An Algebraic Process for Visualization Design

Carlos Scheidegger, Gordon Kindlmann

VIS 2014

Test Suites for Visualization

- How do we know that a visualization is doing the right thing?
  - What is even the right thing?

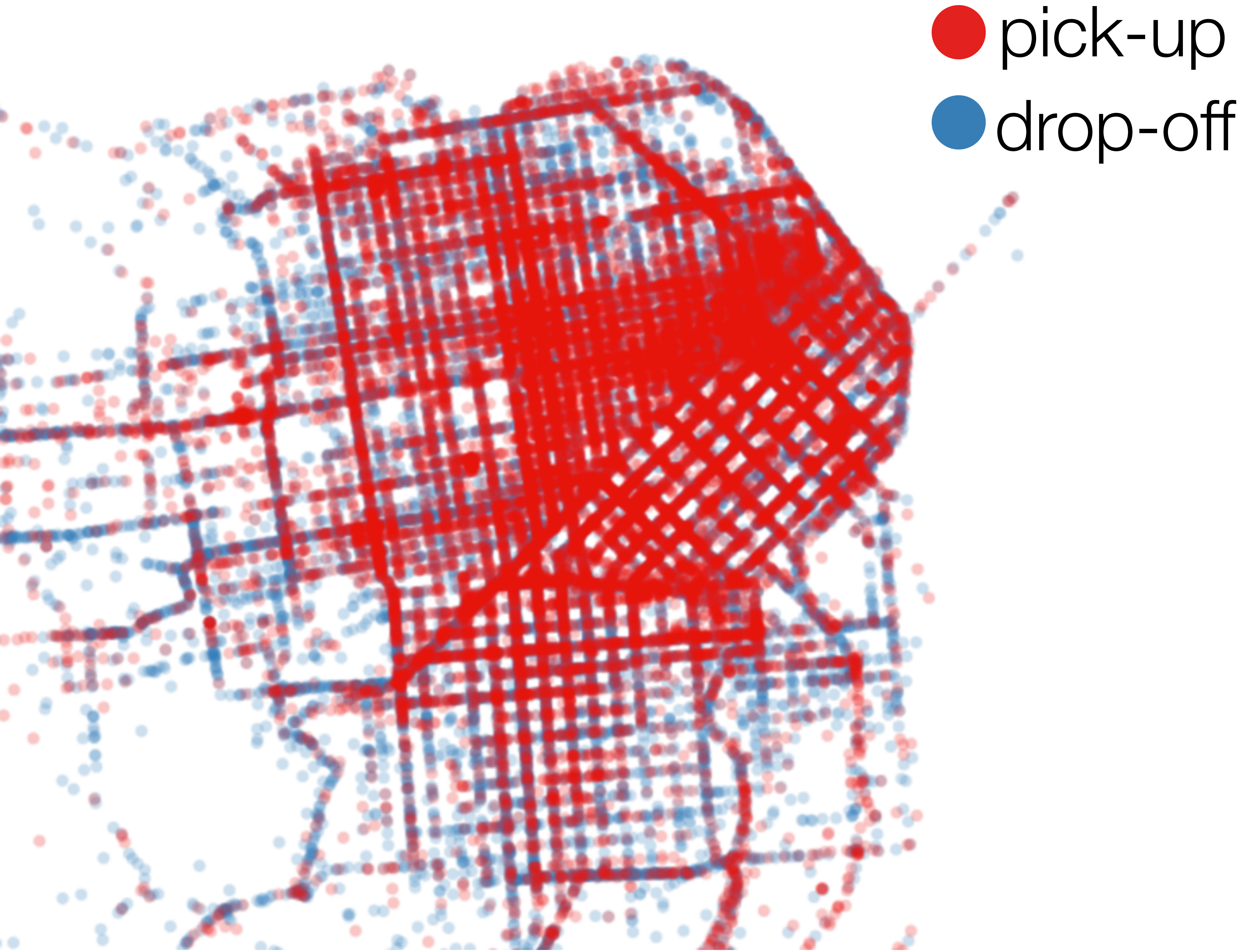
EVALUATION

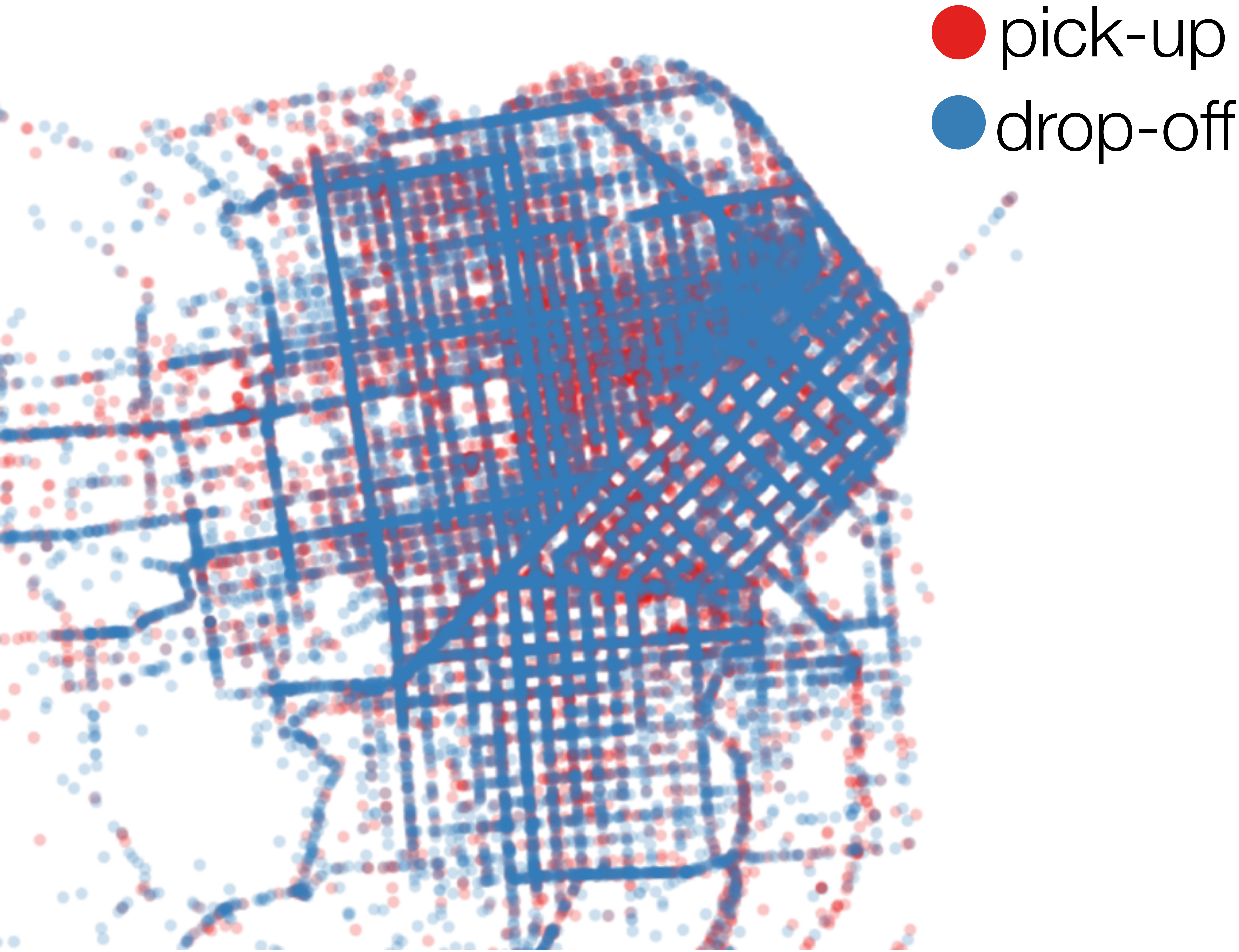
Evaluation through User Studies

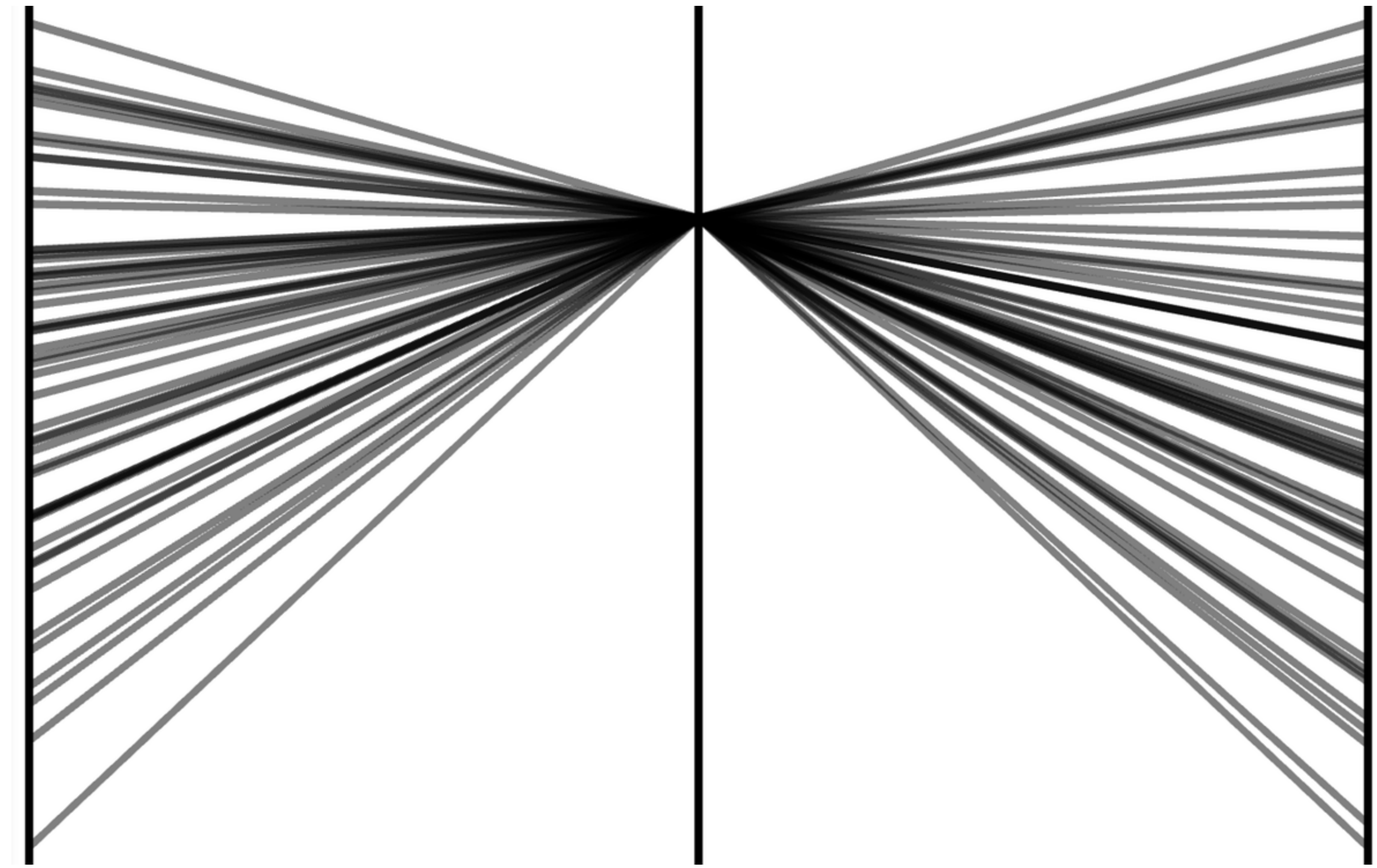
- Define tasks, run user study, measure variable, do stats
    - Very hard to do right, time-consuming, expensive
    - and even harder for conclusions to generalize
- Whole courses are taught entirely about this - we're not going to do that

