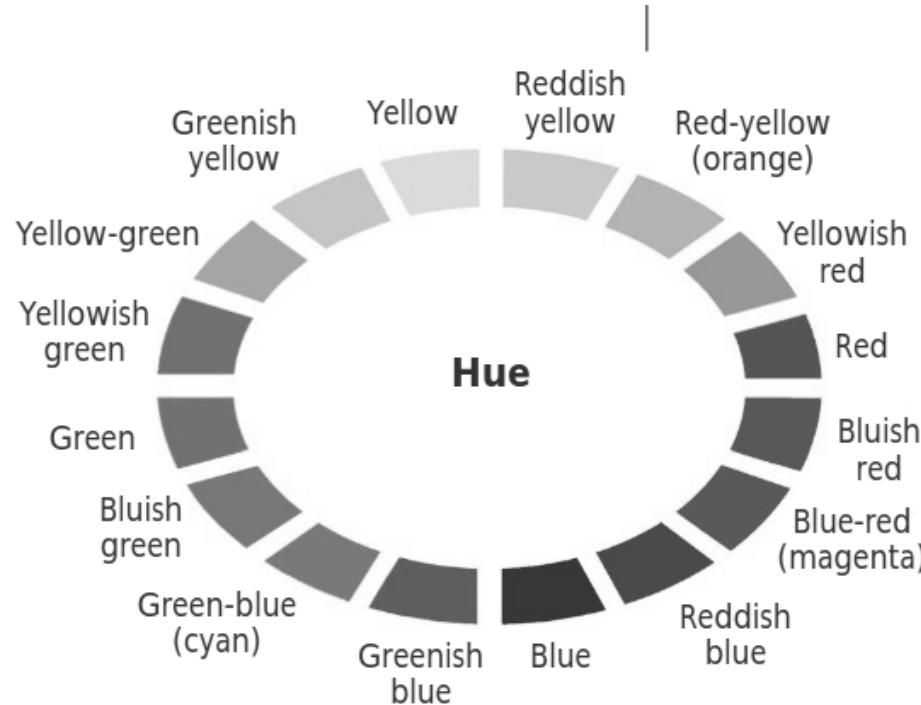
Color: salient but complicated



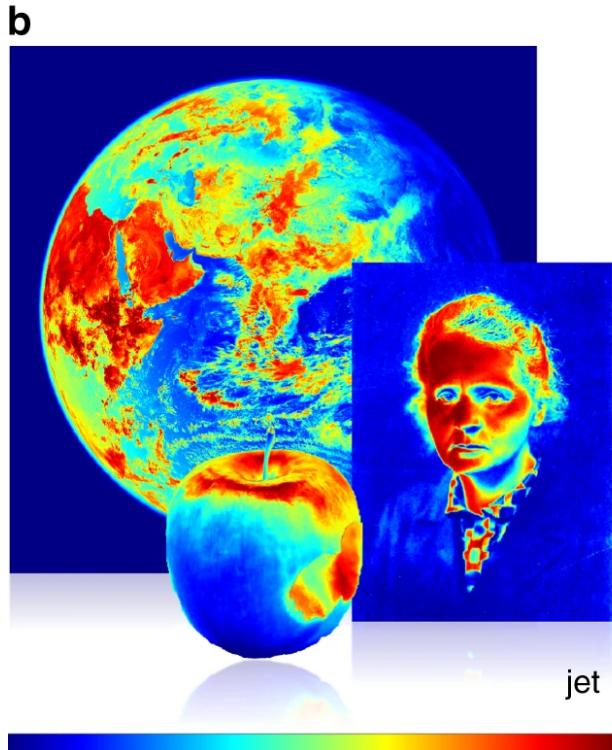
Hue: how we talk about color



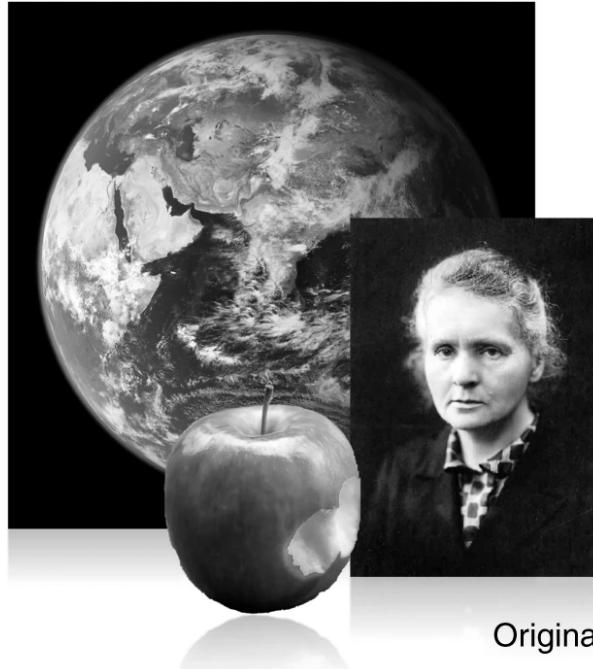
Luminance: an important subconscious cue



Named colors don't work well for ordered values

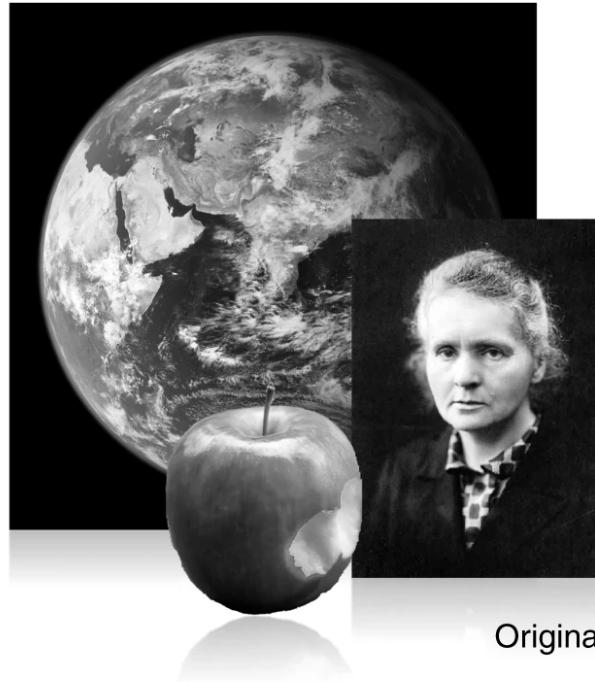
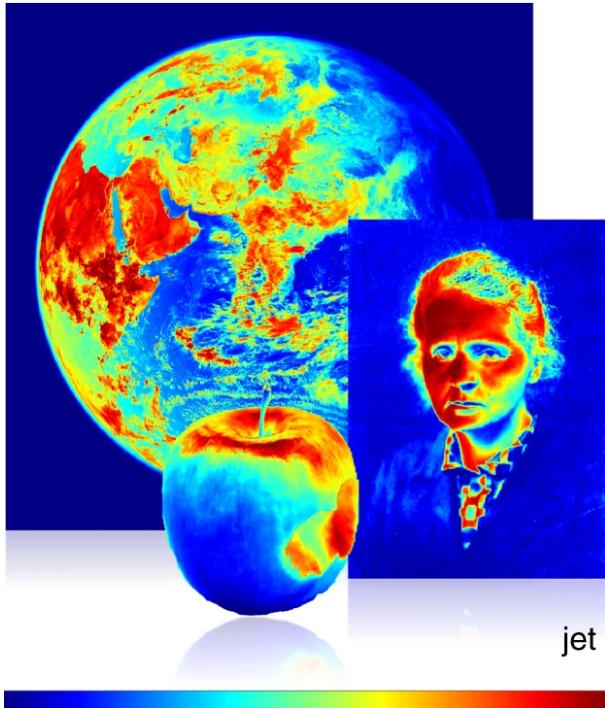


Named colors don't work well for ordered values



Certain colormaps do work

b



c



Certain colormaps do work

b



c



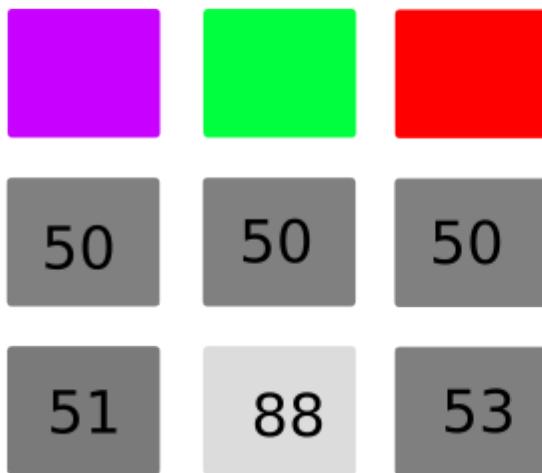
Which is
lightest?



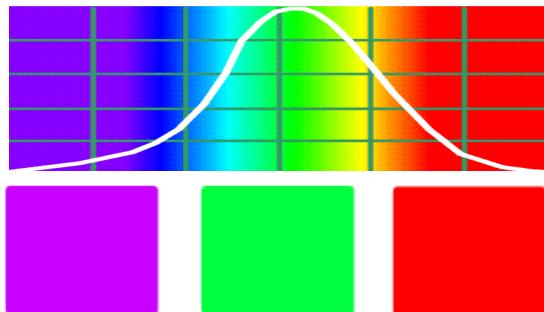


Physical
Lightness





Spectral sensitivity to luminance (lightness)



Physical
Lightness

50	50	50
----	----	----

Perceptual
Lightness

51	88	53
----	----	----

Spectral sensitivity to luminance (lightness)

HSL

(physically accurate)



Physical
Lightness

50

50

50

Perceptual
Lightness

51

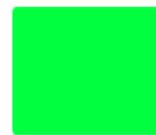
88

53

Spectral sensitivity to luminance (lightness)

HSL

(physically accurate)



Physical
Lightness

50

50

50

49

27

47

Perceptual
Lightness

51

88

53

50

50

50

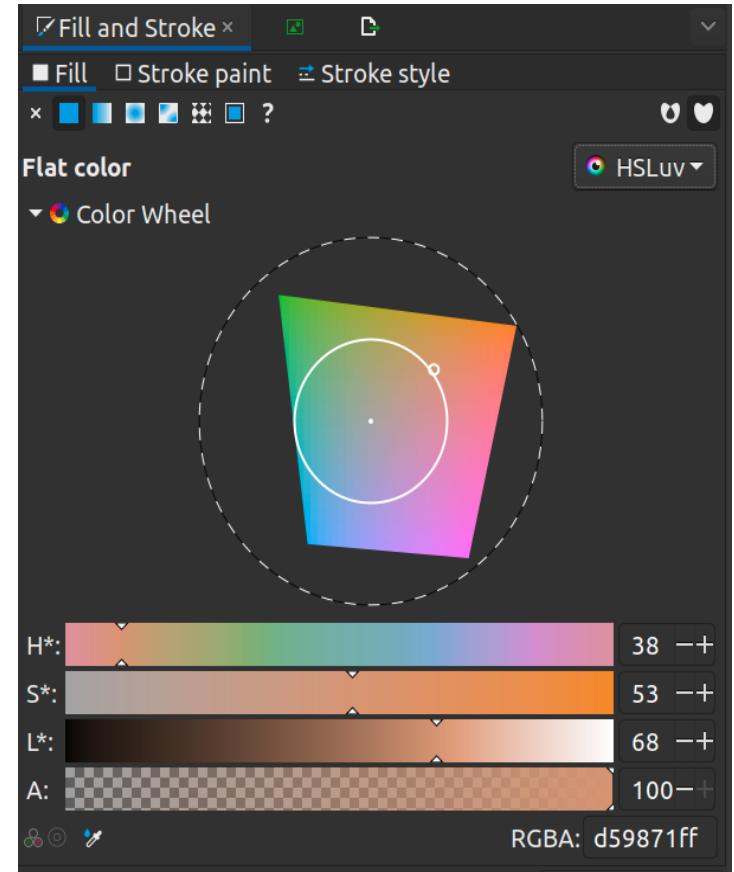
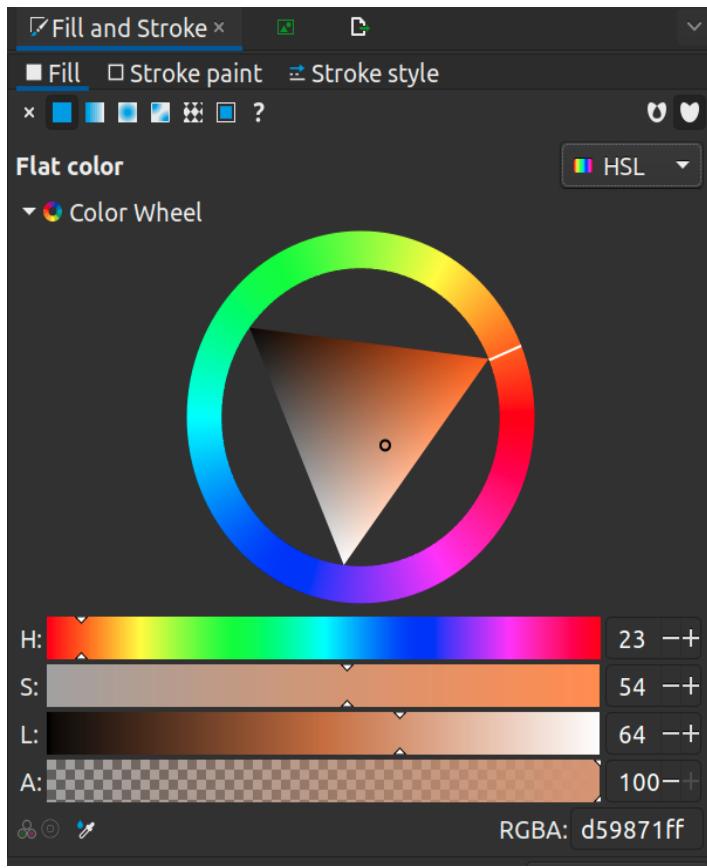
Spectral sensitivity to luminance (lightness)

	HSL (physically accurate)			H*S*L* (psycho physically accurate)		
Physical Lightness	50	50	50	49	27	47
Perceptual Lightness	51	88	53	50	50	50

