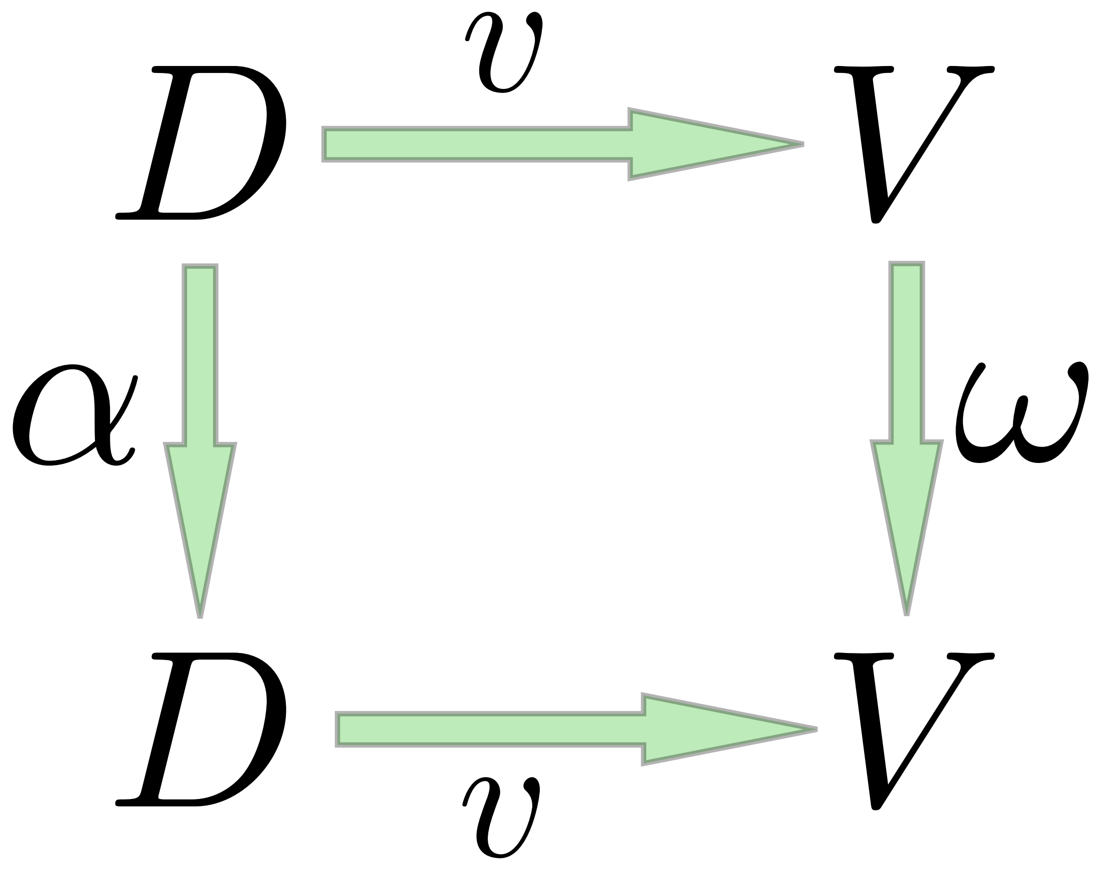
Introduction to an Algebraic Process for Visualization Design

http://algebraicvis.net

Gordon Kindlmann
Computer Science & Computation Institute
University of Chicago
glk@uchicago.edu
@glk1

Carlos Scheidegger
Computer Science
University of Arizona
cscheid@cscheid.net
@scheidegger