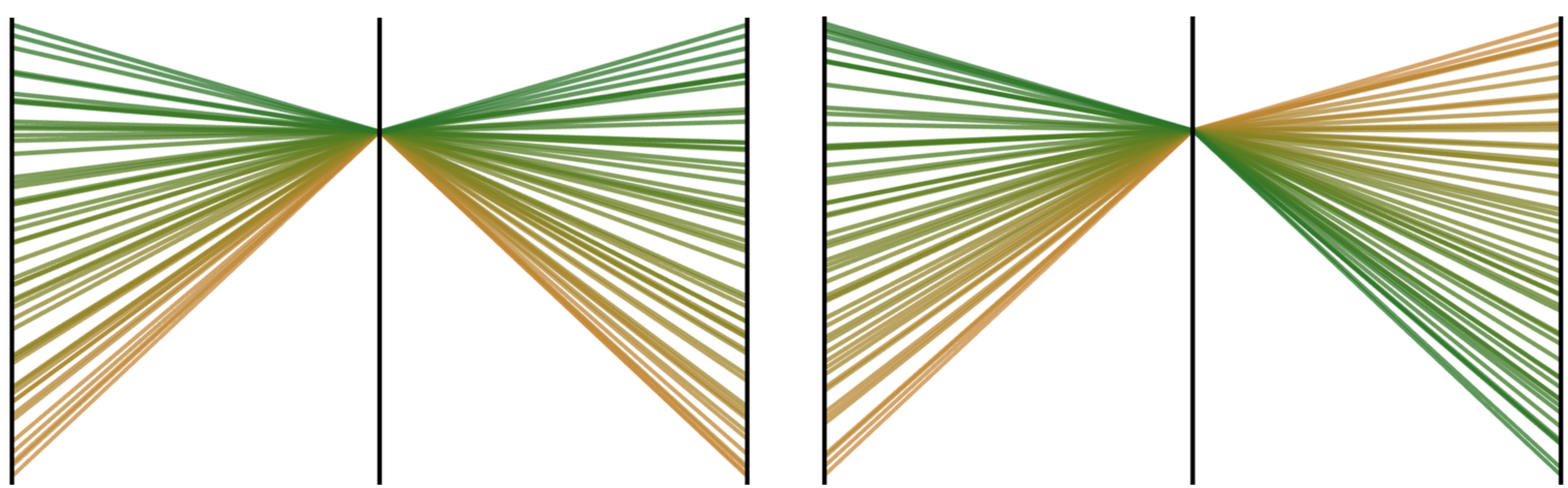
“Evaluation through Imagination”

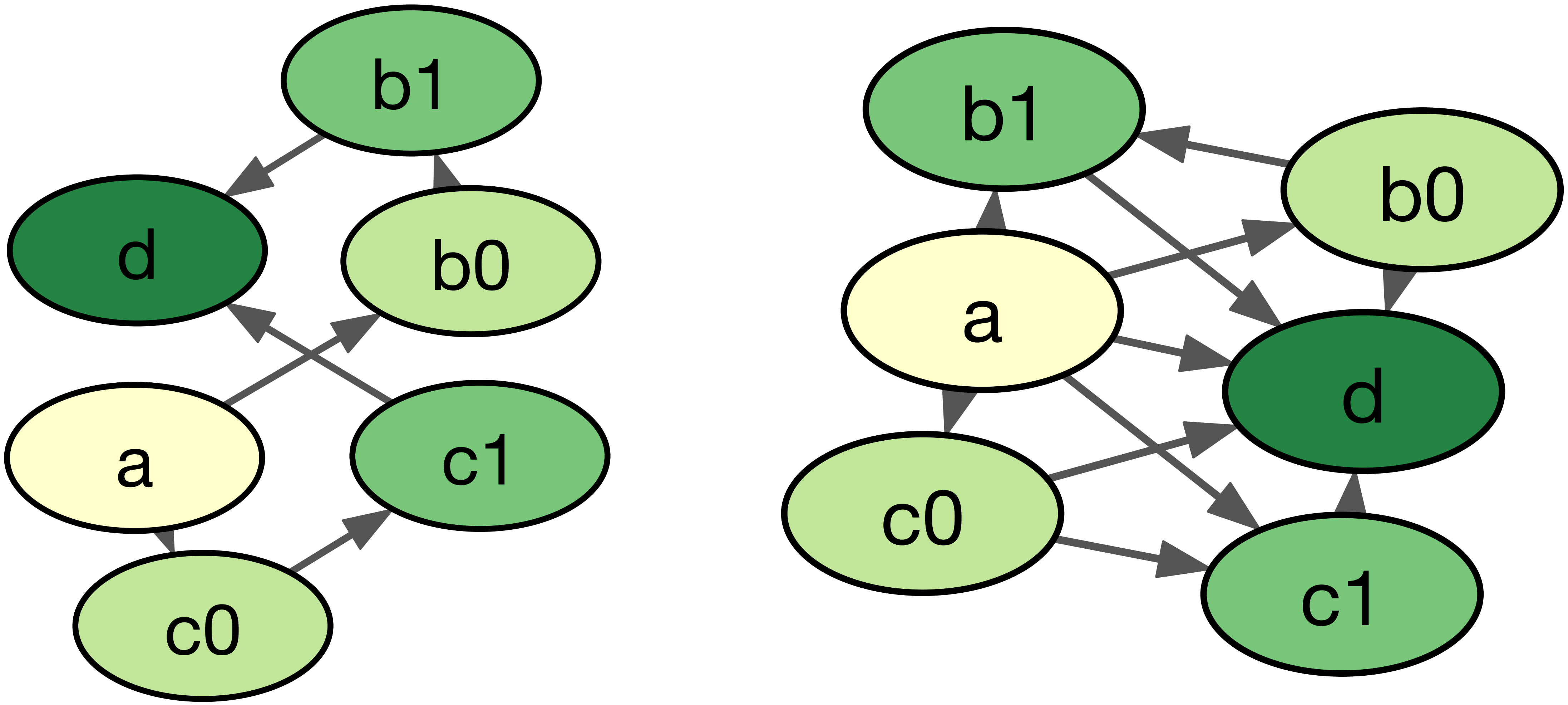
- Instead, we are going to use **thought experiments**:
  - **What if the input were different** - what would this change cause?
  - **What if the picture were different** - how could the input have been different?
- The answers tell us a lot about the visualization
  - Not as good as a good user study, but **practical**