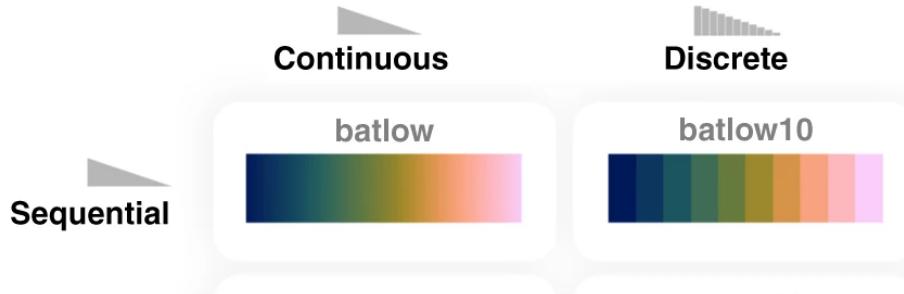


HSL

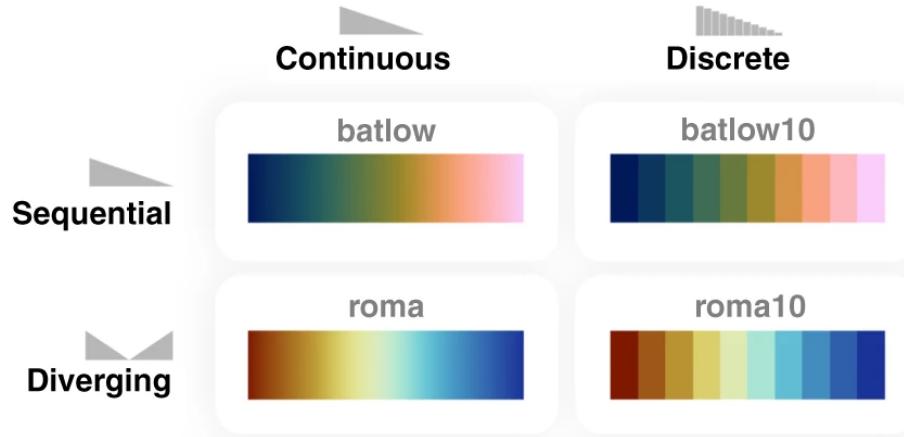
vs H*S*L* (or HSLuv)



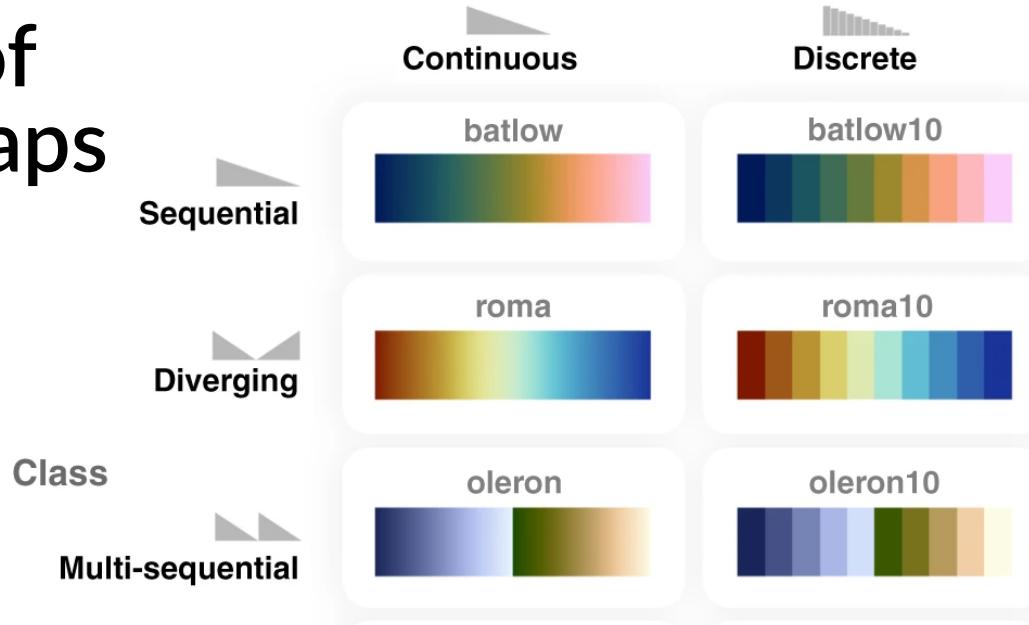
Types of colormaps



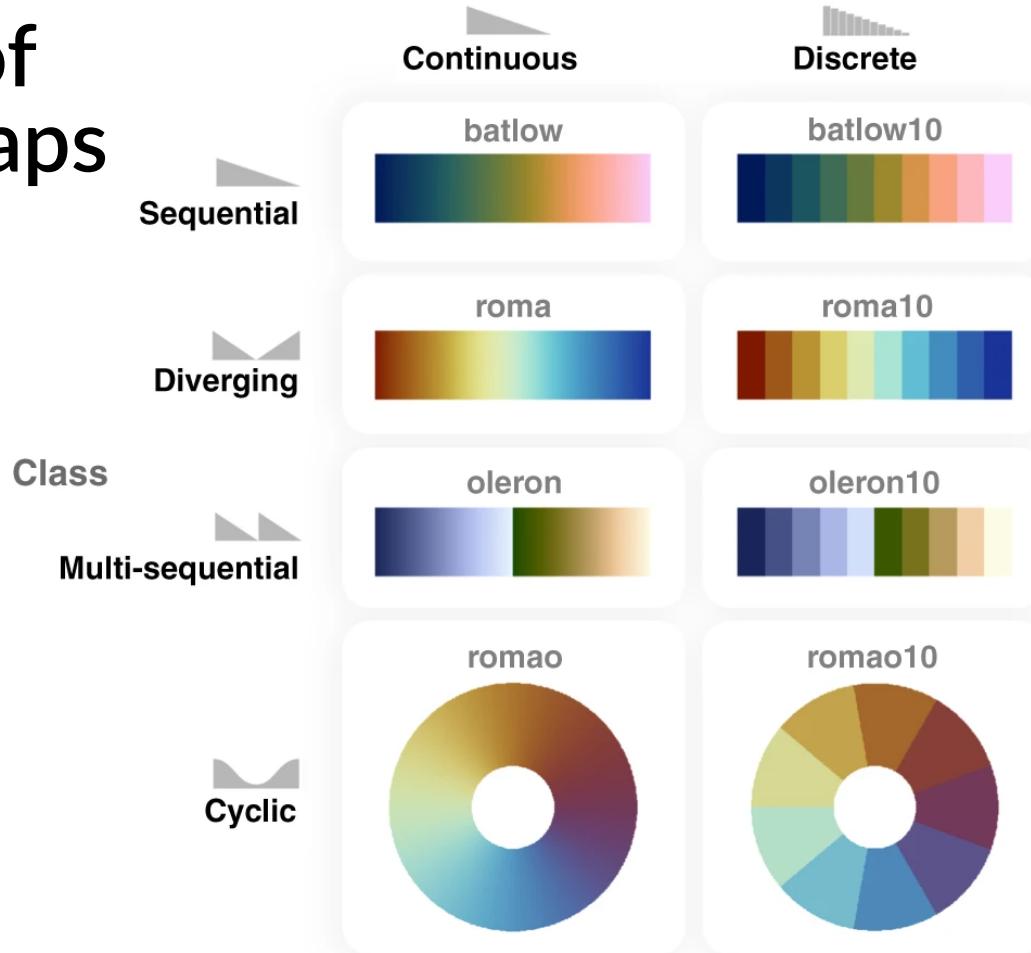
Types of colormaps



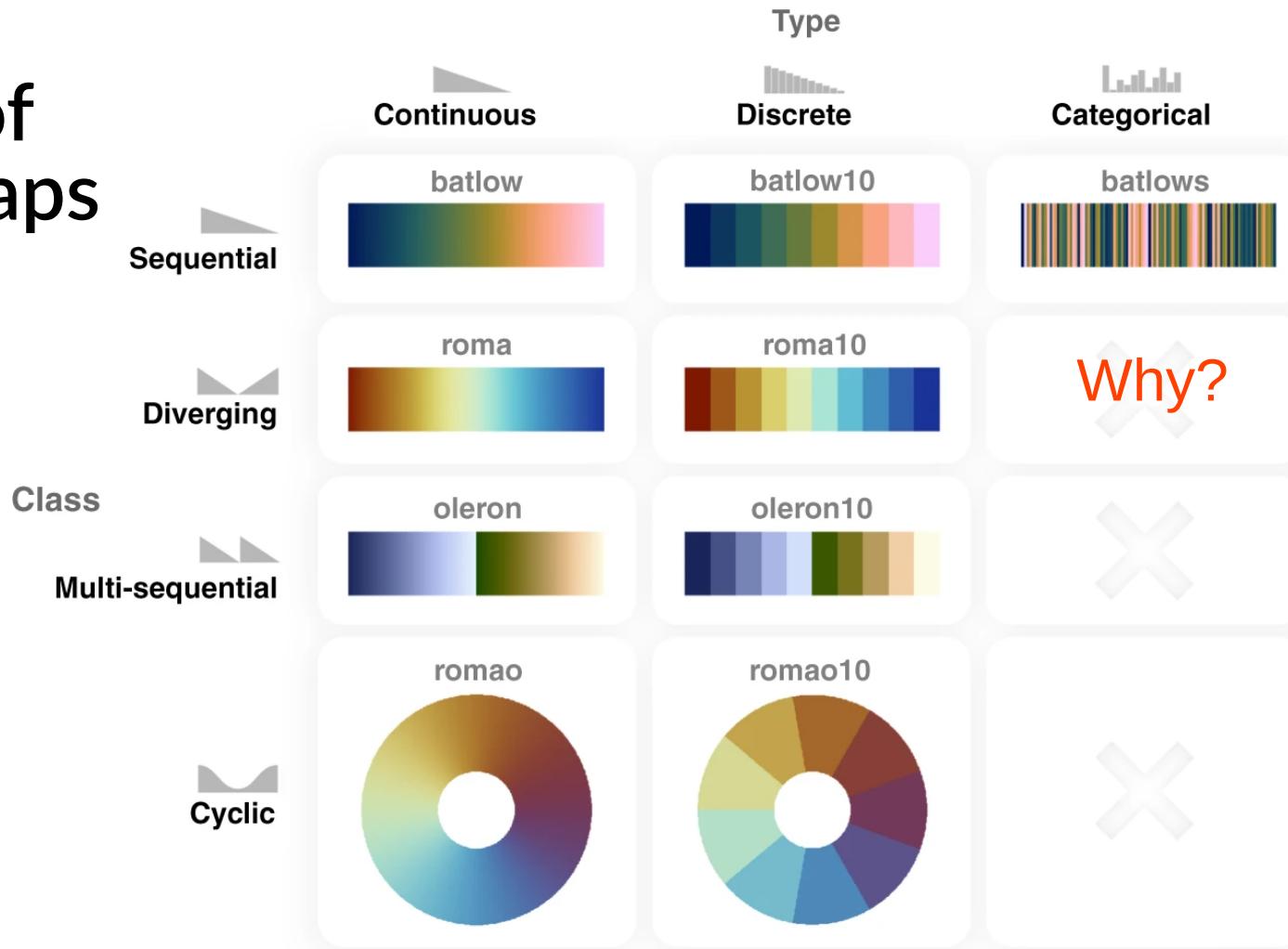
Types of colormaps



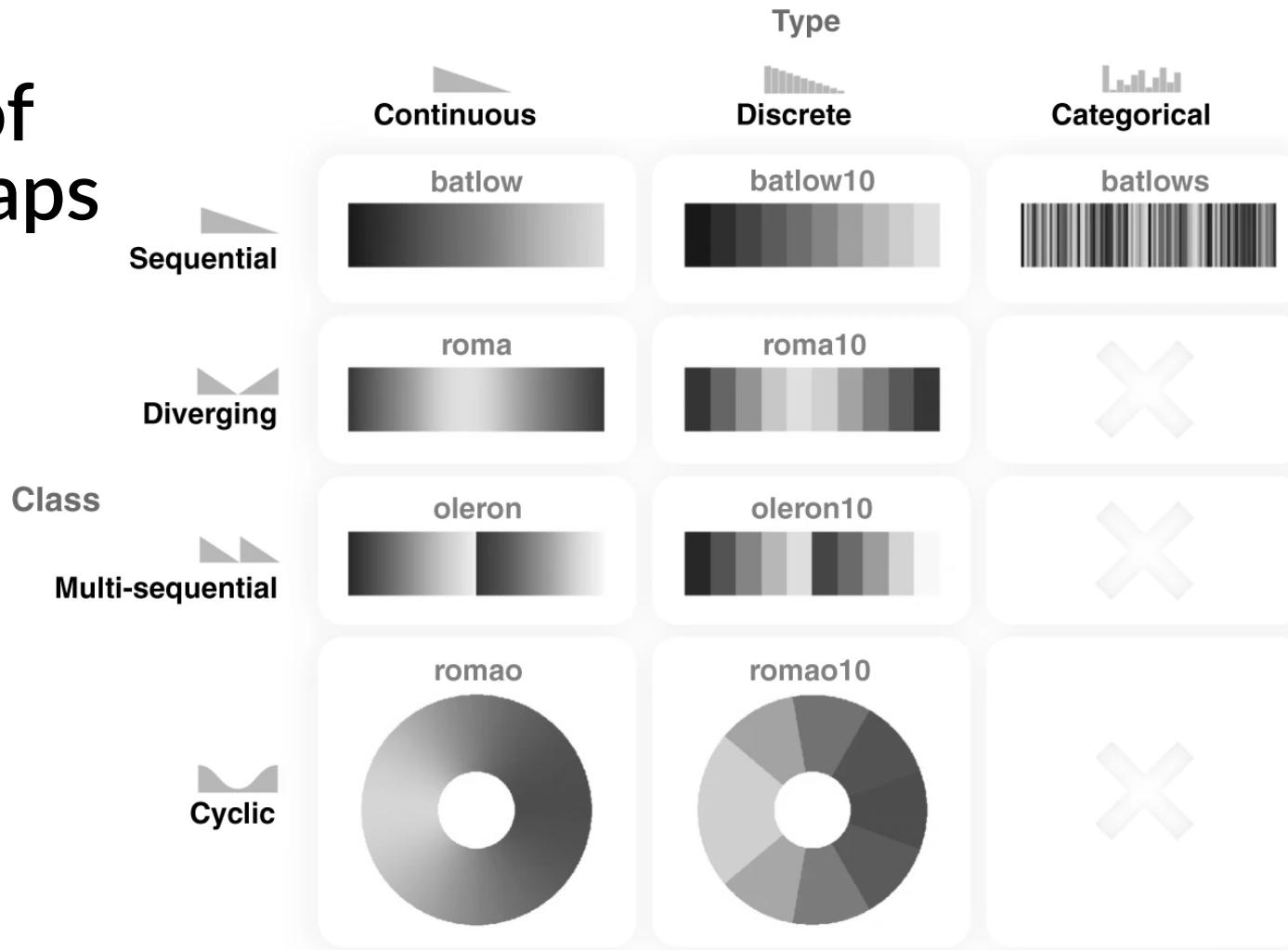
Types of colormaps



Types of colormaps



Types of colormaps



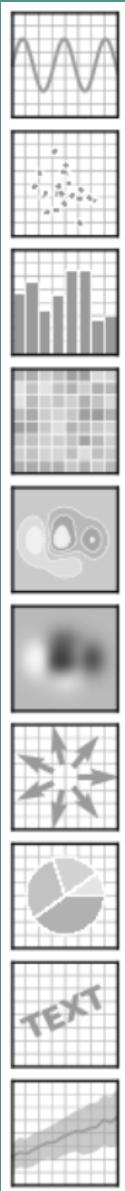
How do I remember all that??