31 July 2016, JSM2016

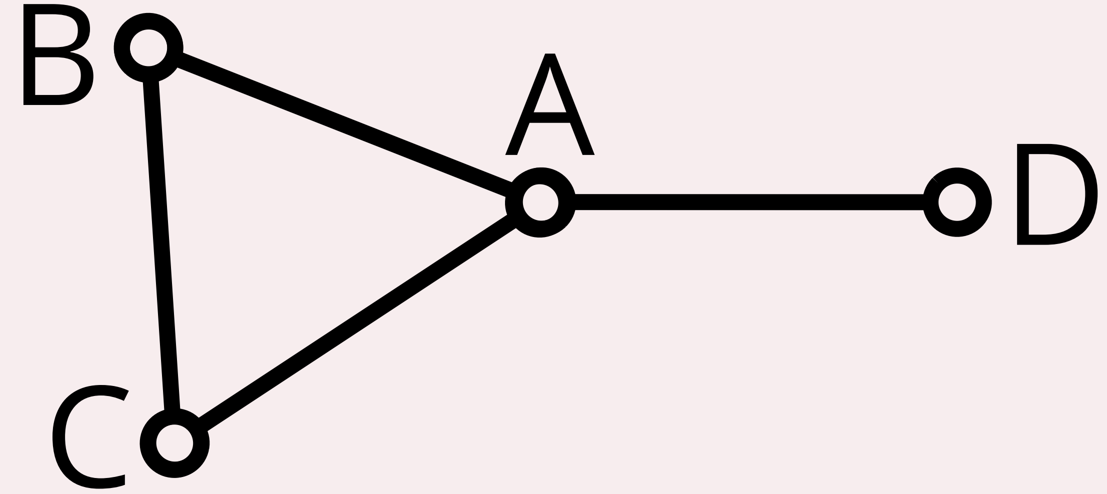
The basic mapping of visualization

Data → Visual

1) How to use 2 planar dimensions? (layout, arrange)
2) What to draw at each location? (encode)

How will these be perceived by the viewer?

example:
(a particular graph on 4 vertices)



Vis methods use **computational** representation

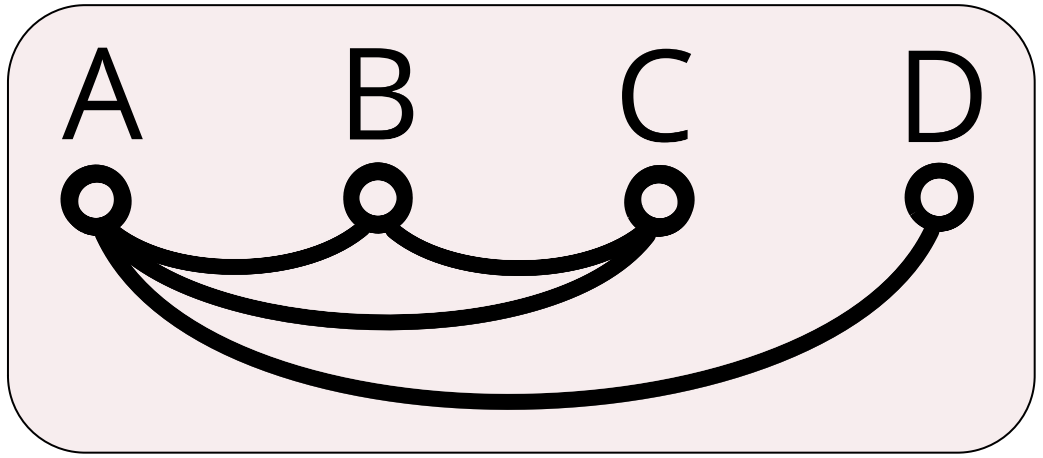
"Data" → Representation → Visual

Underlying thing of interest → How we can measure or store it on computer

"Show data variation, not design variation" [Tufte 1983]

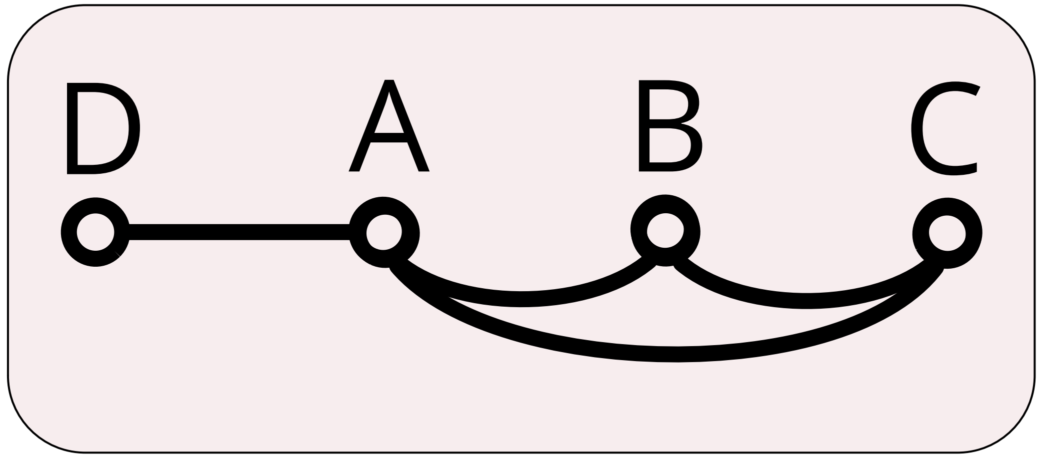
(a graph)