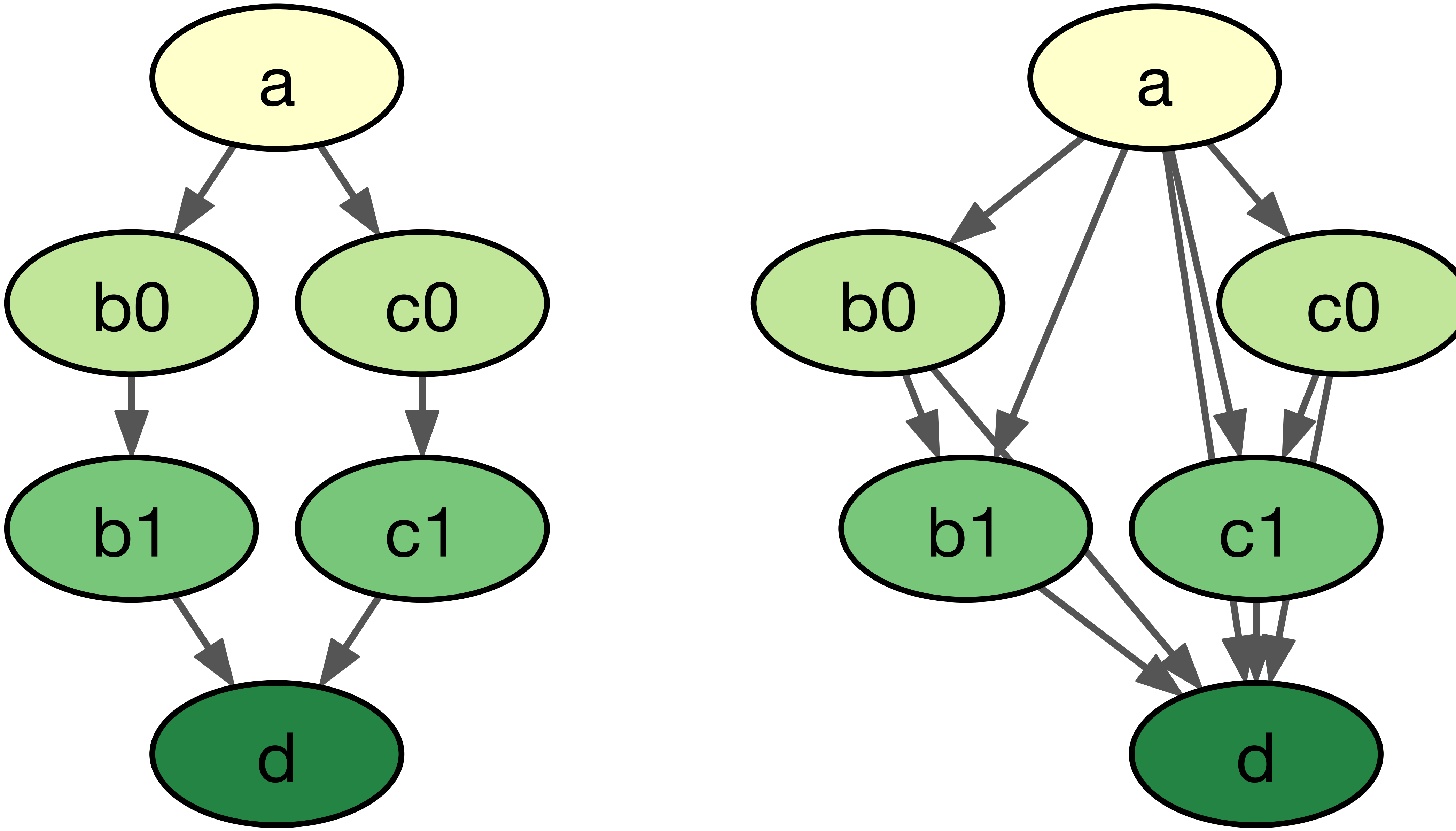

pick-up
drop-off


pick-up
drop-off

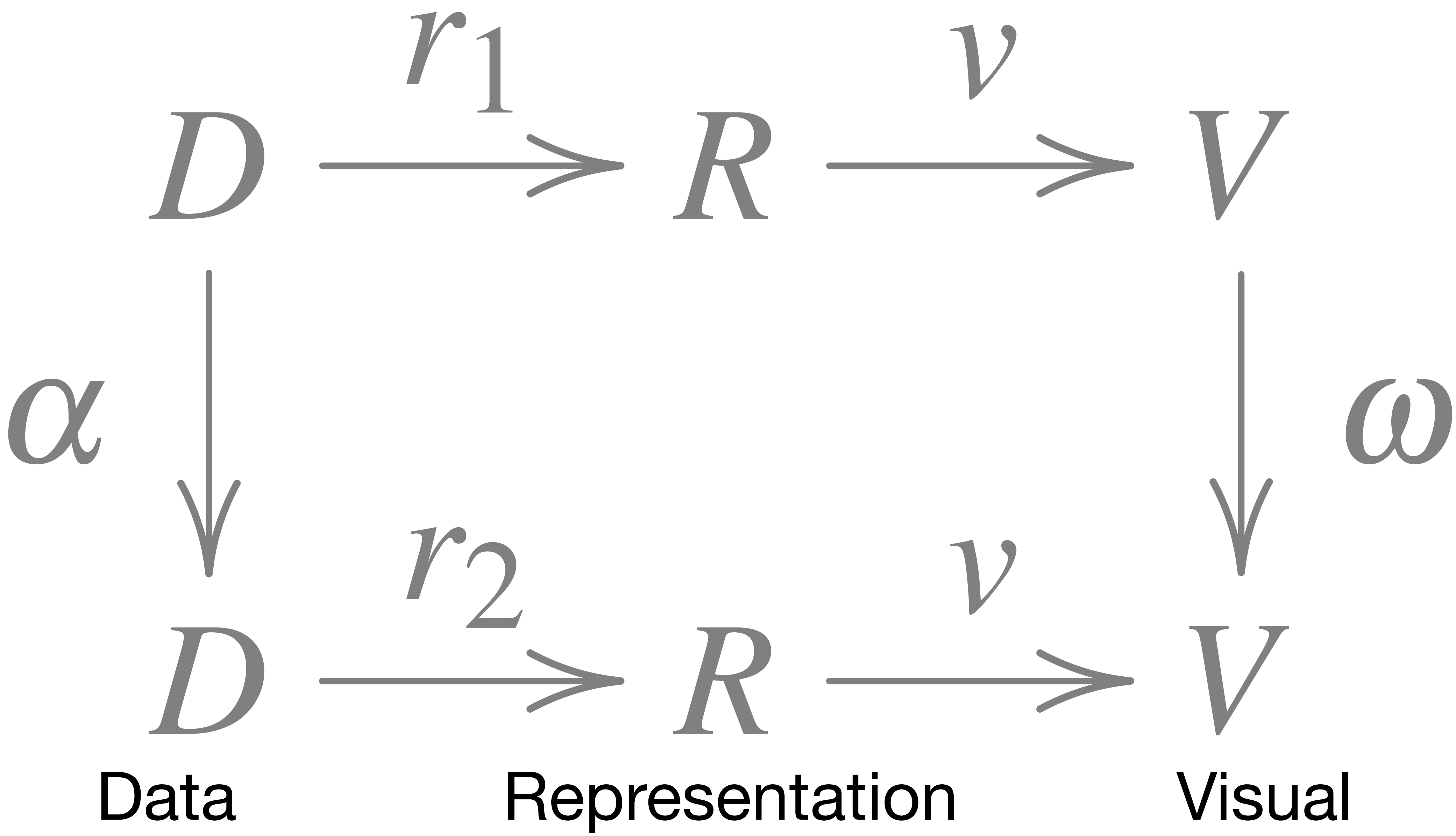





b1
d
b0
a
c1
c0
b1
b0
a
d
c0
c1


a
b0
c0
b1
c1
d
a
b0
c0
b1
c1
d

We want a theory to **explain**, **critique** and **suggest** visualizations



Equation 1

$$v \circ r_2 \circ \alpha = \omega \circ v \circ r_1$$

Failure of The Invariance Principle

(list in some order)

$$D \xrightarrow{r_1} R \xrightarrow{v} V$$

$$(1_D =)\ \alpha \downarrow \qquad \downarrow \omega\ (\neq 1_V)$$

$$D \xrightarrow{r_2} R \xrightarrow{v} V$$

(list in reverse order)

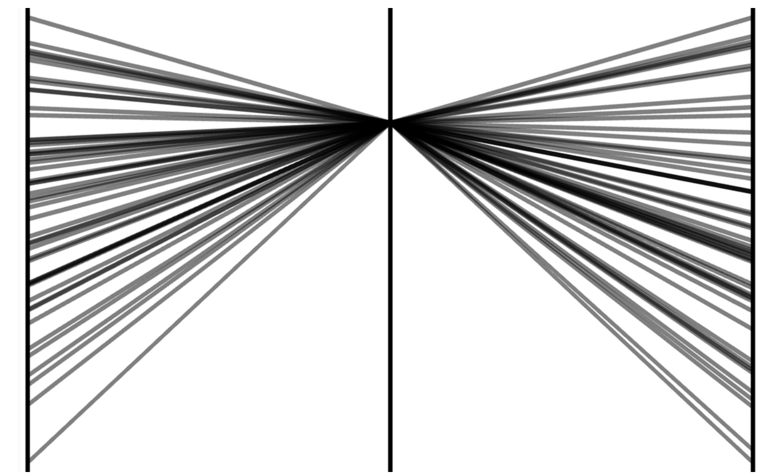
Success of The Invariance Principle

(list in some order)

$$
\begin{array}{ccccc}
D & \xrightarrow{r_1} & R & \xrightarrow{v} & V \\
(1_D =)\, \alpha \downarrow & & & & \downarrow \omega\, (= 1_V) \\
D & \xrightarrow{r_2} & R & \xrightarrow{v} & V
\end{array}
$$

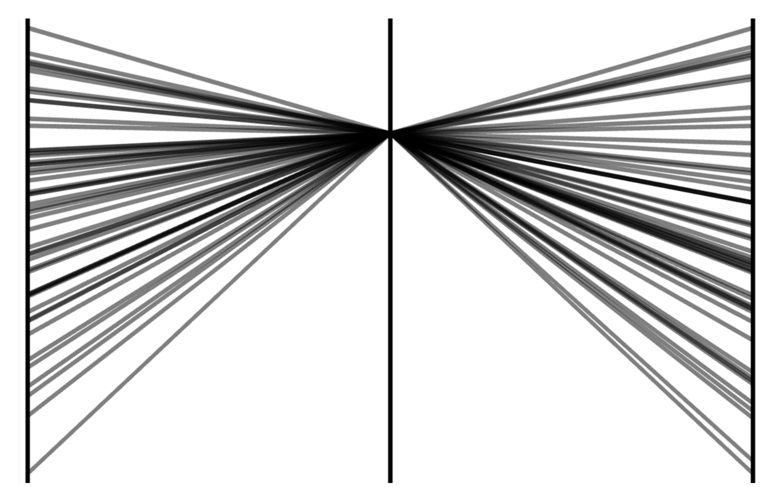
(list in reverse order)

Failure of The Unambiguity Principle

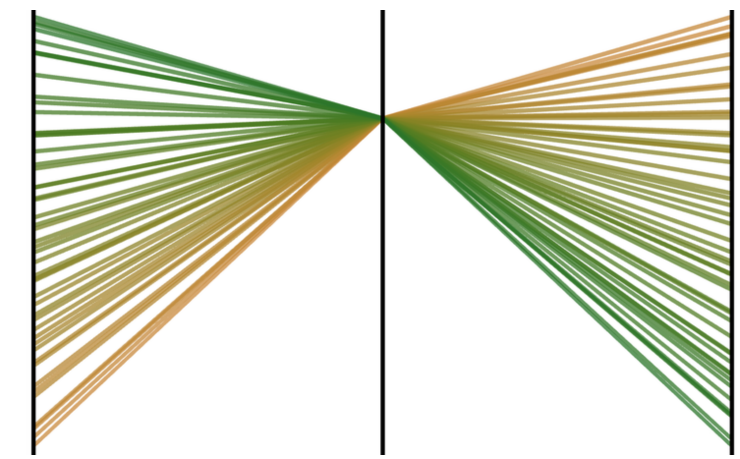


$$
\begin{array}{ccccc}
D & \xrightarrow{r_1} & R & \xrightarrow{v} & V \\
(1_D \neq)\,\alpha \downarrow & & & & \downarrow \omega\,(= 1_V) \\
D & \xrightarrow{r_2} & R & \xrightarrow{v} & V
\end{array}
$$

Reverse third coordinate

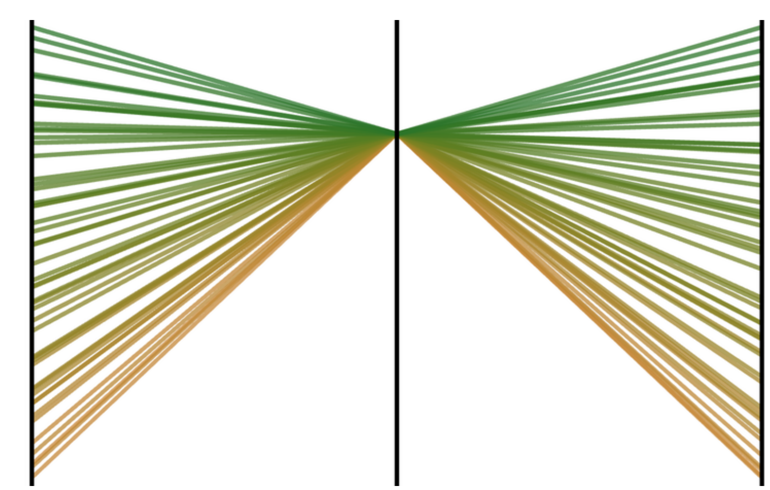


Success of The Unambiguity Principle

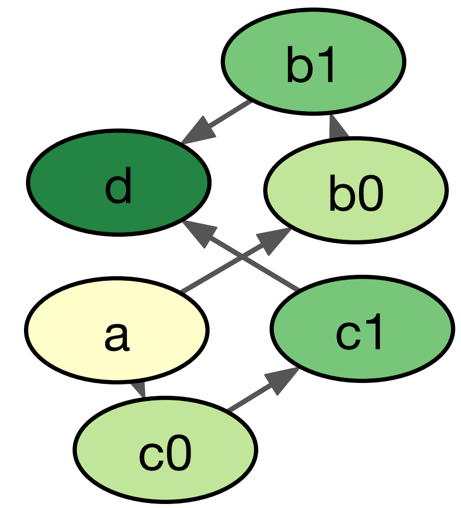


$$\begin{array}{ccccc} D & \xrightarrow{r_1} & R & \xrightarrow{v} & V \\ (1_D \neq)\,\alpha \downarrow & & & & \downarrow \omega\,(\neq 1_V) \\ D & \xrightarrow{r_2} & R & \xrightarrow{v} & V \end{array}$$