- Test yourself, your lab mate, your friend
“Which bar is higher?”
- Look at the image in grayscale



(For future reference)

Choosing effective charts



Channels: Expressiveness Types and Effectiveness Ranks

⇒ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



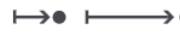
Tilt/angle



Area (2D size)



Depth (3D position)



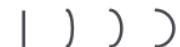
Color luminance



Color saturation



Curvature



Volume (3D size)



⇒ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



Shape



Text

Fr
Ca

Canada France Canada France

https://matplotlib.org/cheatsheets/_images/cheatsheets-1.png

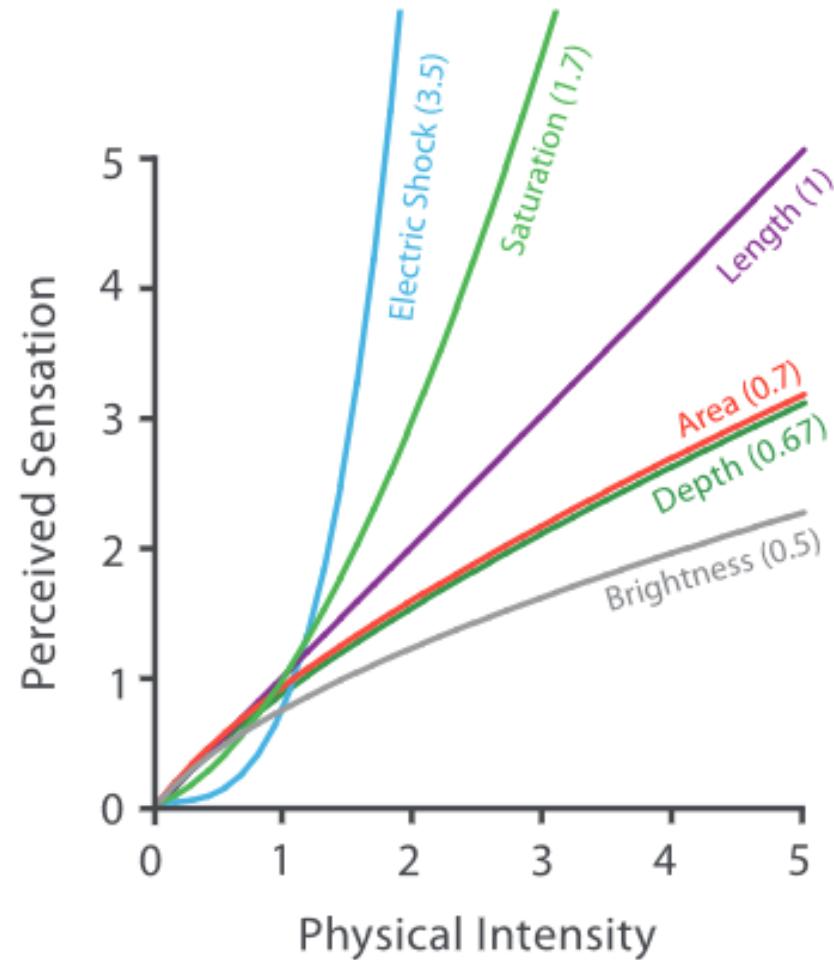
Munzner, T. (2014). Visualization analysis and design. CRC press.

Lineplot

Ordered

Categorical

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



Lineplot

Ordered

Position on a common scale

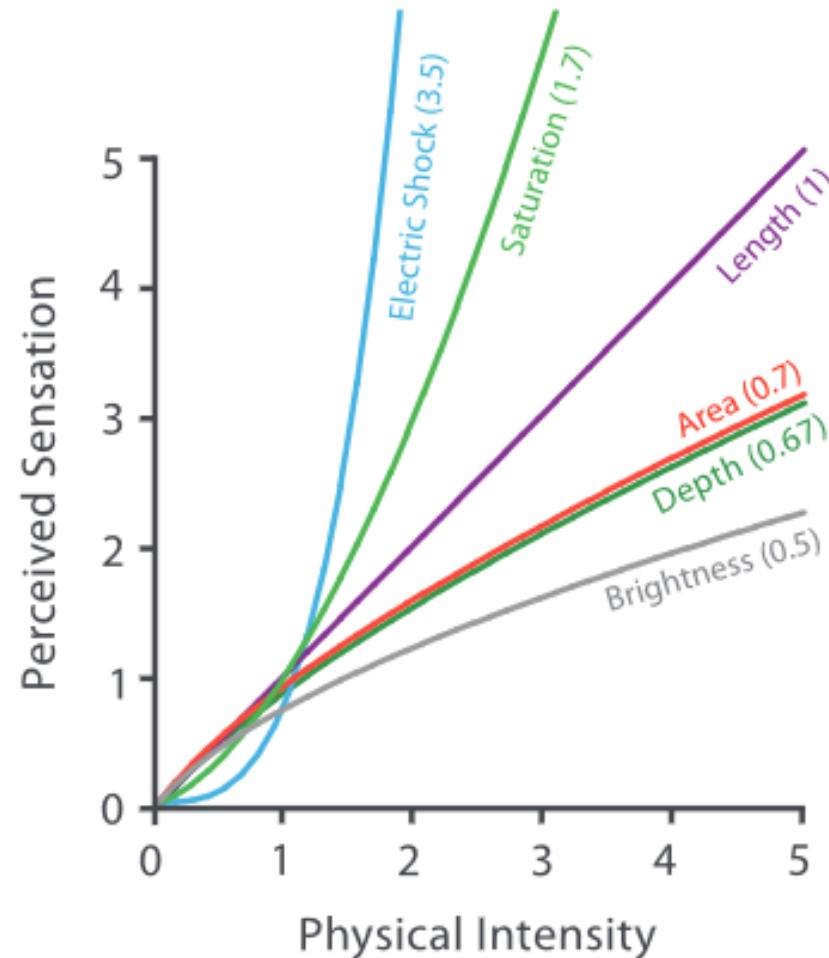


Canada Canada Canada
France France France

Text (equation)

Categorical

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



Lineplot

Ordered

Position on a common scale



Canada Canada Canada
France France France

Text (equation)

Categorical

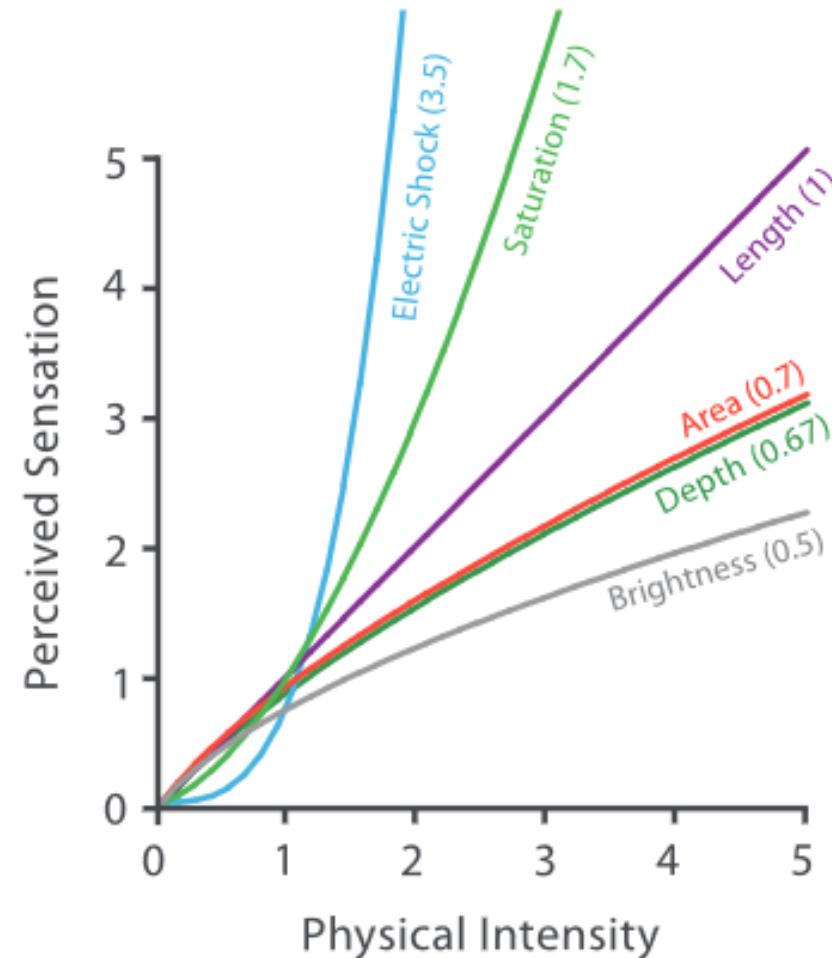
Hue



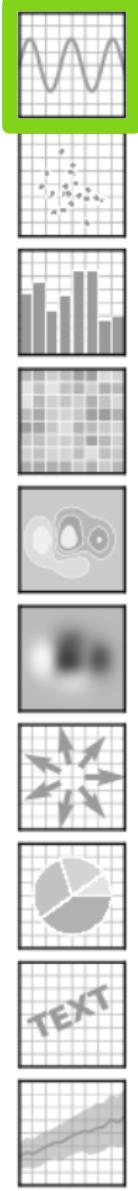
Text (identity)

Canada Canada Canada
France France France

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



Lineplot



Channels: Expressiveness Types and Effectiveness Ranks

⇒ **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)



⇒ **Identity** Channels: **Categorical Attributes**

Spatial region



Color hue



Motion



Shape



Text



Most

Effective

Least

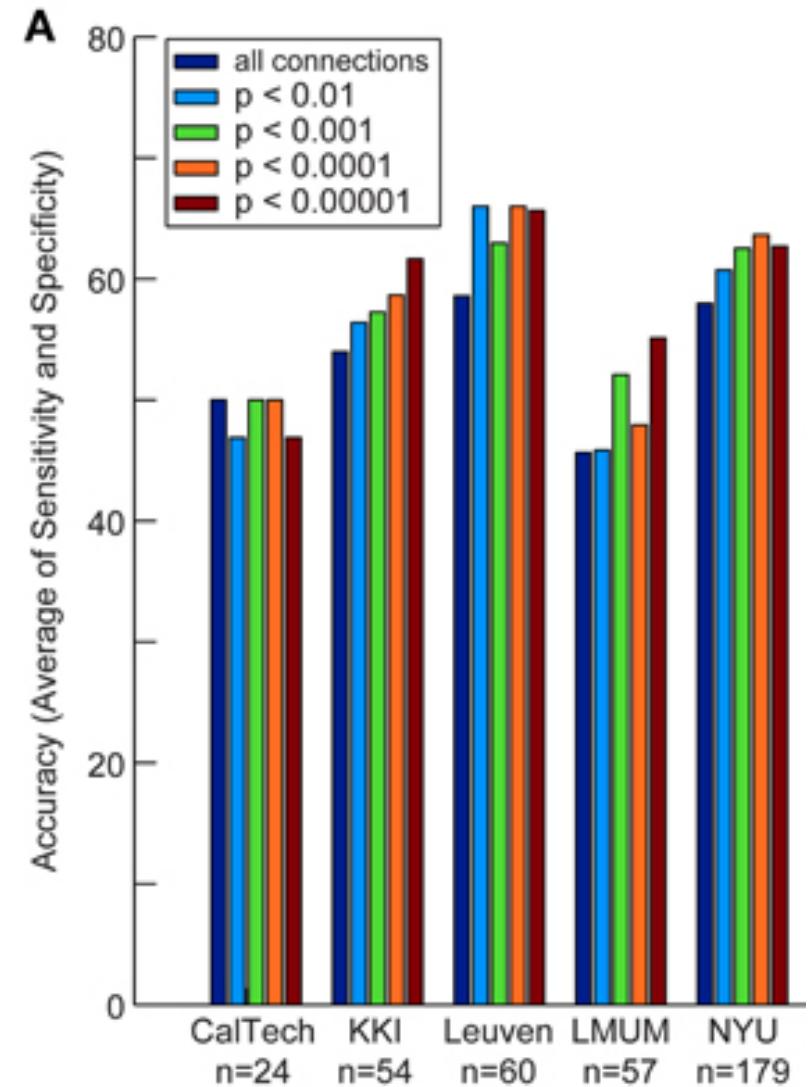
https://matplotlib.org/cheatsheets/_images/cheatsheets-1.png

Munzner, T. (2014). Visualization analysis and design. CRC press.

Barplot

Ordered

Categorical

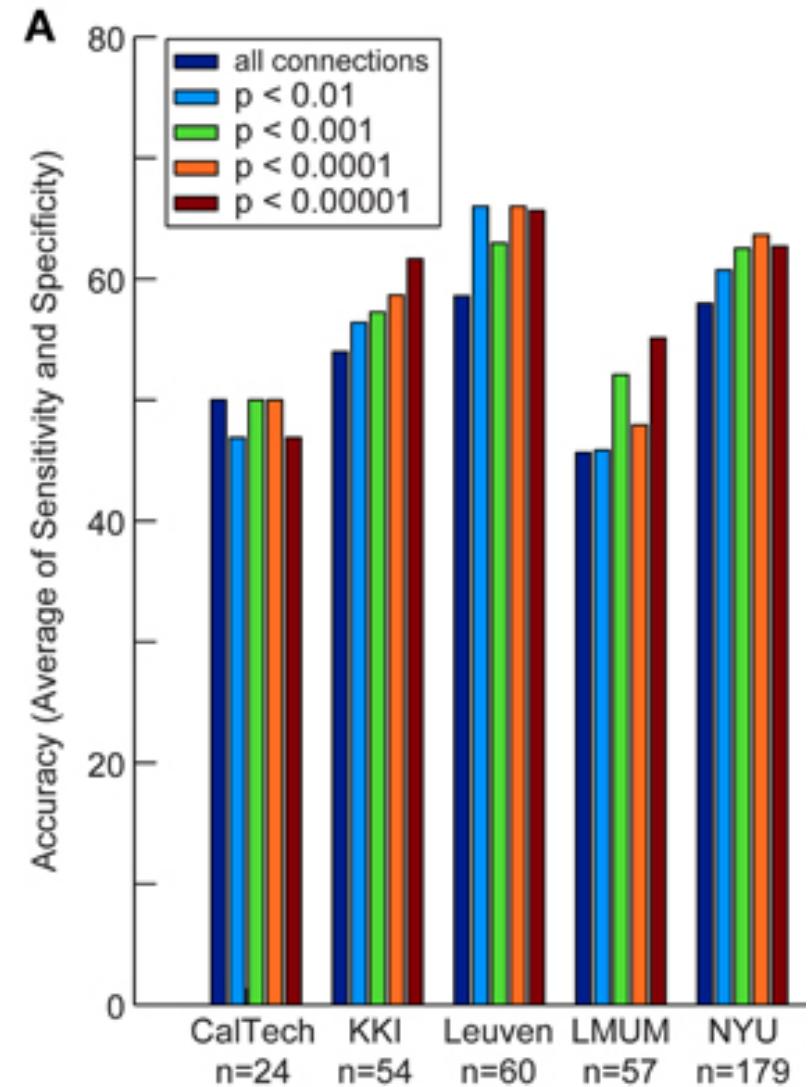


Barplot

Ordered

Position on a
common scale

Categorical



Barplot

Ordered

Position on a
common scale

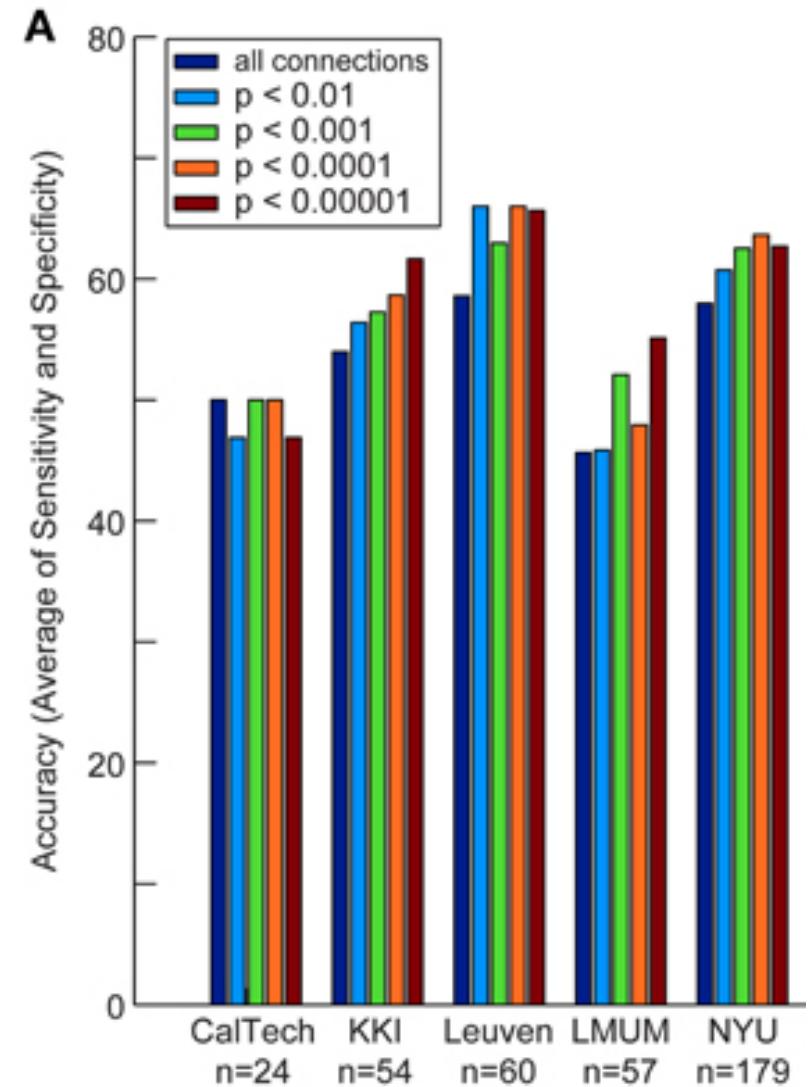


Categorical

Spatial region



Hue



Barplot

Ordered

Position on a
common scale

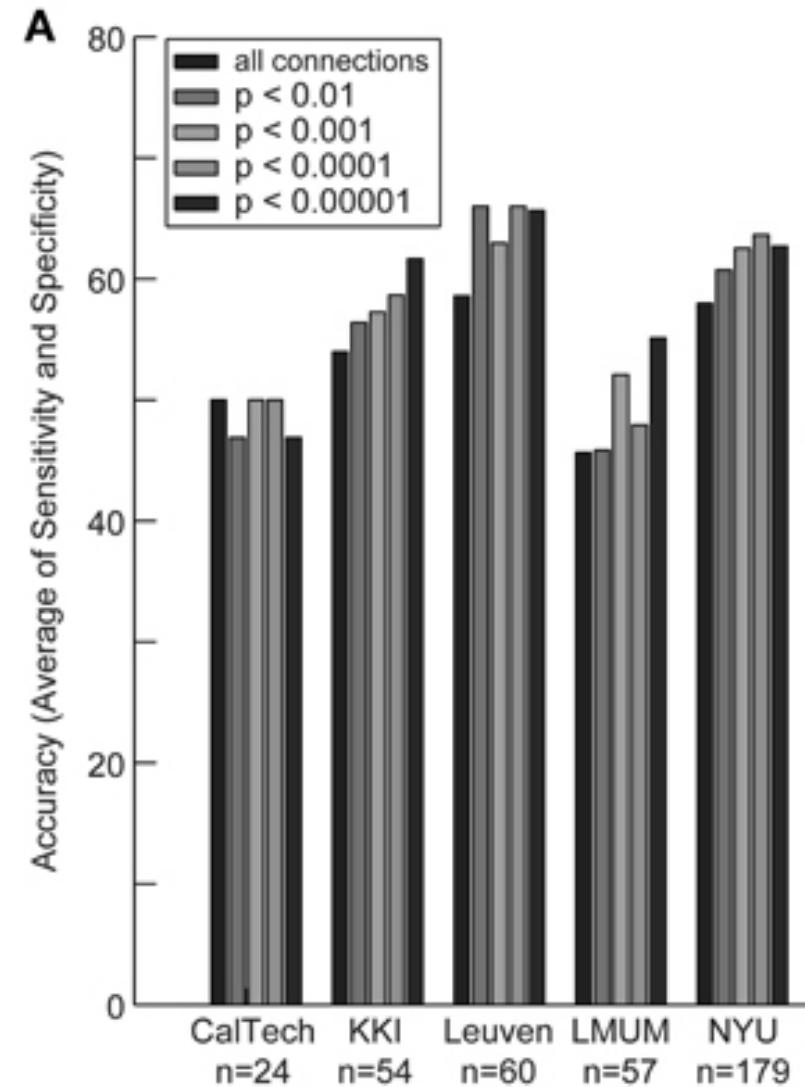


Categorical

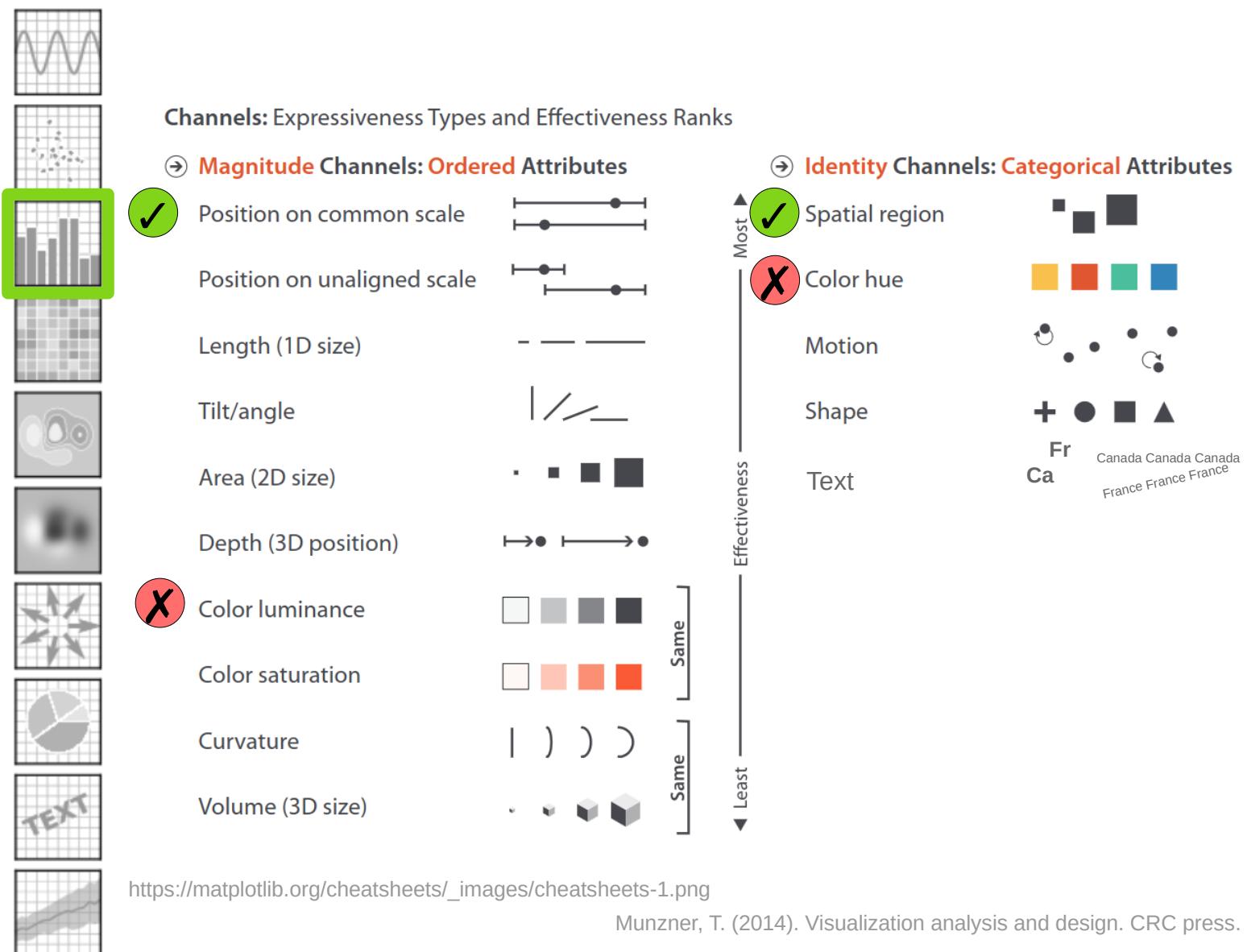
Spatial region



Hue



Barplot



Barplot

Ordered

Position on a common scale



Hue



Luminance

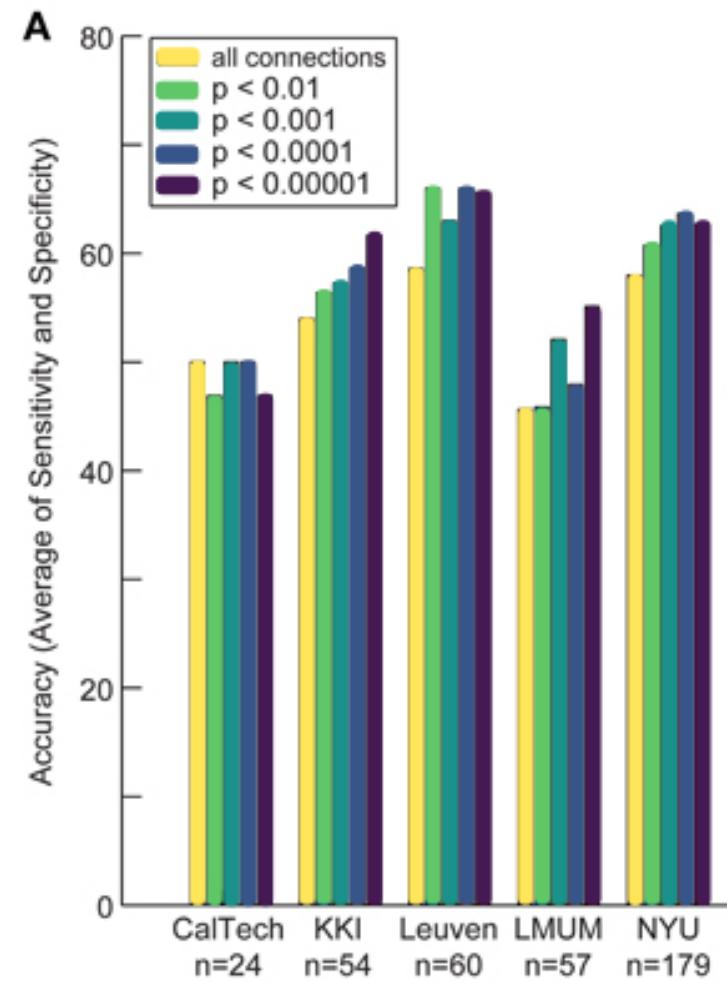


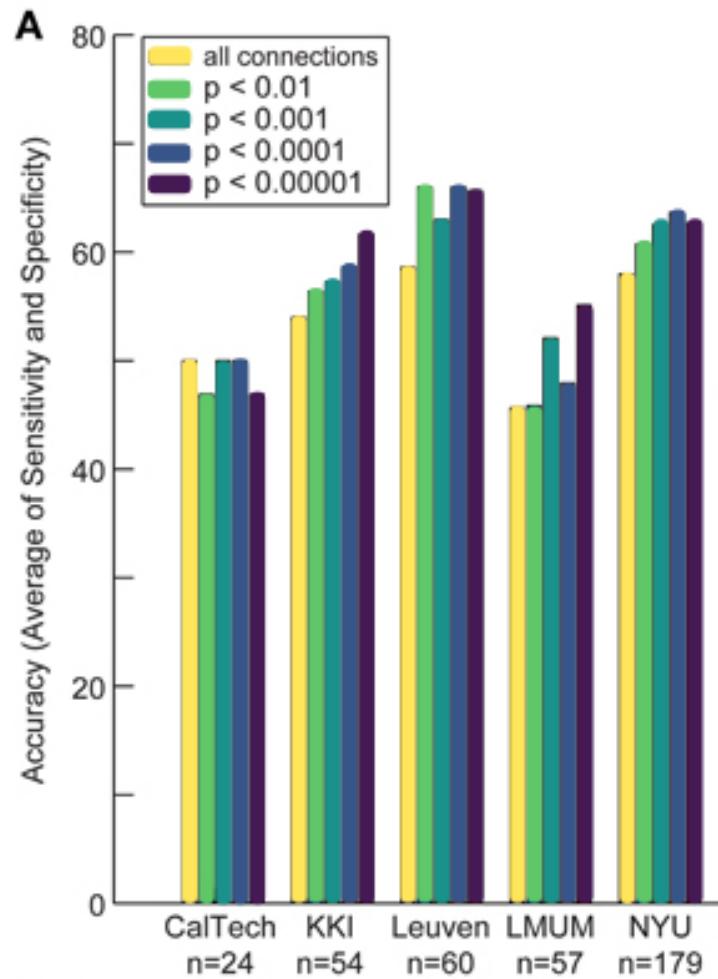
Categorical

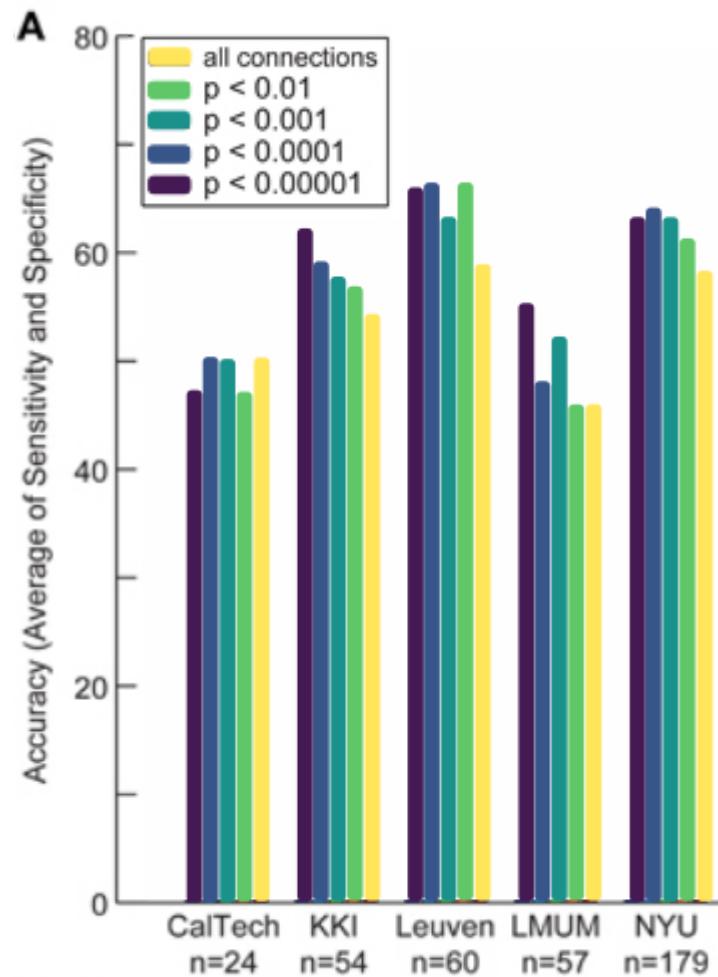
Spatial region

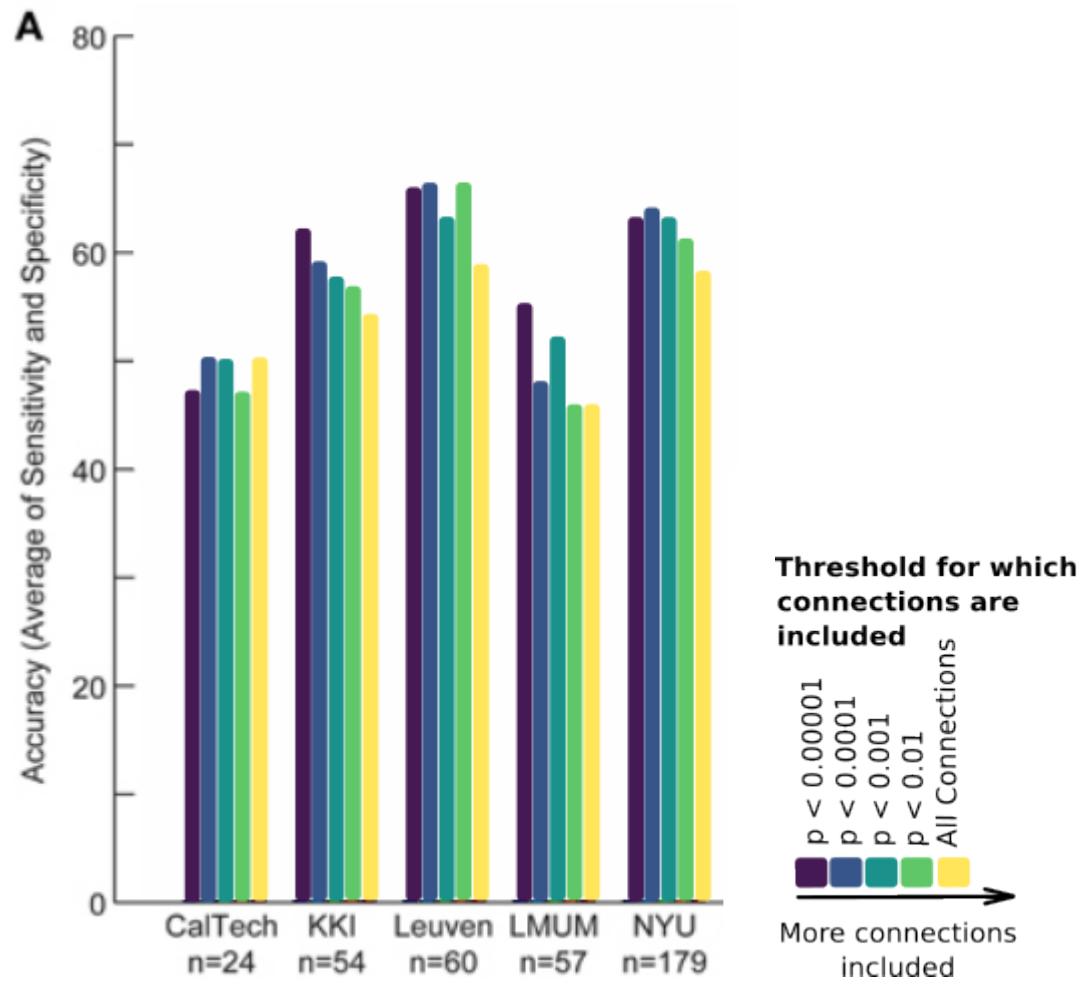


Hue



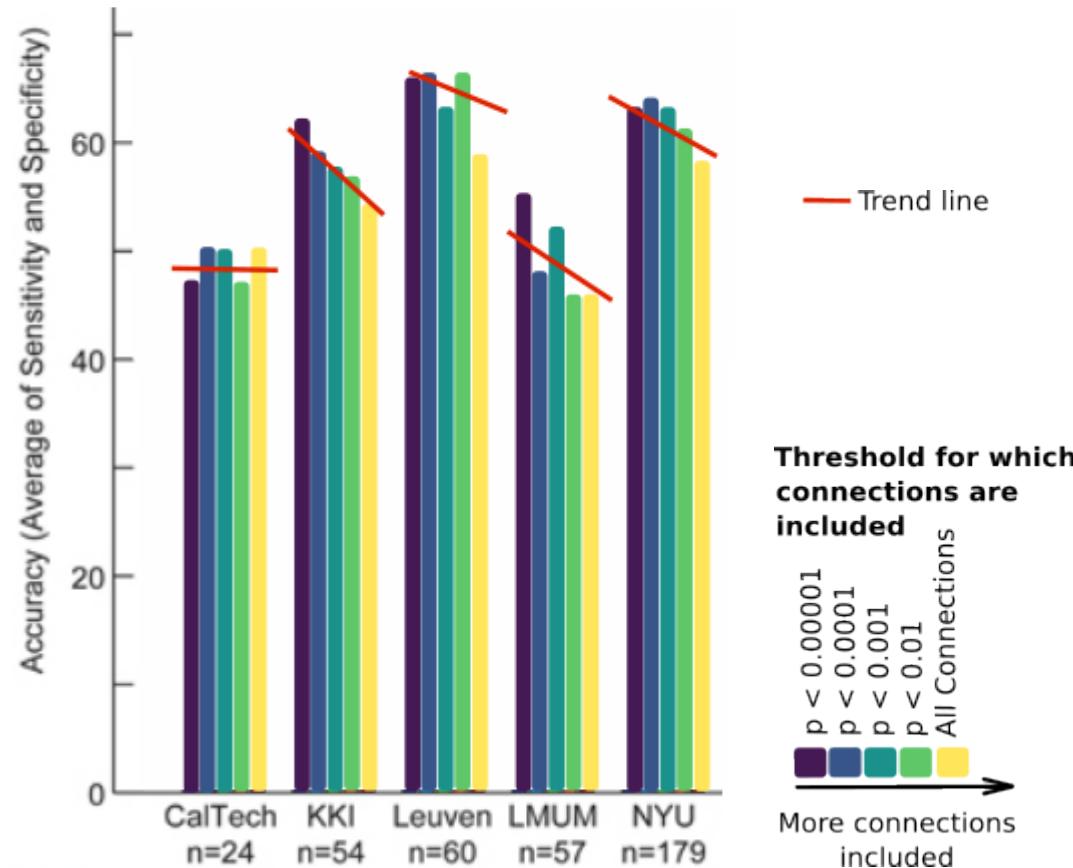






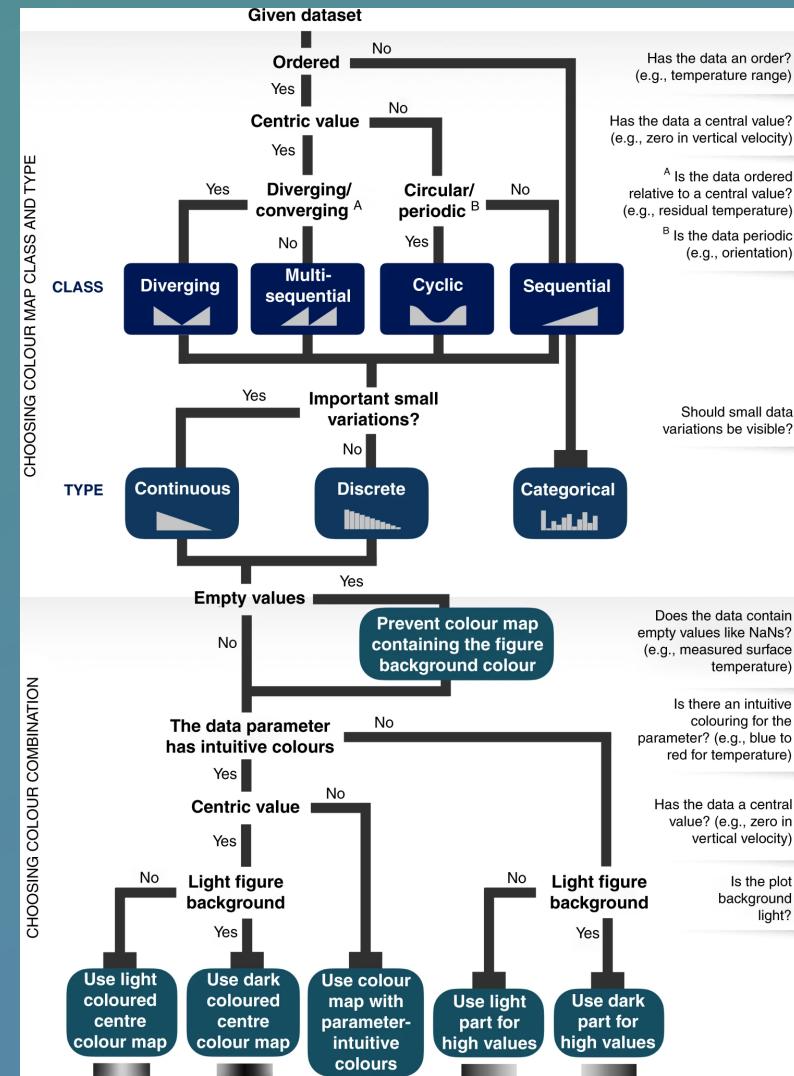
Across locations, the diagnostic accuracy is fairly consistent around 55%.

The diagnostic accuracy tends to decrease as we include more connections in the connectomes.



(For future reference)

Choosing effective colormaps



Activities Firefox Web Browser • Apr 26 7:51 PM • 100 %

Frontiers | rogowitz a IEEE Xplor Data visual IEEE Xplor Fig. 2: Colo Choosing + - X

https://matplotlib.org/stable/tutorials/colors/colormaps.html#sphx-glr-tutorials-colors-colormaps

matplotlib

Plot types Examples Tutorials Reference User guide Develop Releases stable

Section Navigation

- Introductory
- Intermediate
- Advanced
- Colors
 - Specifying colors
 - Customized Colorbars Tutorial
 - Creating Colormaps in Matplotlib
 - Colormap Normalization
- Choosing Colormaps in Matplotlib