Reverse third coordinate

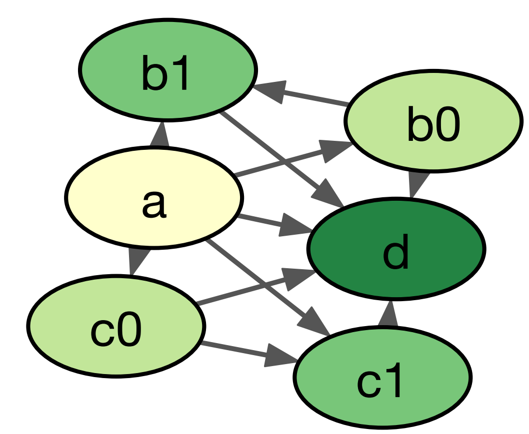


Failure of The Correspondence Principle

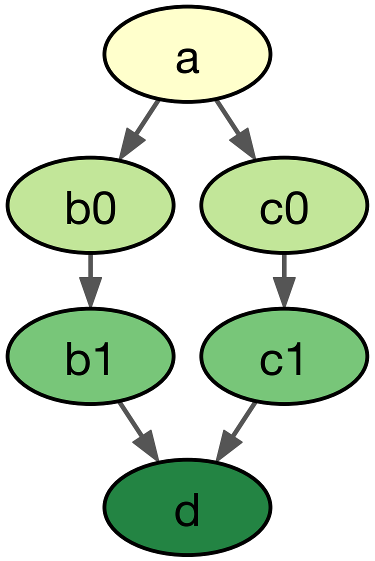


$$\begin{array}{ccccc} D & \xrightarrow{r_1} & R & \xrightarrow{v} & V \\ (1_D \neq)\ \alpha \downarrow & & & & \downarrow \omega\ (\neq 1_V) \\ D & \xrightarrow{r_2} & R & \xrightarrow{v} & V \end{array}$$

Add edges monotonically



Success of The Correspondence Principle

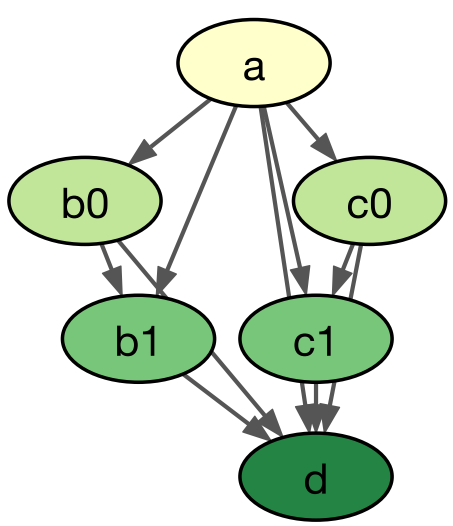


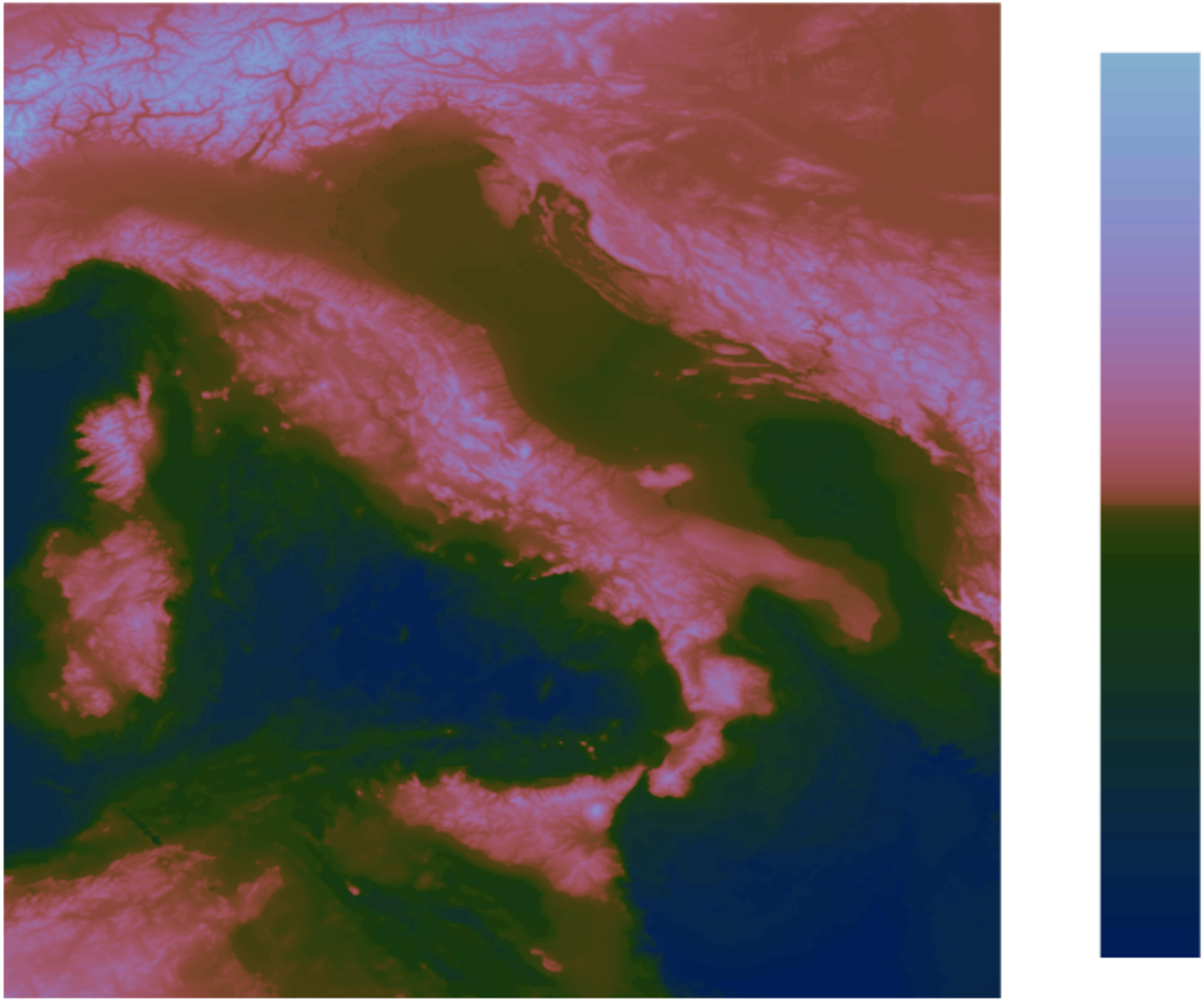
$$D \xrightarrow{r_1} R \xrightarrow{v} V$$