 - Text
 - Toolkits
 - Provisional

Sequential

For the Sequential plots, the lightness value increases monotonically through the colormaps. This is good. Some of the L^* values in the colormaps span from 0 to 100 (binary and the other grayscale), and others start around $L^* = 20$. Those that have a smaller range of L^* will accordingly have a smaller perceptual range. Note also that the L^* function varies amongst the colormaps: some are approximately linear in L^* and others are more curved.

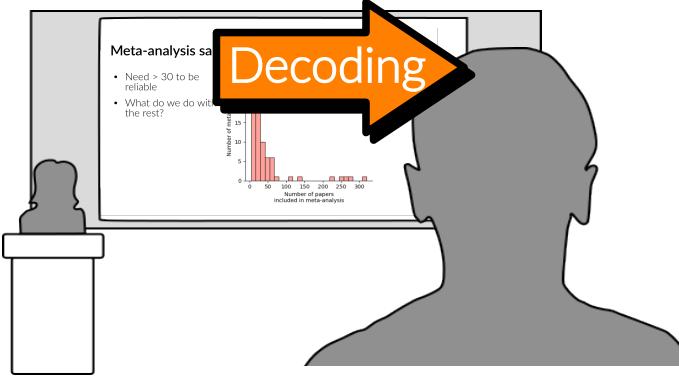
```
plot_color_gradients('Perceptually Uniform Sequential',
                     ['viridis', 'plasma', 'inferno', 'magma', 'cividis'])
```

Perceptually Uniform Sequential colormaps

```
plot_color_gradients('Sequential',
                     ['Greys', 'Purples', 'Blues', 'Greens', 'Oranges', 'Reds',
                      'YlOrRd', 'OrRd', 'PuRd', 'RdPu', 'BuPu', 'GnBu',
                      'PuBu', 'RdGy', 'BuGn', 'GnBu_r', 'PuBu_r', 'RdGy_r',
                      'BuGn_r'])
```

On this page

- Overview
- Classes of colormaps
 - Sequential
 - Sequential2
 - Diverging
 - Cyclic
 - Qualitative
 - Miscellaneous
- Lightness of Matplotlib colormaps
- Grayscale conversion
- Color vision deficiencies
- References

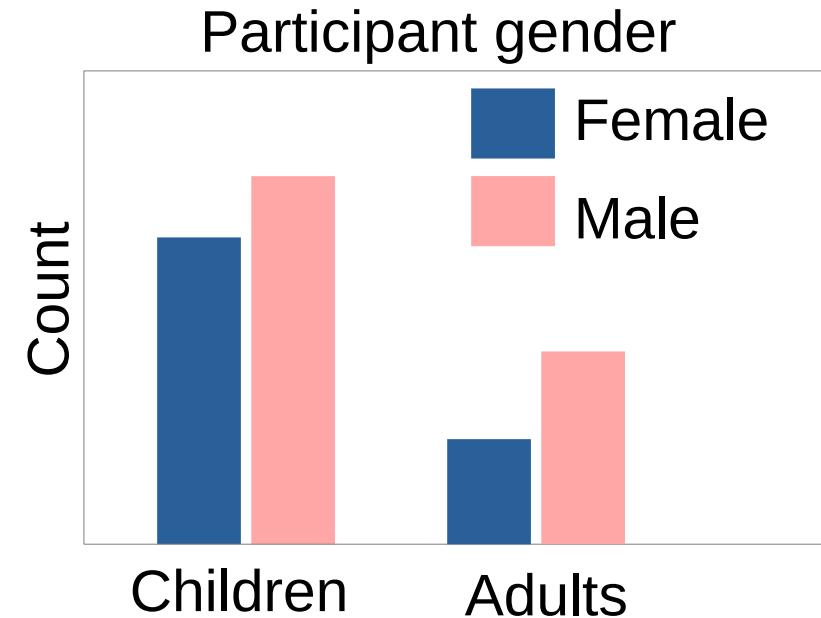


Decoding

To plan an **effective visualization**, we need to think about

- **Message**
 - What we want to communicate
- **Perception**
 - How best to communicate it
- **Conventions**
 - How it's usually communicated
- **Context**
 - Where it will be seen

Culture: Color associations



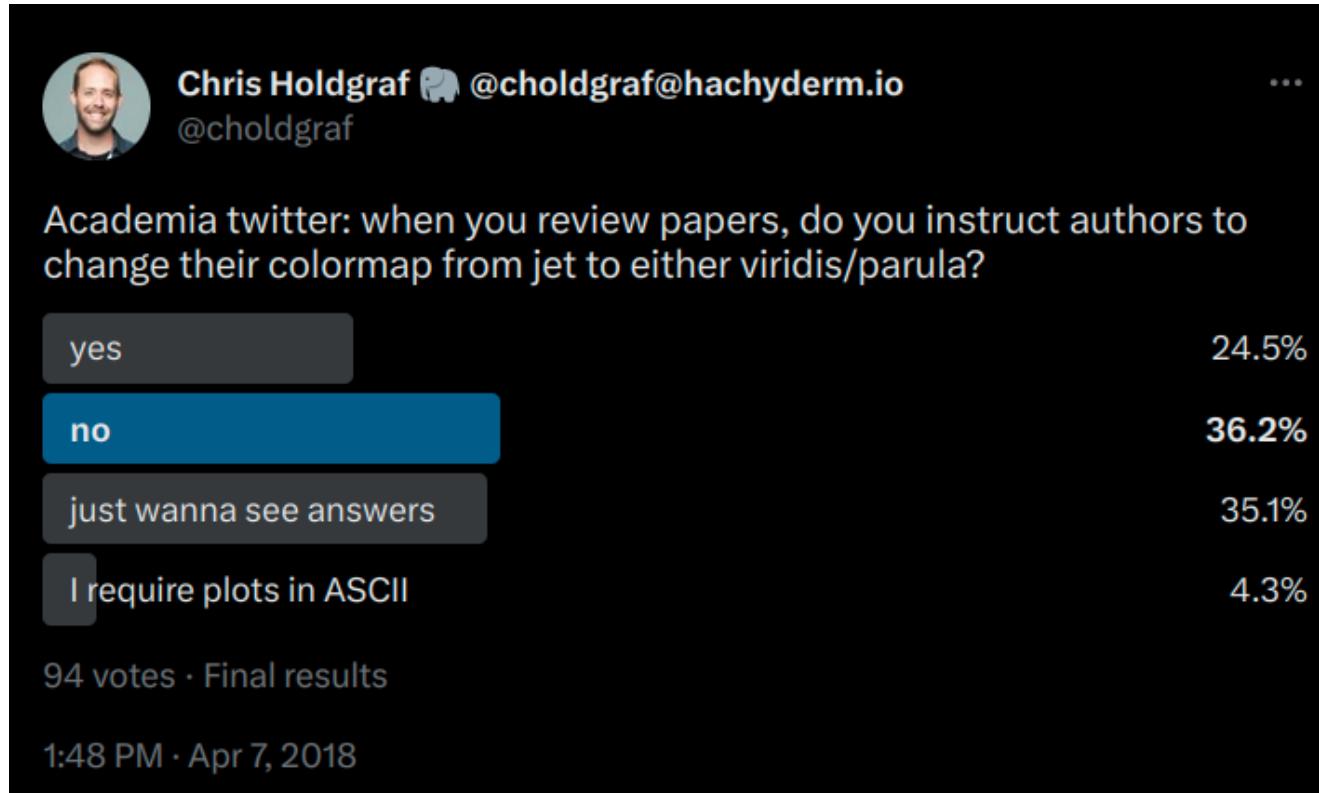
Science: Rainbow colormaps are disliked

← **Tweet**

 **Patrick Mineault** 
@patrickmineault

Does anybody else have a visceral negative reaction when they see the jet colormap?

Science: Rainbow colormaps are disliked



Chris Holdgraf 🐘 @choldgraf@hachyderm.io
@choldgraf

...
Academia twitter: when you review papers, do you instruct authors to change their colormap from jet to either viridis/parula?

Response	Percentage
yes	24.5%
no	36.2%
just wanna see answers	35.1%
I require plots in ASCII	4.3%

94 votes · Final results
1:48 PM · Apr 7, 2018

Science: Rainbow colormaps are disliked?

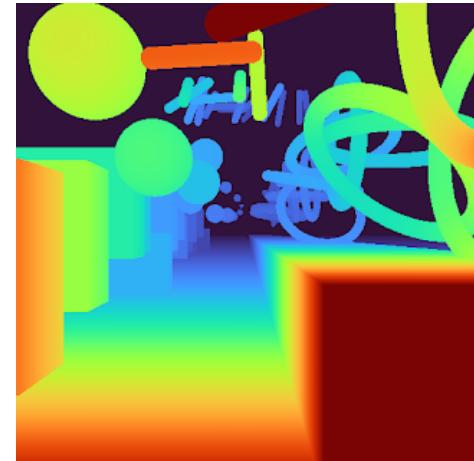


Simon Eickhoff
@INM7_ISN

Turbo, a colormap that looks like jet without its downsides

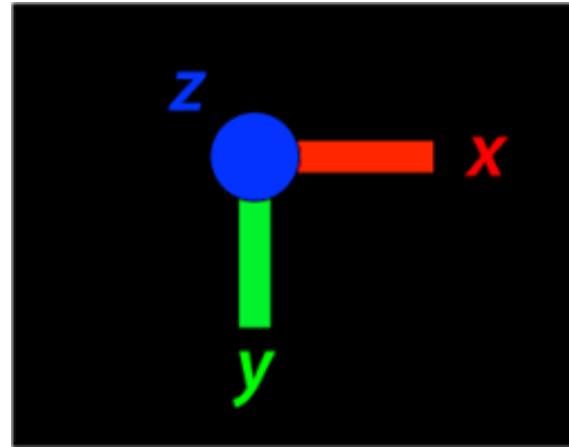
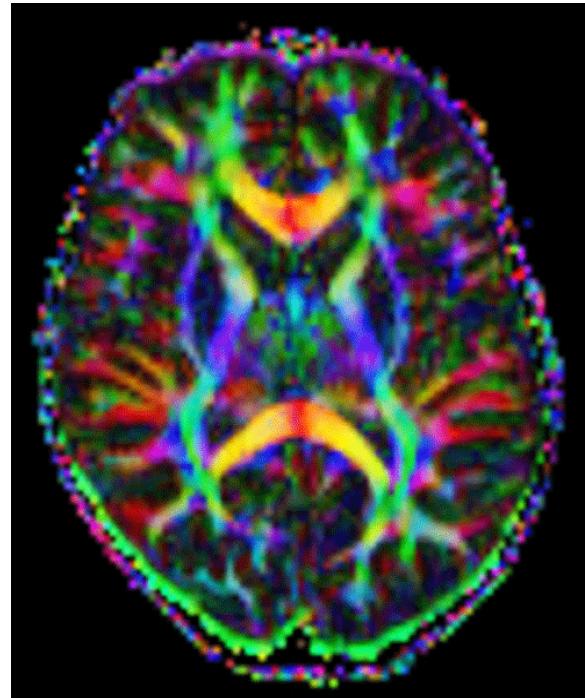
Great news for somebody like me, who likes jet as nameable colors are
indispensable in many applications: Try explaining anybody, which
shade of parula actually denotes the interesting finding

buff.ly/30h9N8O



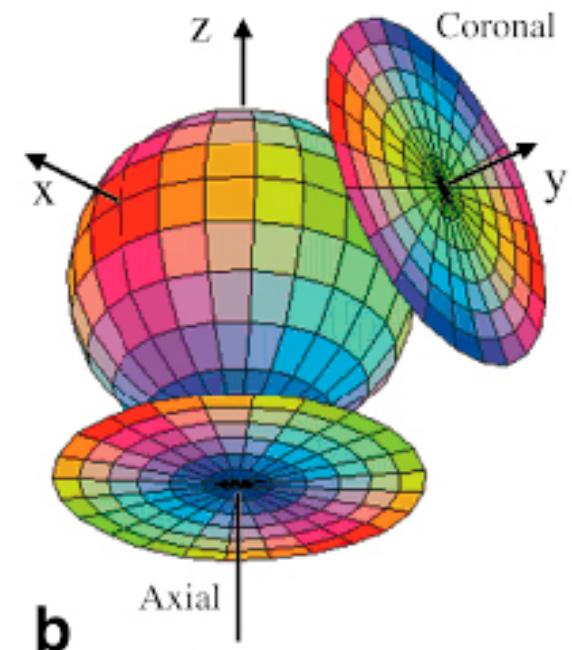
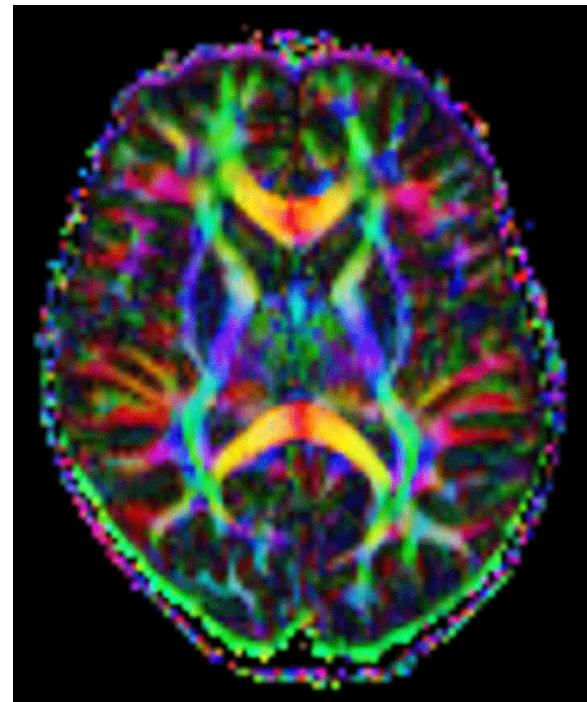
Neuroscience: Colors = 3D direction

Principal Diffusion Direction



Neuroscience: Colors = 3D direction

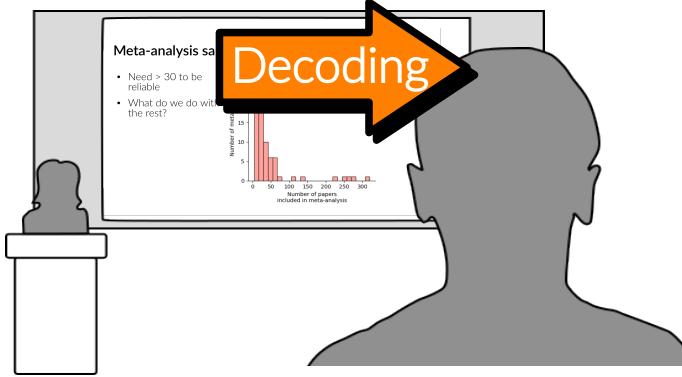
Principal Diffusion Direction



(For future reference)

Examples of visualization conventions

- Culture
 - Pink is female
 - Time goes left to right
- Science
 - Rainbow colormaps are disliked
- Neuroscience
 - MRI in grayscale
 - DWI colors = directions

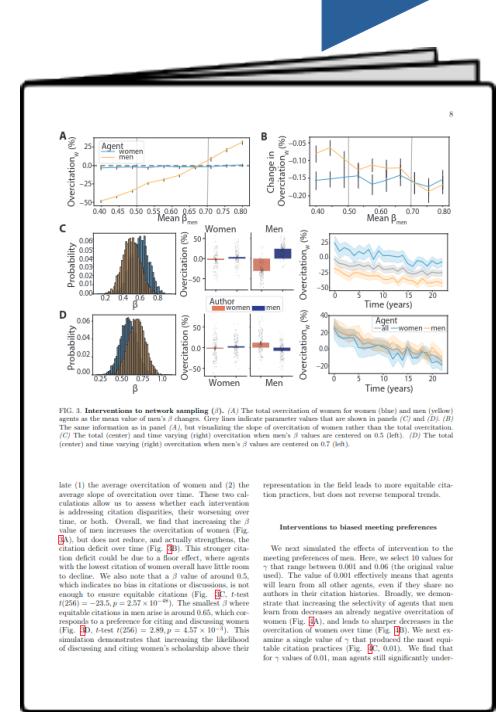
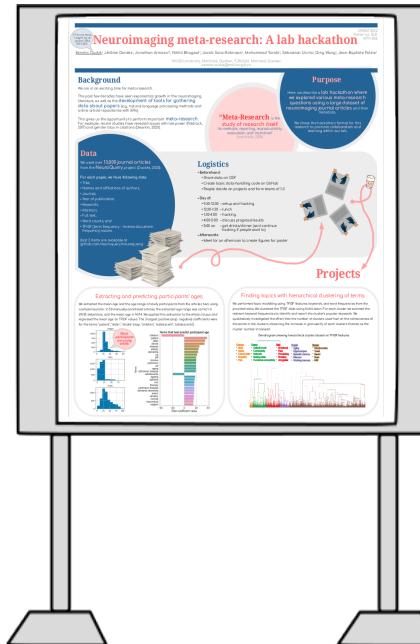
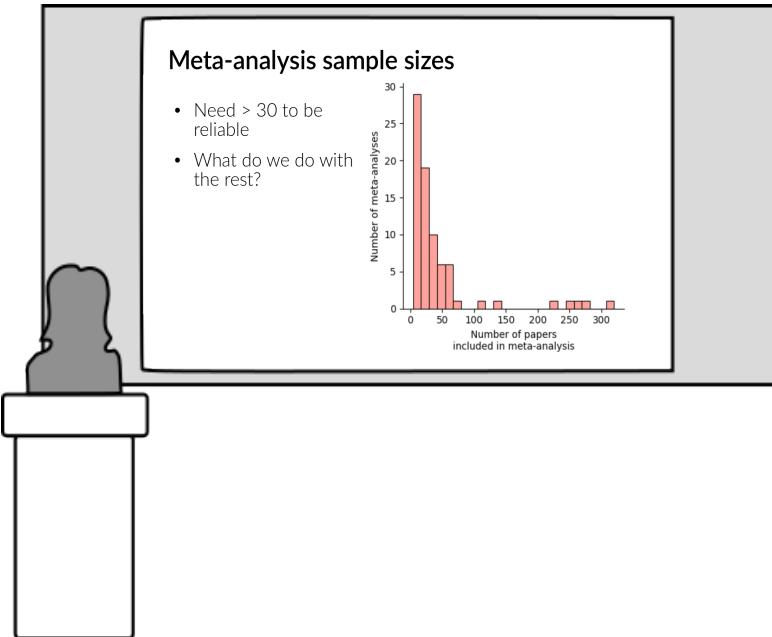