$(1_D \neq)\ \alpha \downarrow$ $\qquad\qquad$ $\downarrow \omega\ (\neq 1_V)$

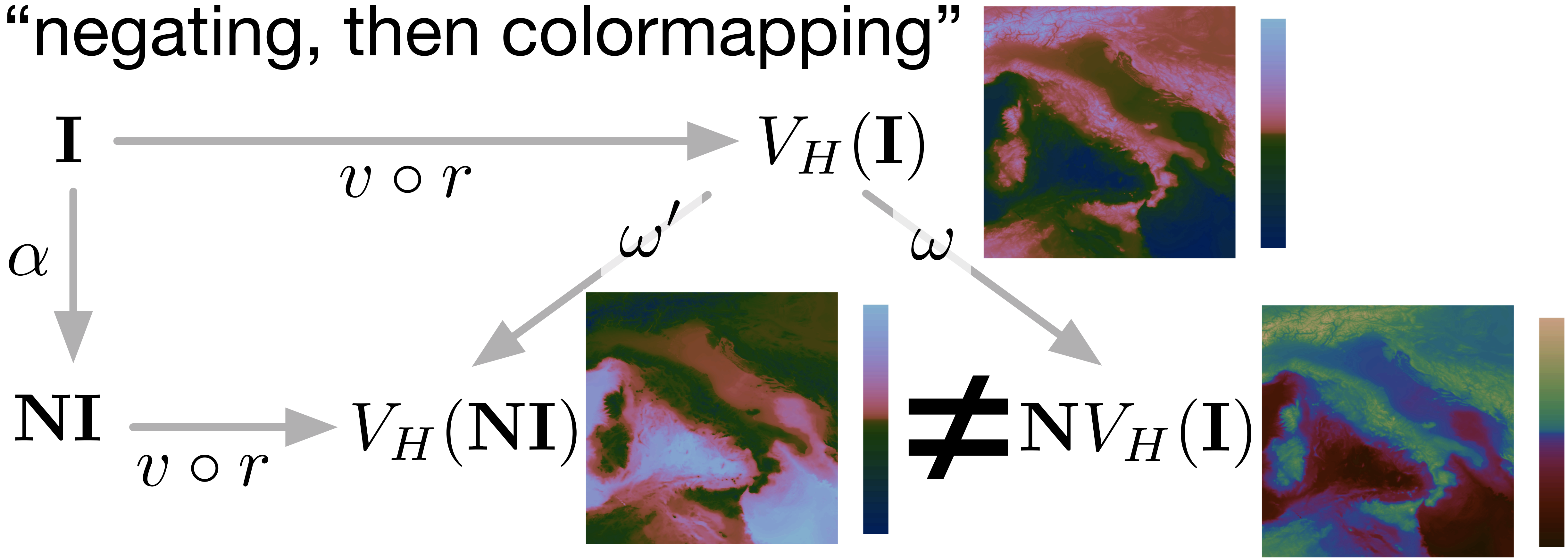
Add edges monotonically

$$D \xrightarrow{r_2} R \xrightarrow{v} V$$

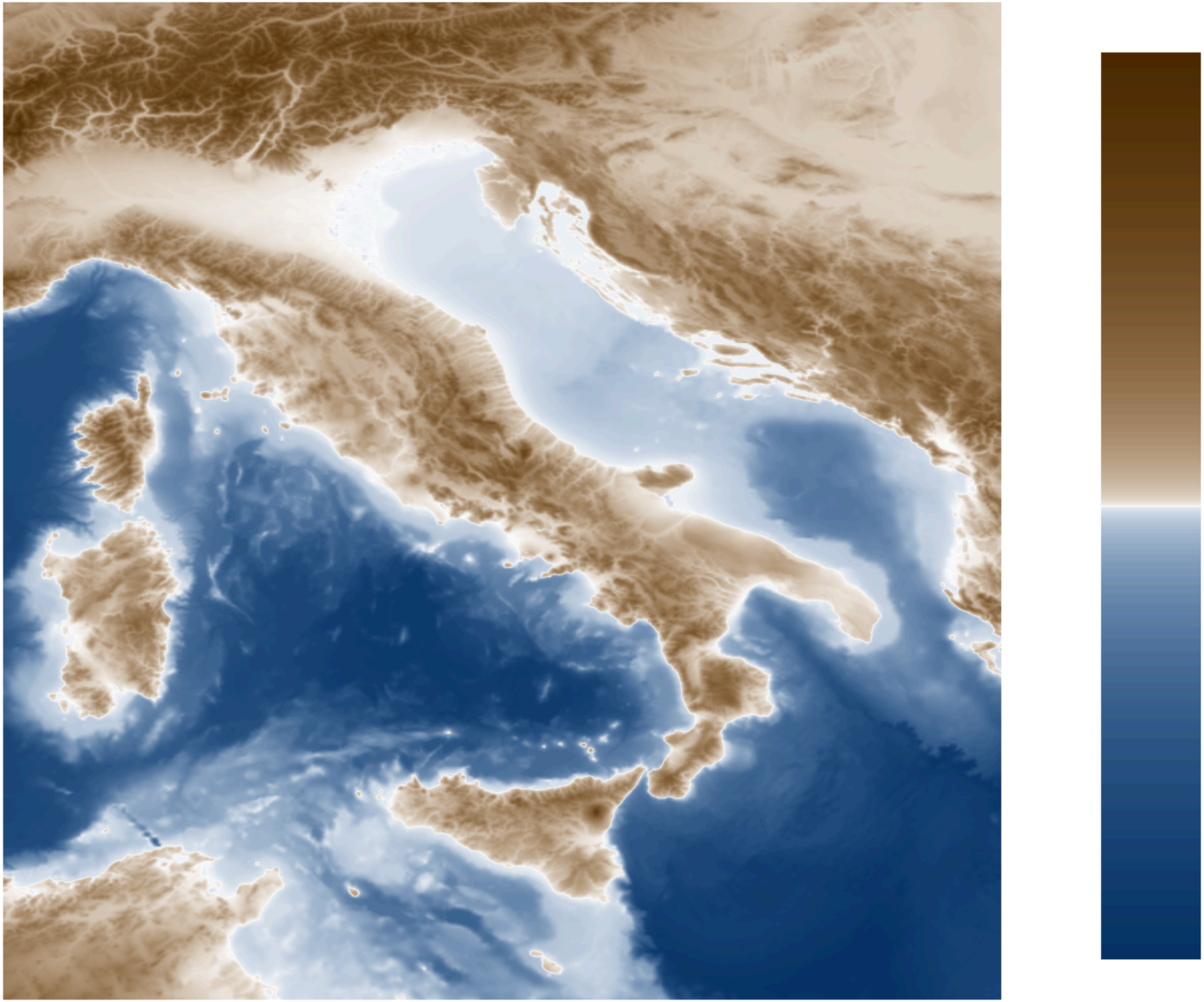


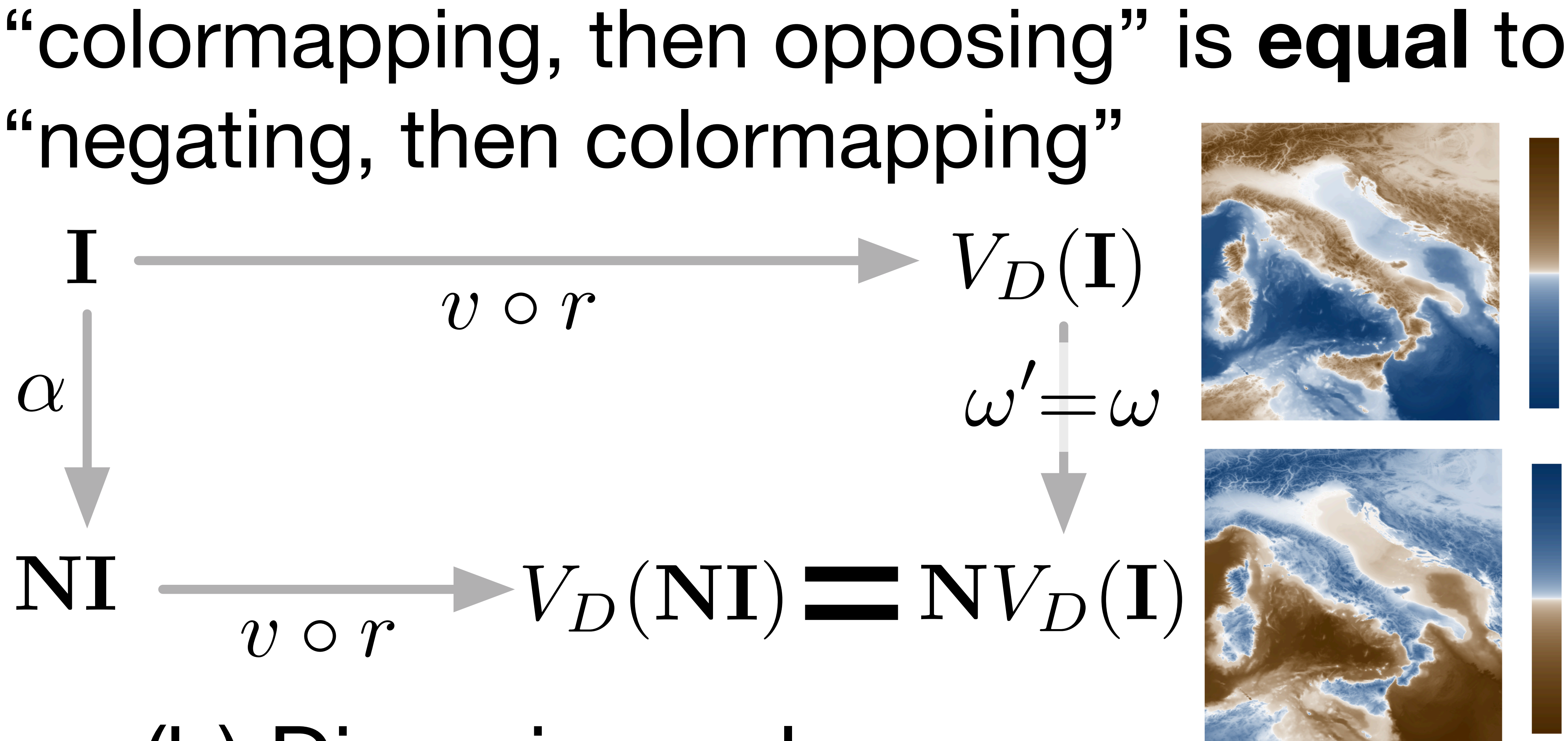


“colormapping, then opposing” **differs** from “negating, then colormapping”



(a) Hue+Luminance colormap





(b) Diverging colormap

The algebraic process:

1) Pick spaces of interest
2) Pick transformations of interest
3) Study what happens on the other side

(We want a theory to **explain**, **critique** and **suggest** visualizations)

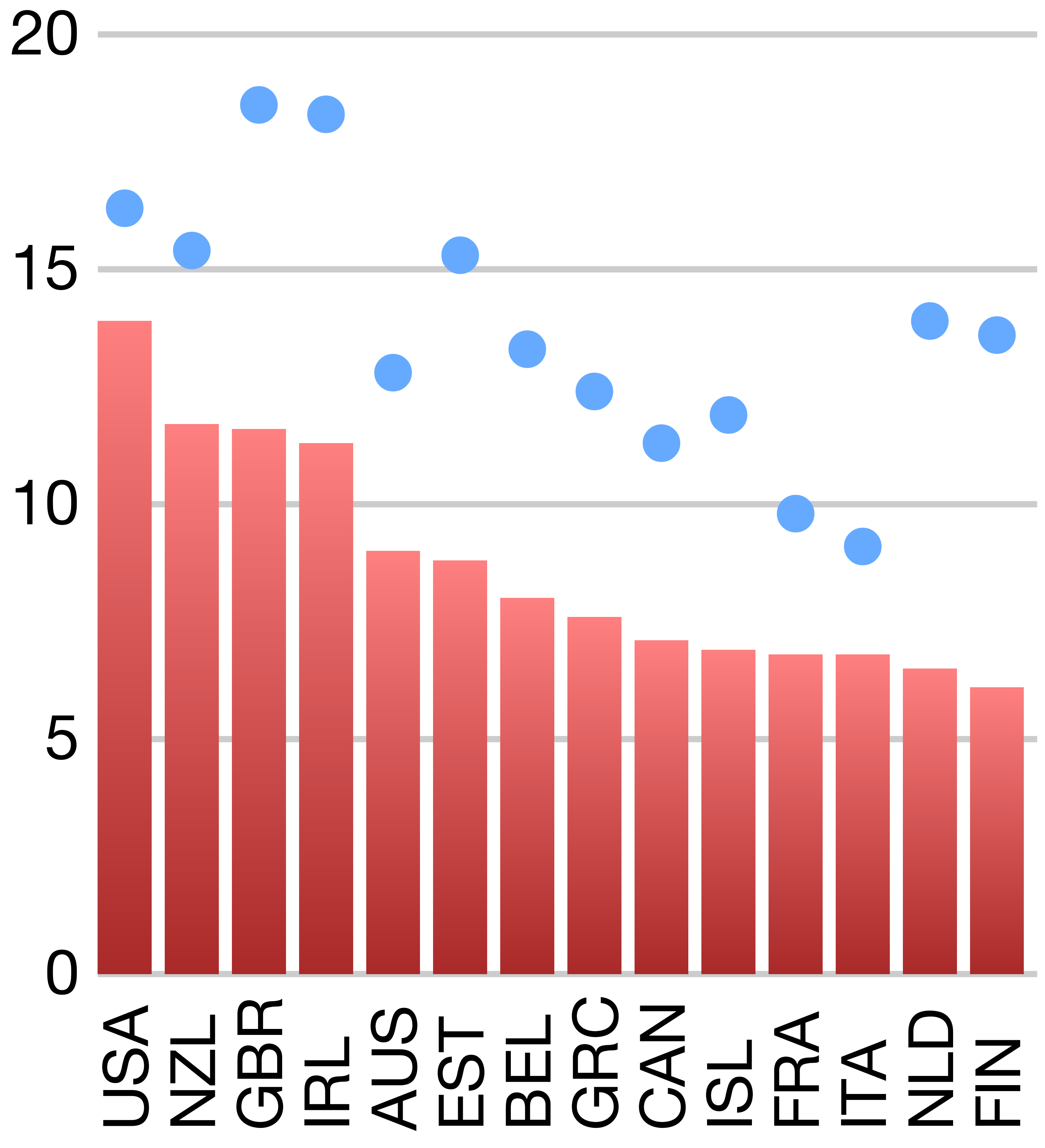
Interlude: Cleveland and McGill

Position: Good
Length: Good
Angle: Not so good
Saturation: Not so good

Case Study:

Employment rates
across countries and genders

Original visualizations by
Jonathan Schwabish and
NYT's Catherine Rampell

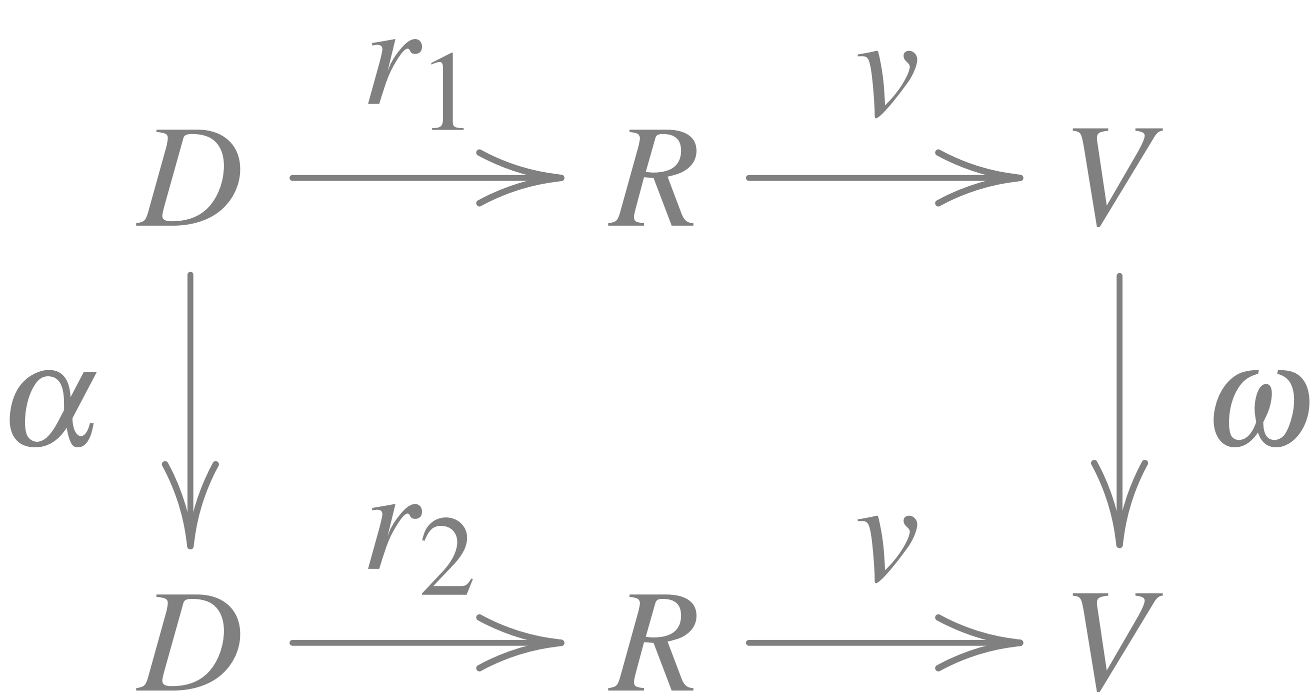


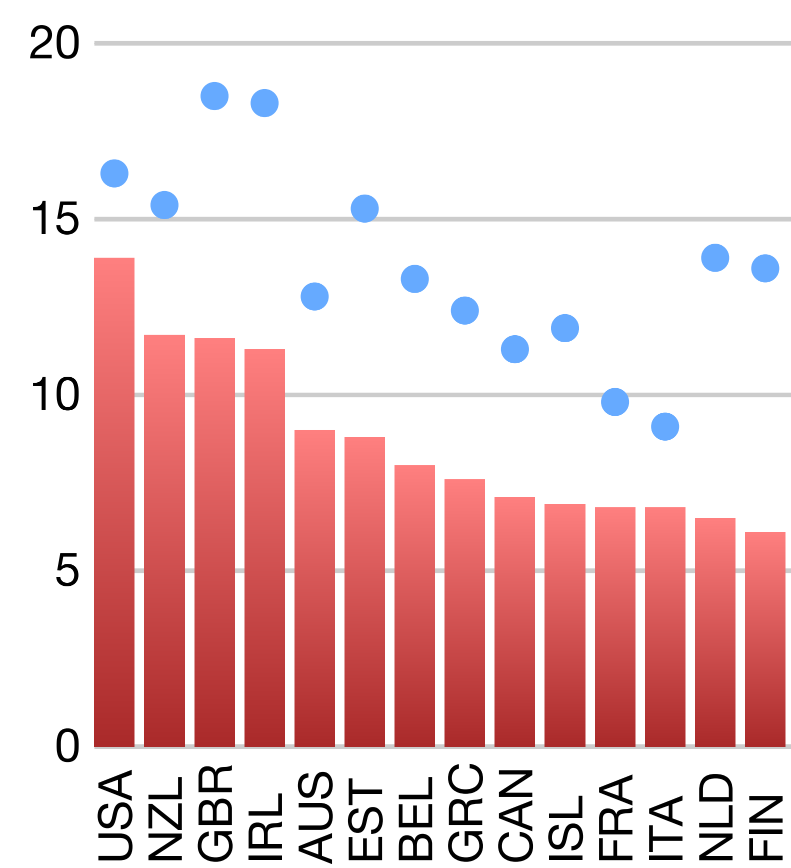
20
15
10
5
0
USA
NZL
GBR
IRL
AUS
EST
BEL
GRC
CAN
ISL
FRA
ITA
NLD
FIN

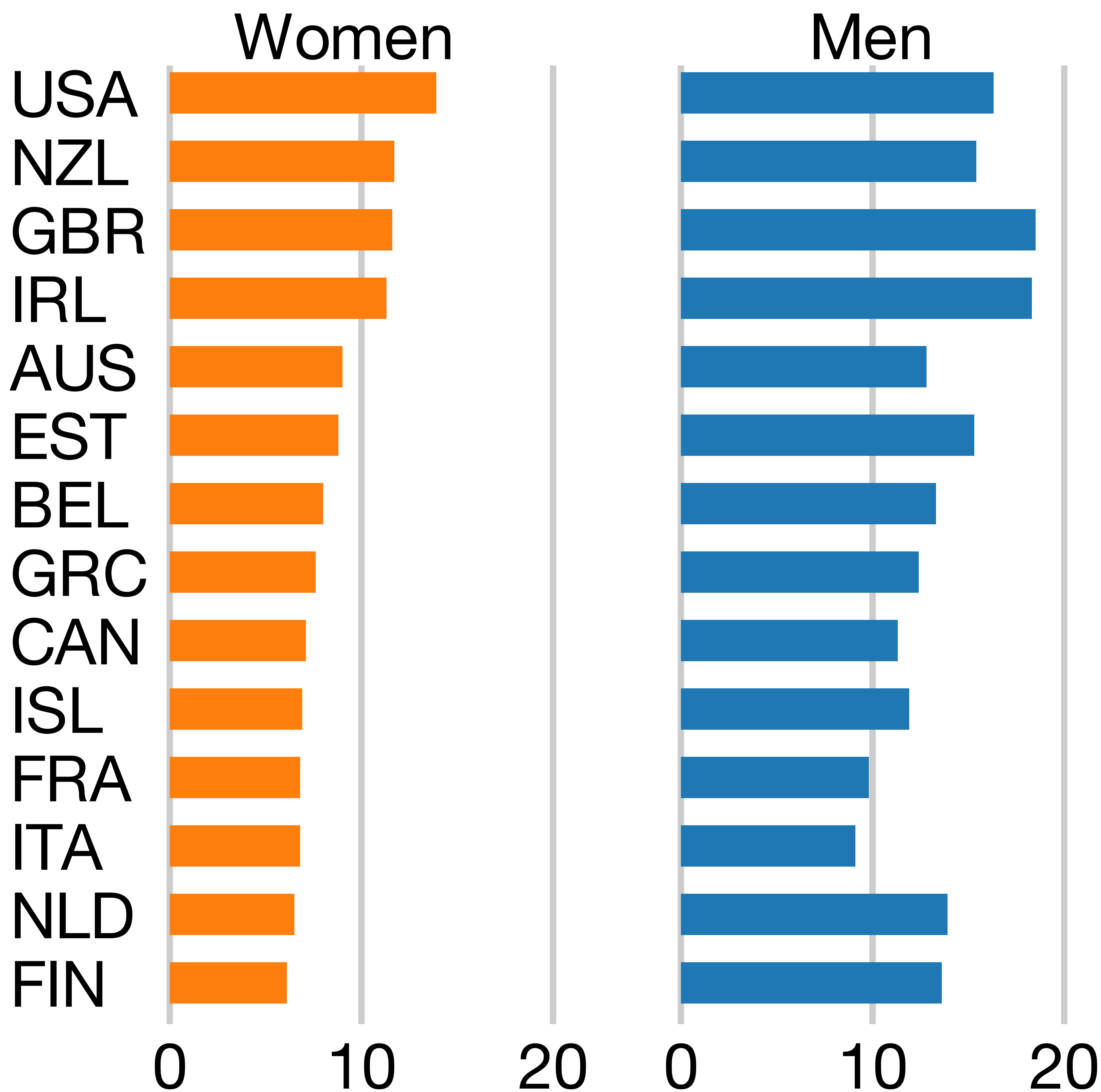
How do we use this?

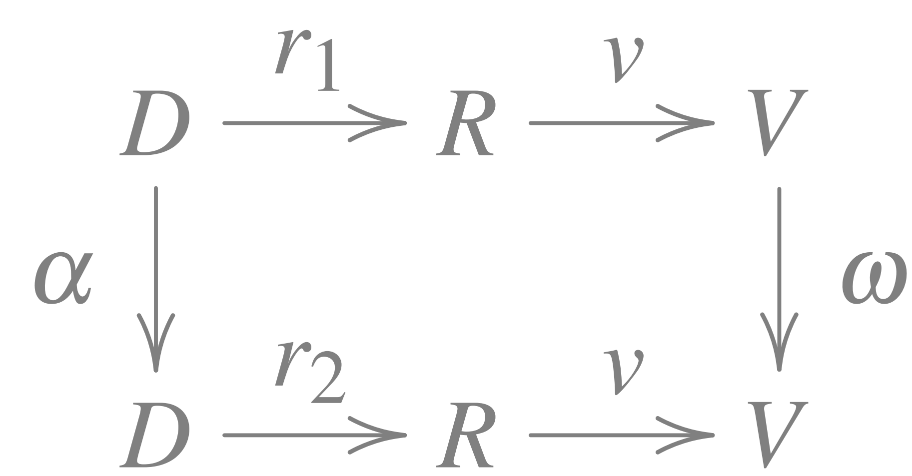
$$
\begin{array}{ccccc}
D & \xrightarrow{r_1} & R & \xrightarrow{\nu} & V \\
\alpha \downarrow & & & & \downarrow \omega \\
D & \xrightarrow{r_2} & R & \xrightarrow{\nu} & V
\end{array}
$$

1. α_1: What if the rate was different for just one gender?
   Either $x'_W = x_W + k$ and $x'_M = x_M$, or, $x'_M = x_M + k$ and $x'_W = x_W$.
2. α_2: What if the rates for men and women were switched?
   $x'_M = x_W$ and $x'_W = x_M$.
3. α_3: What if the gender gap in the rate was different?
   $x'_M = x_M + k$ and $x'_W = x_W - k$.
4. α_4: What if the overall rate was different (the same gender gap)?
   $x'_M = x_M + k$ and $x'_W = x_W + k$.