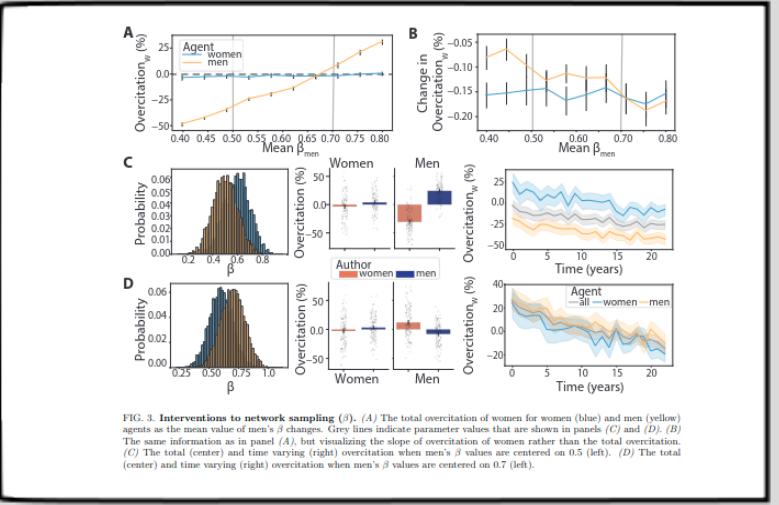
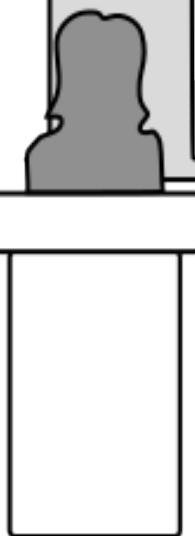
Decoding

To plan an **effective visualization**, we need to think about

- **Message**
 - What we want to communicate
- **Perception**
 - How best to communicate it
- **Conventions**
 - How it's usually communicated
- **Context**
 - Where it will be seen

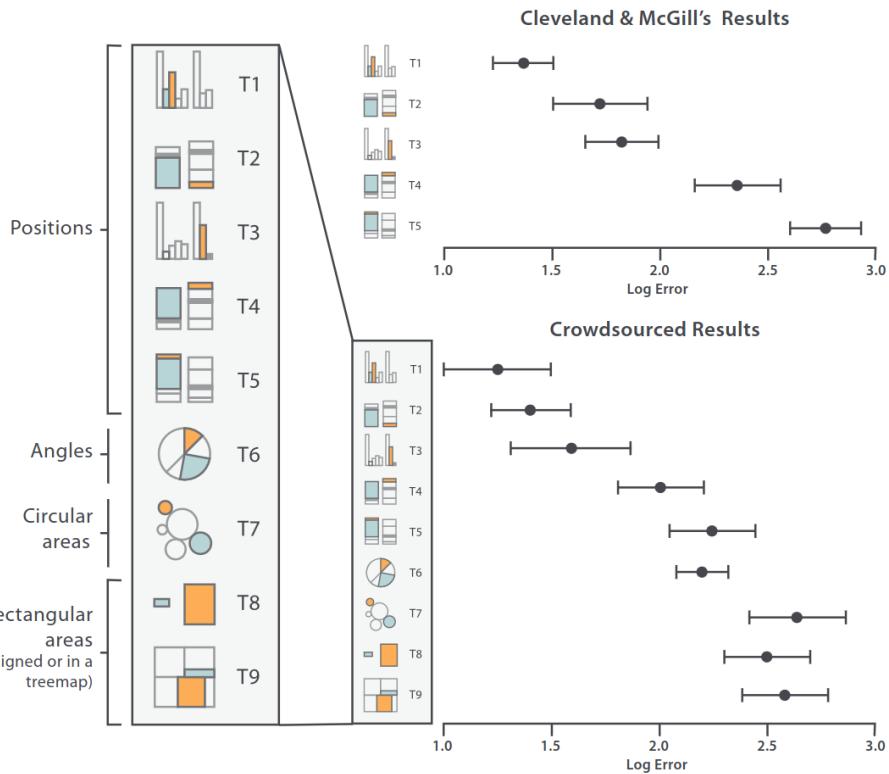
More self-explanatory





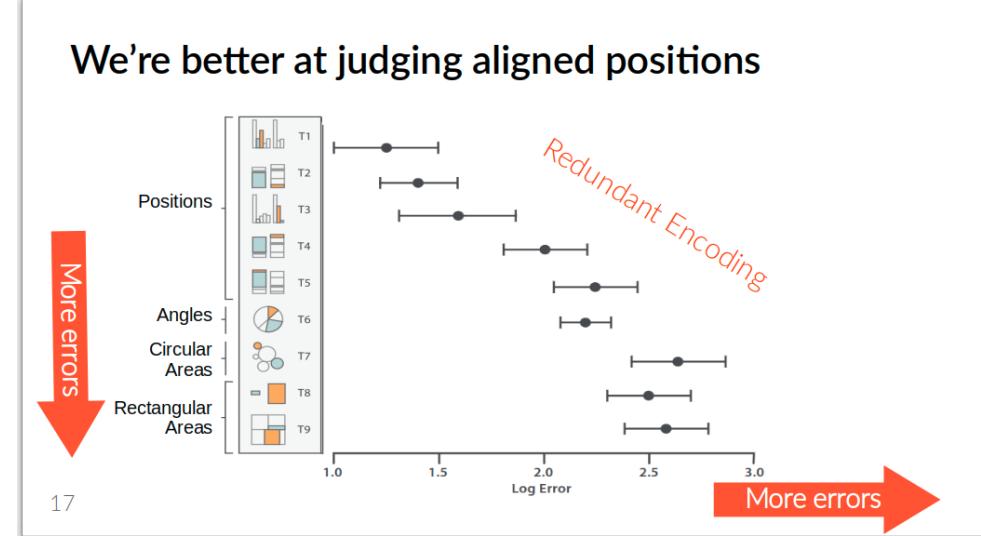
- How long would it take them to understand this slide?
- How long will they see the slide?

E.g., Figure in textbook



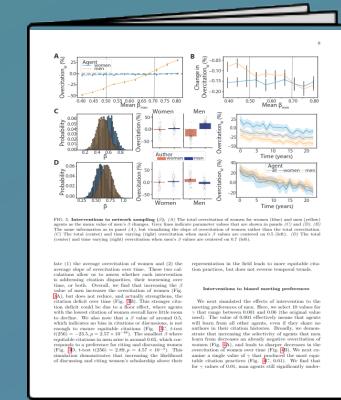
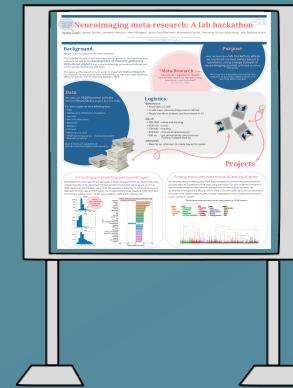
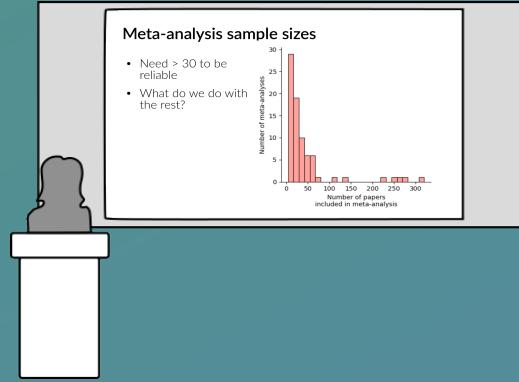
My slide

We're better at judging aligned positions



(For future reference)

Contextual adjustments



Time to understand

Proximity to viz

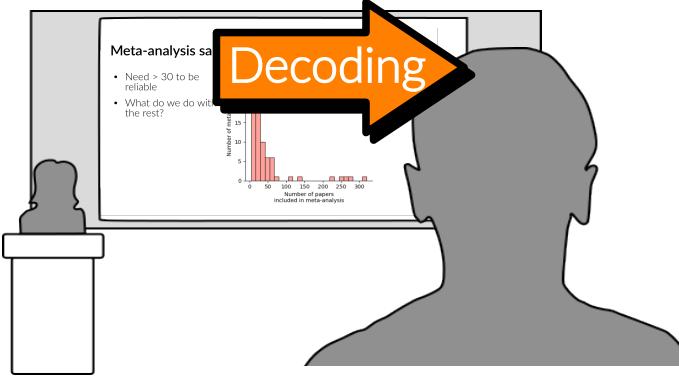
Less

 More

Text explanations

Self-explanatory

Visual complexity



To plan an **effective visualization**, we need to think about

- **Message**
 - **What** we want to communicate
- **Perception**
 - **How best** to communicate it
- **Conventions**
 - **How** it's usually communicated
- **Context**
 - **Where** it will be seen

Example 1 – Simple data and figure

- Original paper on the ABIDE dataset
 - 964 subjects
 - 396 male
 - 51 female

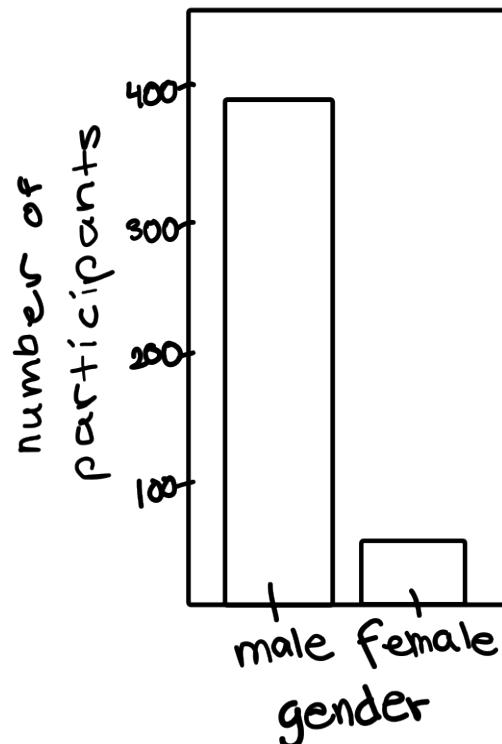
← Lets visualize this

What's our message?

- More male than female participants
 - 2 categories, 2 values

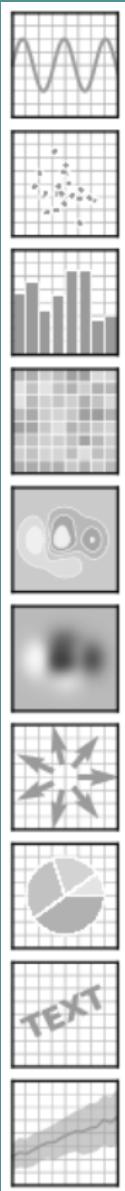
How should we communicate this message?

- Clearly show 2 categories and 2 different values



(For future reference)

Choosing effective charts



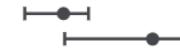
Channels: Expressiveness Types and Effectiveness Ranks

⇒ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



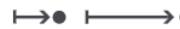
Tilt/angle



Area (2D size)



Depth (3D position)



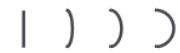
Color luminance



Color saturation



Curvature



Volume (3D size)



⇒ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



Shape



Text

Fr
Ca

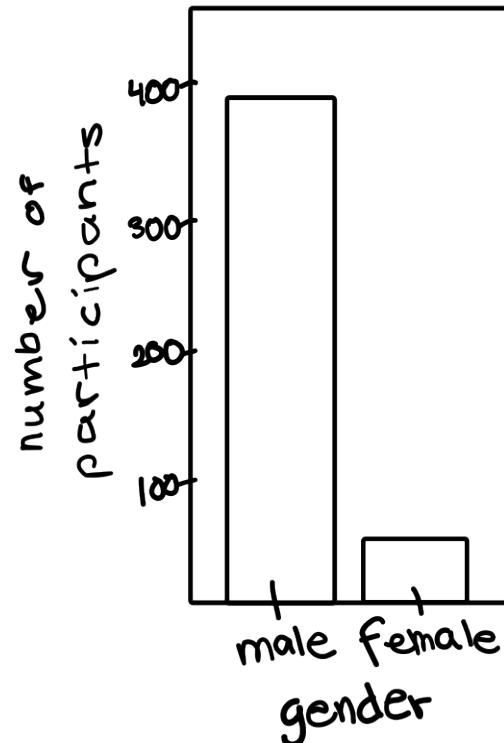
Canada Canada Canada
France France France

https://matplotlib.org/cheatsheets/_images/cheatsheets-1.png

Munzner, T. (2014). Visualization analysis and design. CRC press.

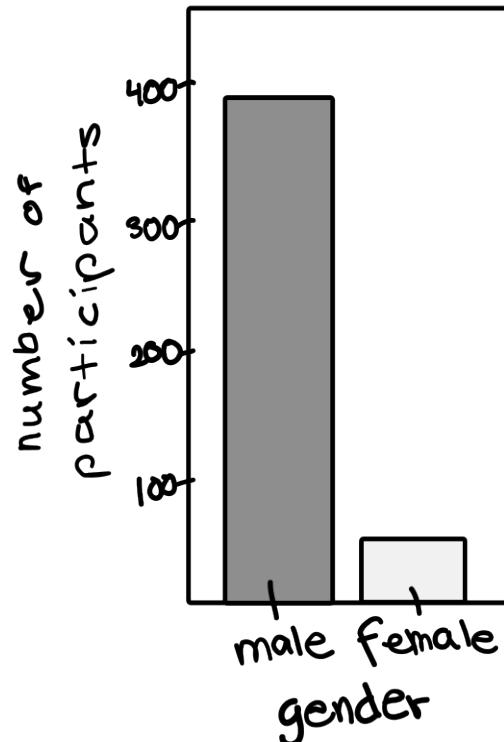
How should we communicate this message?

- Clearly show 2 categories and 2 different values



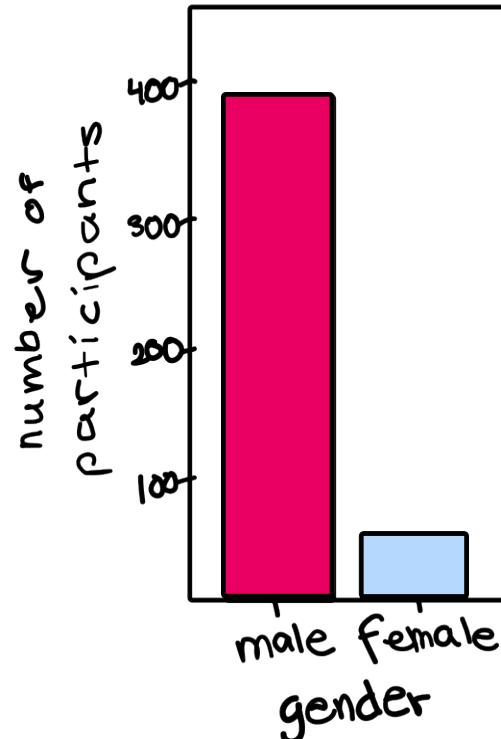
How should we communicate this message?

- Clearly show 2 categories and 2 different values



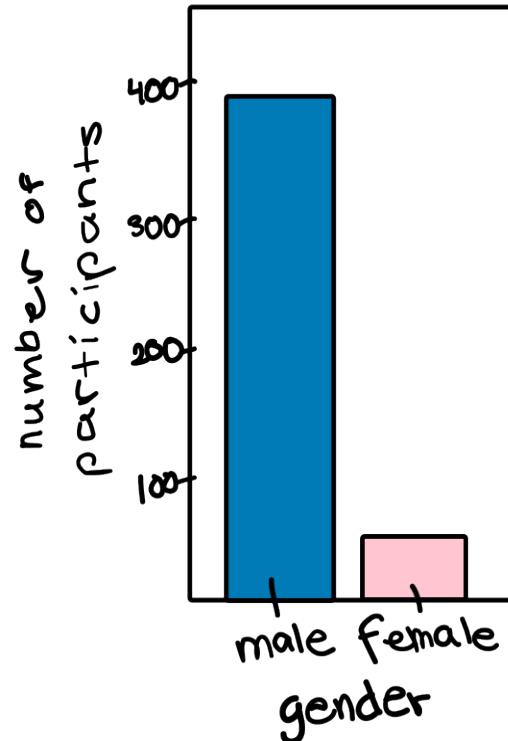
How should we communicate this message?

- Clearly show 2 categories and 2 different values



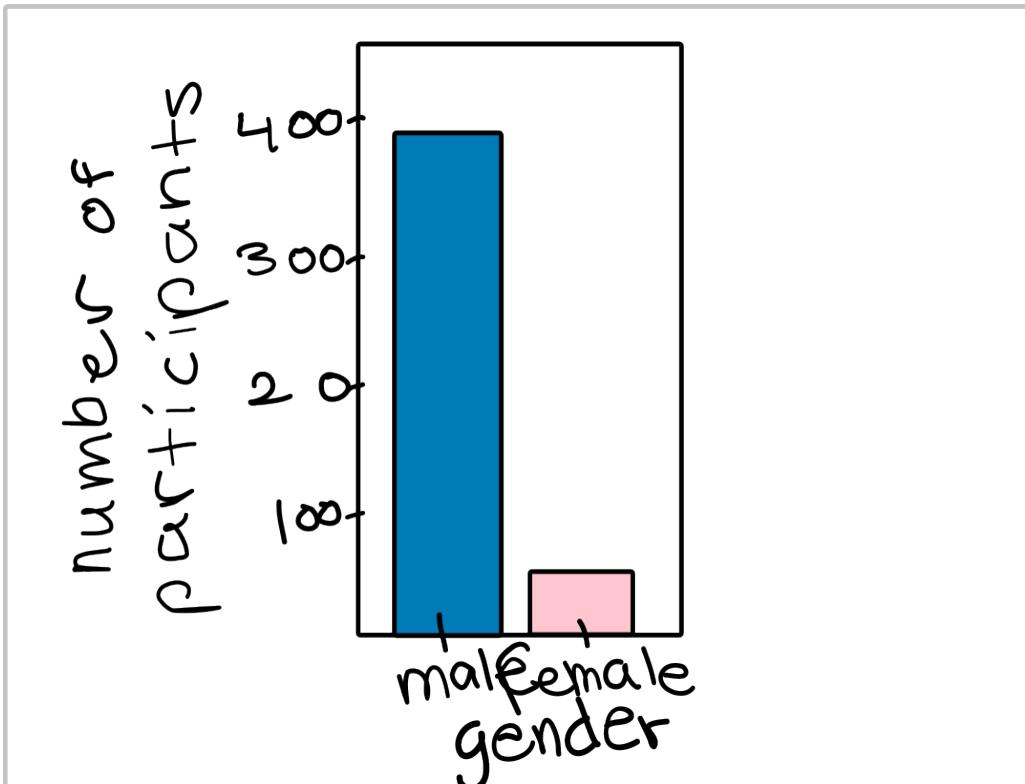
Are there any conventions we should think about?

- Gender & color



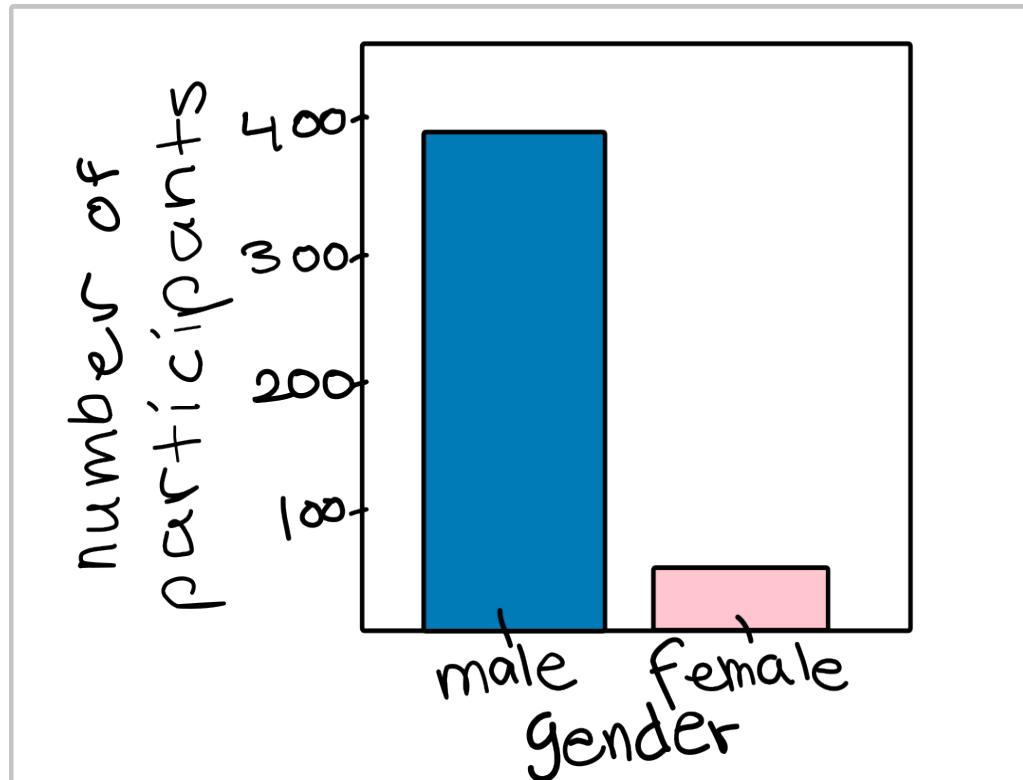
Do we need to adapt it to the context?

- Academic presentation



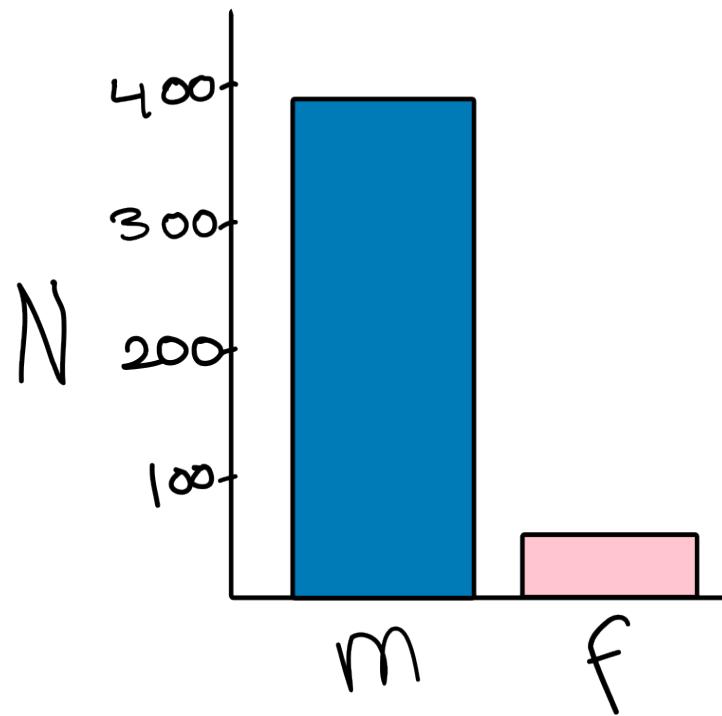
Do we need to adapt it to the context?

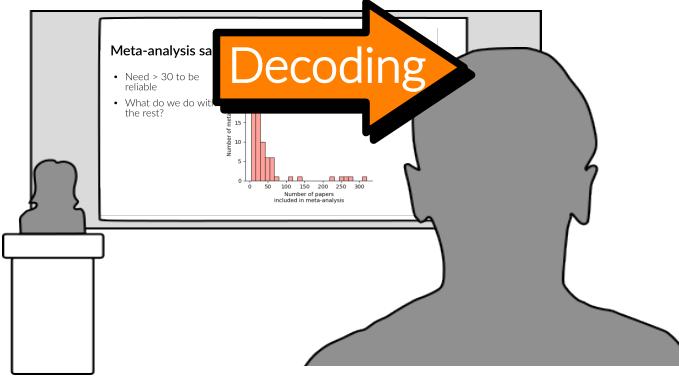
- Academic presentation



Do we need to adapt it to the context?

- Academic presentation





To plan an **effective visualization**, we need to think about

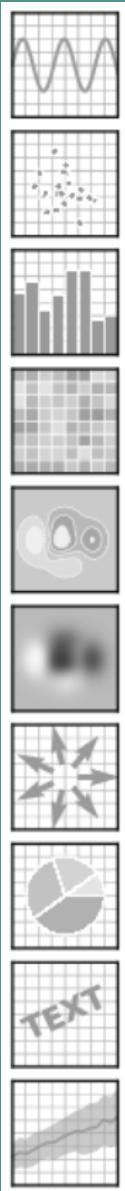
- **Message**
 - **What** we want to communicate
- **Perception**
 - **How best** to communicate it
- **Conventions**
 - **How** it's usually communicated
- **Context**
 - **Where** it will be seen

The end of part 1

Reference slides...

(For future reference)

Choosing effective charts



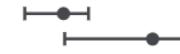
Channels: Expressiveness Types and Effectiveness Ranks

⇒ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



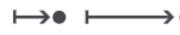
Tilt/angle



Area (2D size)



Depth (3D position)



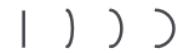
Color luminance



Color saturation



Curvature



Volume (3D size)



⇒ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



Shape



Text

Fr
Ca

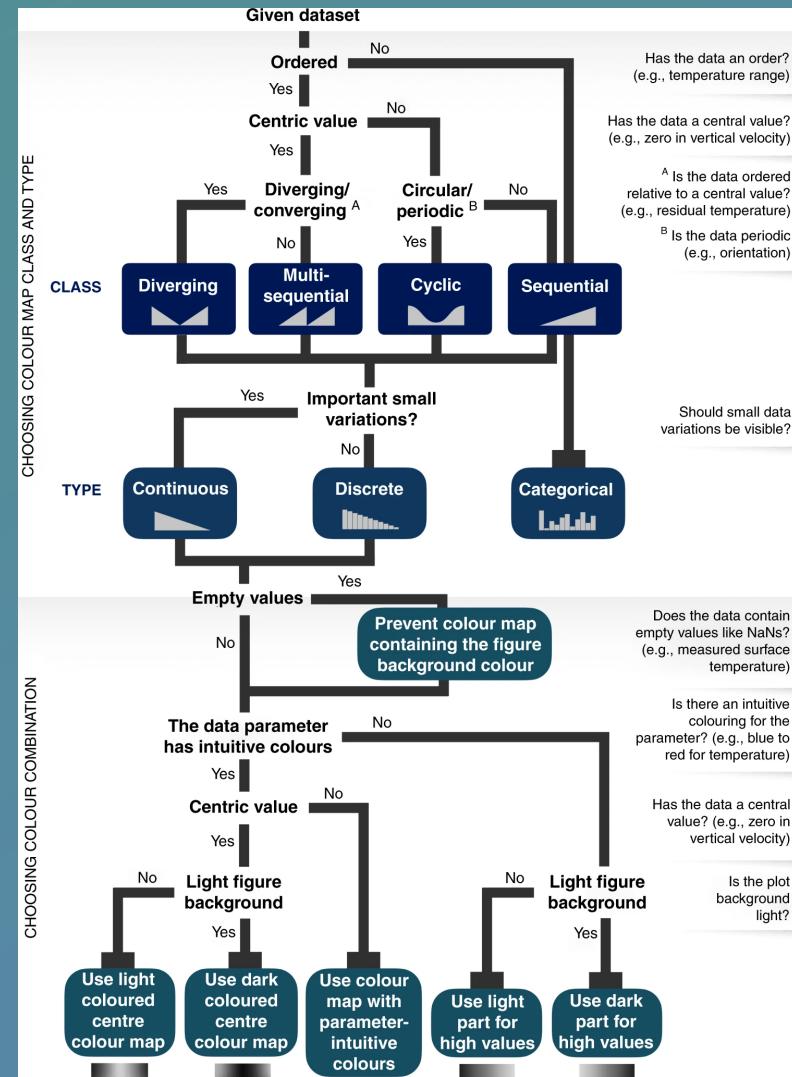
Canada Canada Canada
France France France

https://matplotlib.org/cheatsheets/_images/cheatsheets-1.png

Munzner, T. (2014). Visualization analysis and design. CRC press.

(For future reference)

Choosing effective colormaps



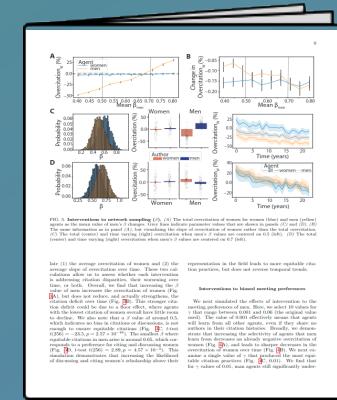
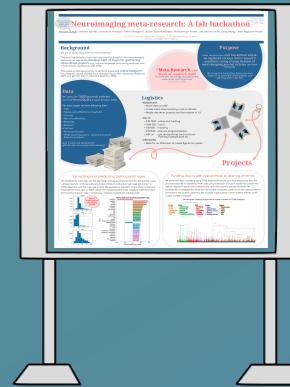
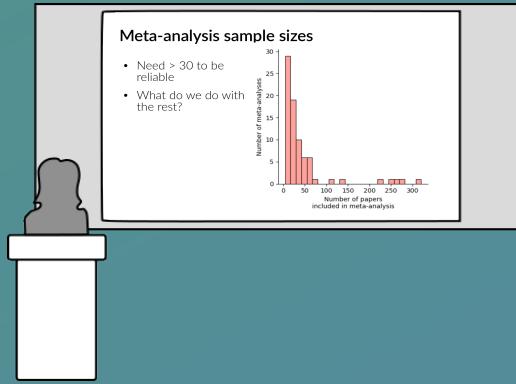
(For future reference)

Examples of visualization conventions

- Culture
 - Pink is female
 - Time goes left to right
- Science
 - Rainbow colormaps are disliked
- Neuroscience
 - MRI in grayscale
 - DWI colors = directions

(For future reference)

Contextual adjustments



Time to understand

Proximity to viz

Less

Text explanations

More

Self-explanatory

Visual complexity