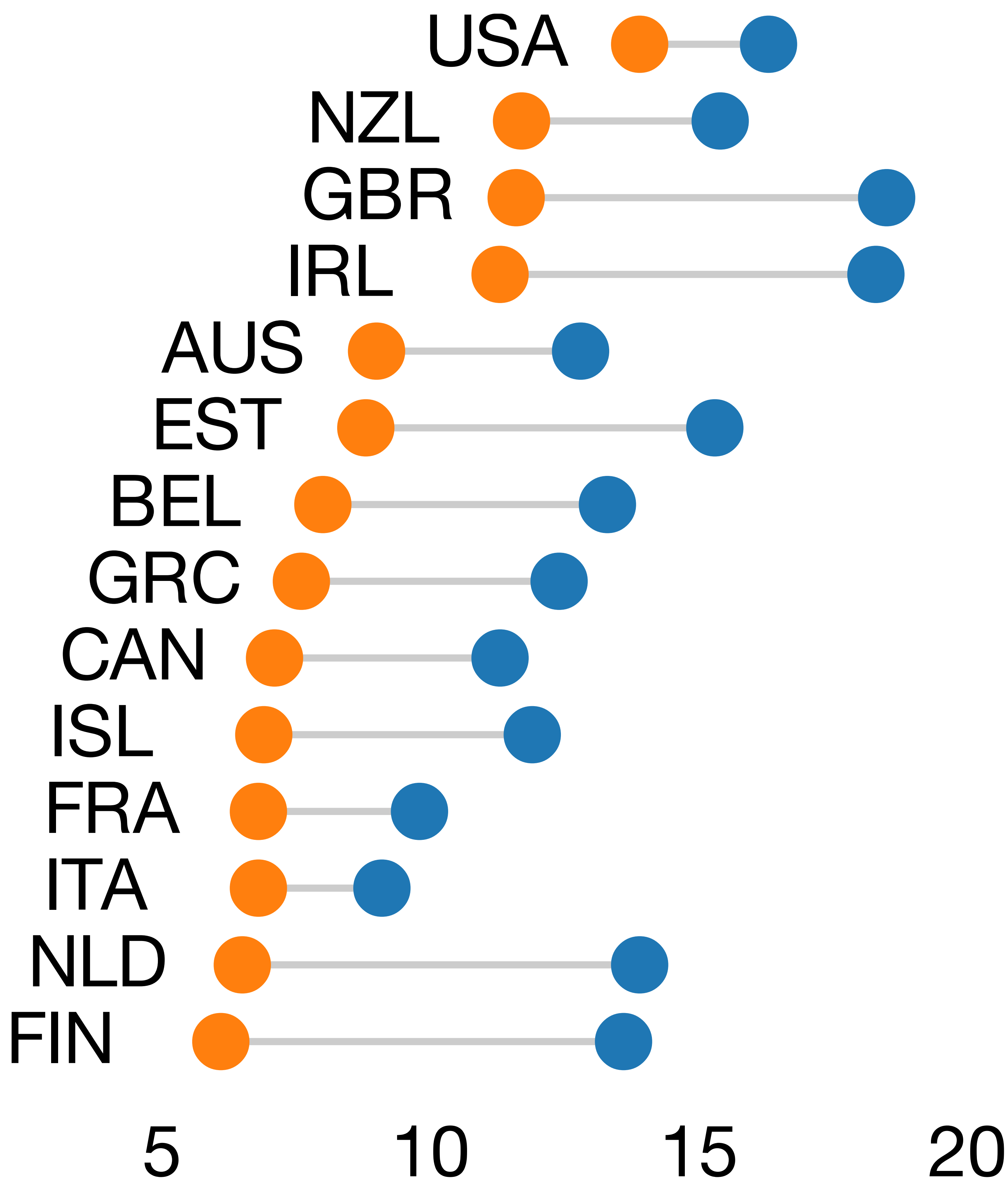


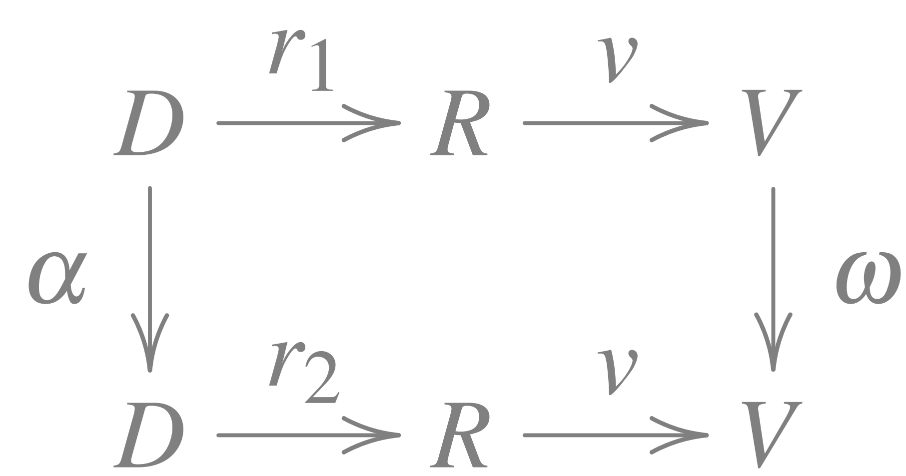




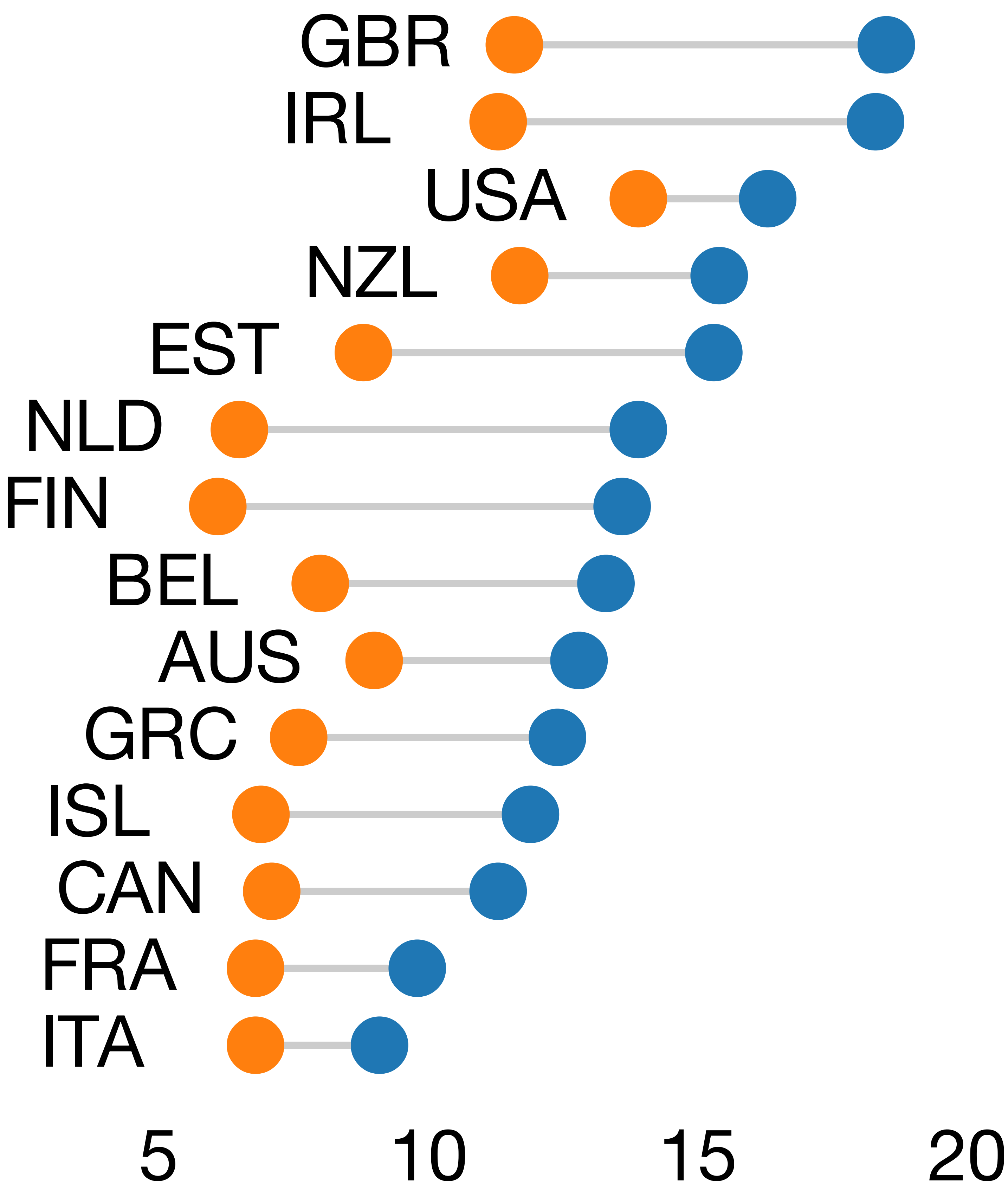


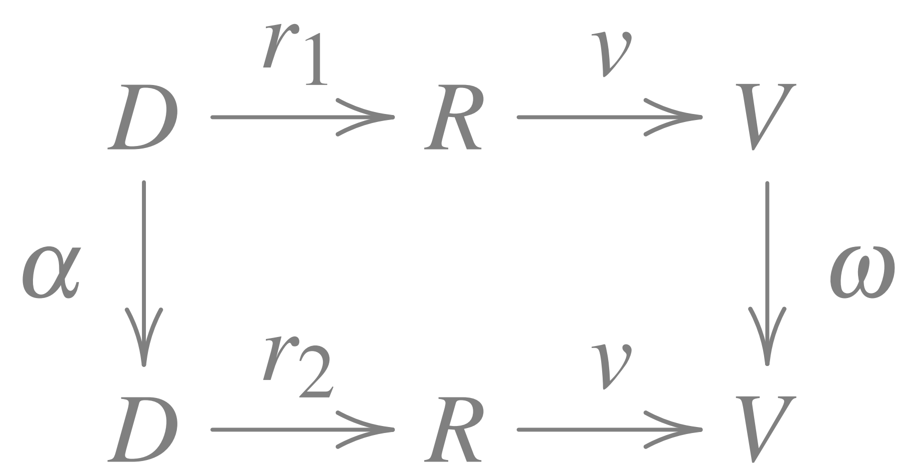
1. α_1: What if the rate was dif
   Either $x'_W = x_W + k$ and x'_M
2. α_2: What if the rates for me
   $x'_M = x_W$ and $x'_W = x_M$.
3. α_3: What if the gender gap
   $x'_M = x_M + k$ and $x'_W = x_W$ -
4. α_4: What if the overall rate
   $x'_M = x_M + k$ and $x'_W = x_W$ -



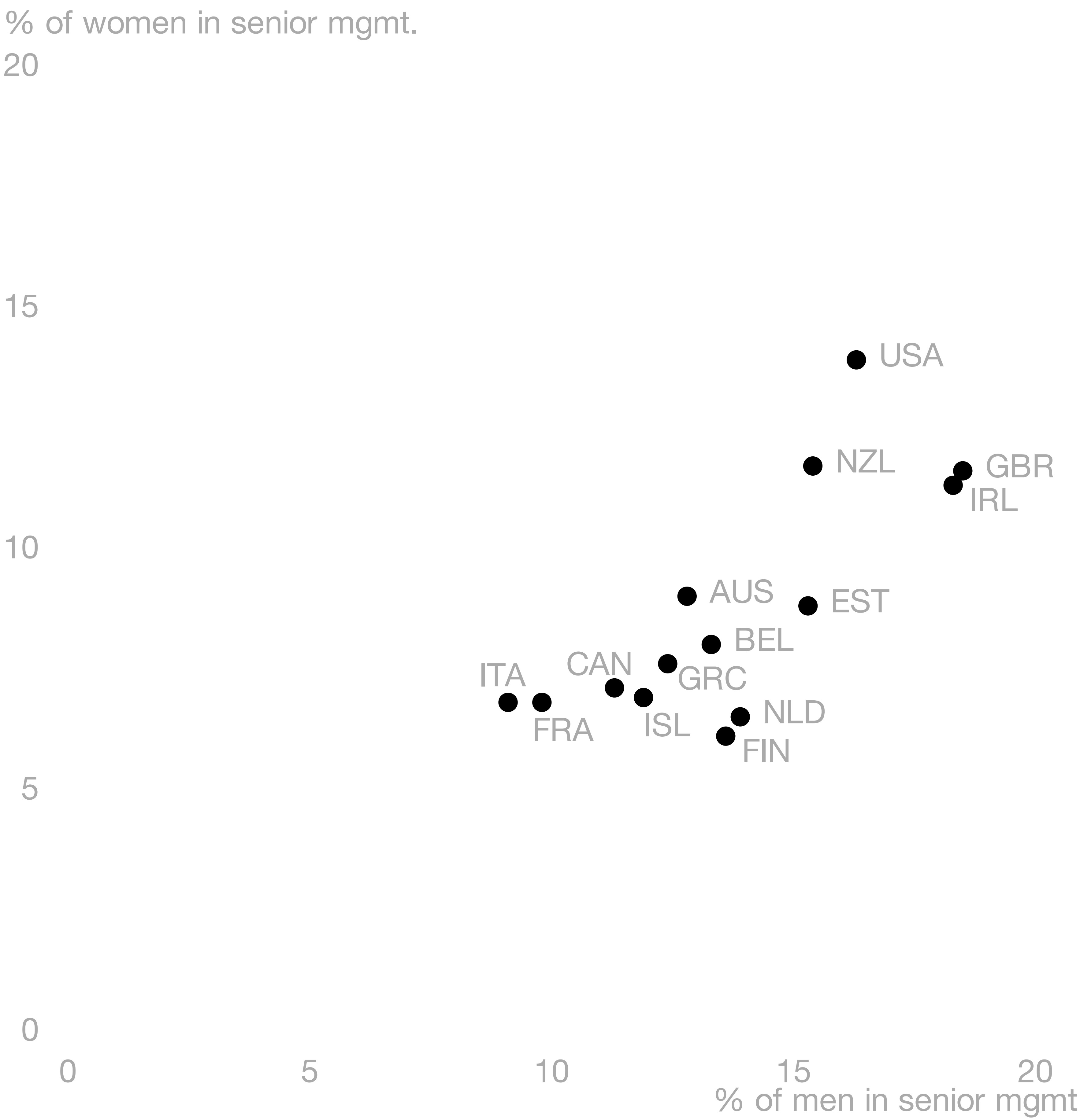


1. α_1: What if the rate was dif
   Either $x'_W = x_W + k$ and x'_M
2. α_2: What if the rates for me
   $x'_M = x_W$ and $x'_W = x_M$.
3. α_3: What if the gender gap
   $x'_M = x_M + k$ and $x'_W = x_W$ -
4. α_4: What if the overall rate
   $x'_M = x_M + k$ and $x'_W = x_W$ -



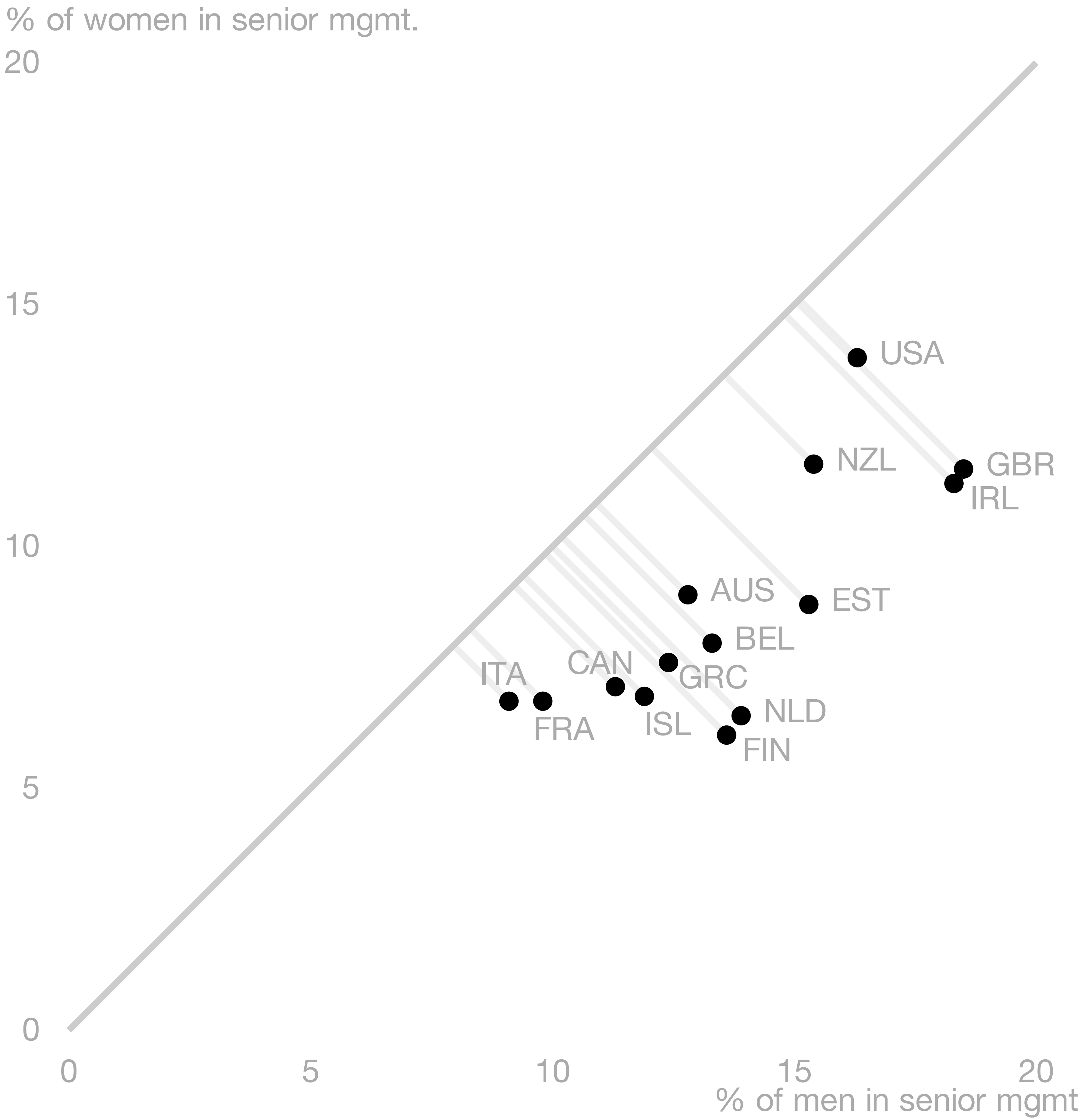


1. α_1: What if the rate was dif
   Either $x'_W = x_W + k$ and x'_M
2. α_2: What if the rates for me
   $x'_M = x_W$ and $x'_W = x_M$.
3. α_3: What if the gender gap
   $x'_M = x_M + k$ and $x'_W = x_W$ -
4. α_4: What if the overall rate
   $x'_M = x_M + k$ and $x'_W = x_W$ -



$$\begin{array}{ccccc} D & \xrightarrow{r_1} & R & \xrightarrow{v} & V \\ \alpha \downarrow & & & & \downarrow \omega \\ D & \xrightarrow{r_2} & R & \xrightarrow{v} & V \end{array}$$

1. α_1: What if the rate was dif
   Either $x'_W = x_W + k$ and x'_M
2. α_2: What if the rates for me
   $x'_M = x_W$ and $x'_W = x_M$.
3. α_3: What if the gender gap
   $x'_M = x_M + k$ and $x'_W = x_W$ -
4. α_4: What if the overall rate
   $x'_M = x_M + k$ and $x'_W = x_W$ -



$$\begin{array}{ccccc} D & \xrightarrow{r_1} & R & \xrightarrow{v} & V \\ \alpha \downarrow & & & & \downarrow \omega \\ D & \xrightarrow{r_2} & R & \xrightarrow{v} & V \end{array}$$

1. α_1: What if the rate was dif
   Either $x'_W = x_W + k$ and x'_M
2. α_2: What if the rates for me
   $x'_M = x_W$ and $x'_W = x_M$.
3. α_3: What if the gender gap
   $x'_M = x_M + k$ and $x'_W = x_W$ -
4. α_4: What if the overall rate
   $x'_M = x_M + k$ and $x'_W = x_W$ -

Summary

- To evaluate a visualization:
  - take one instance of the data being visualized, and **think about how the input could have been different**
  - **What this would do to the vis?** Is this a good channel? Is it separable?
  - Conversely, think of the good channels: position, length, luminance - **do changes of these attributes correspond to sensible changes in the data?**