V=(**A**,**B**,**C**,**D**);
E=(**A-B**,**B-C**,**A-C**,**A-D**)



not equal: bug?

V=(**D**,**A**,**B**,**C**);
E=(**A-B**,**B-C**,**A-C**,**A-D**)



Basic ideas of Algebraic Vis Design

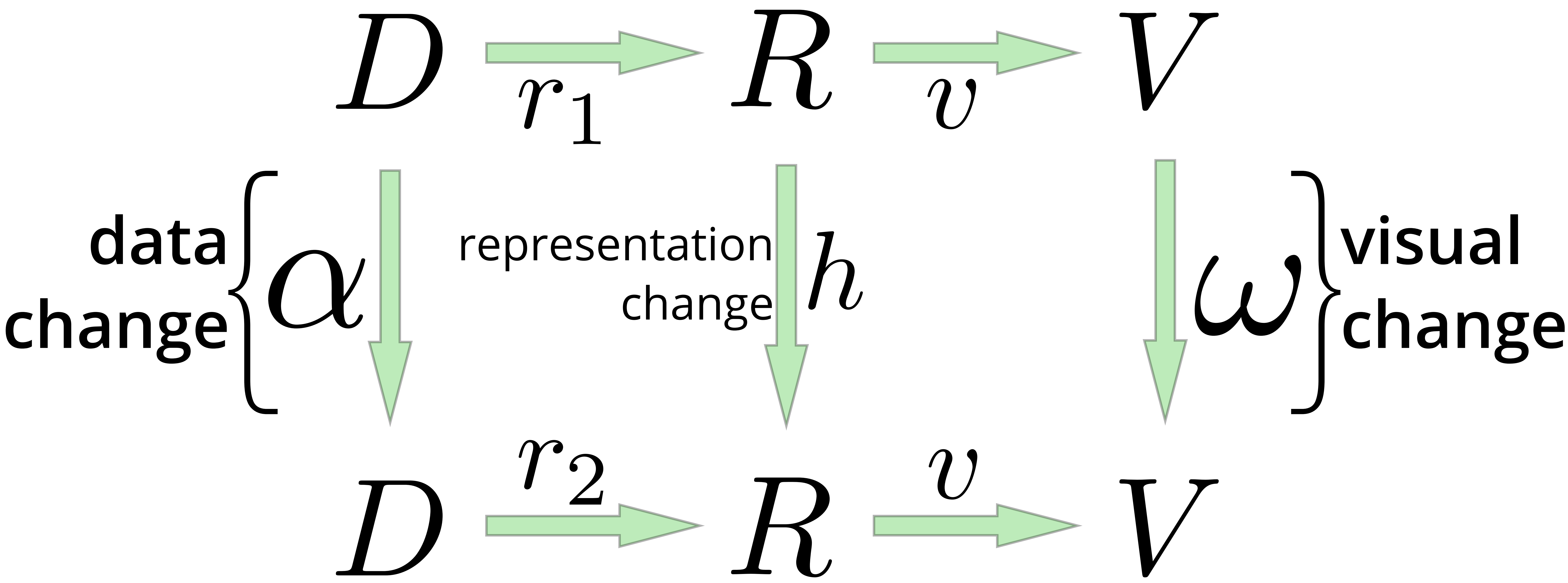
[Kindlmann & Scheidegger 2014]

Are important data changes well-matched with obvious visual changes?

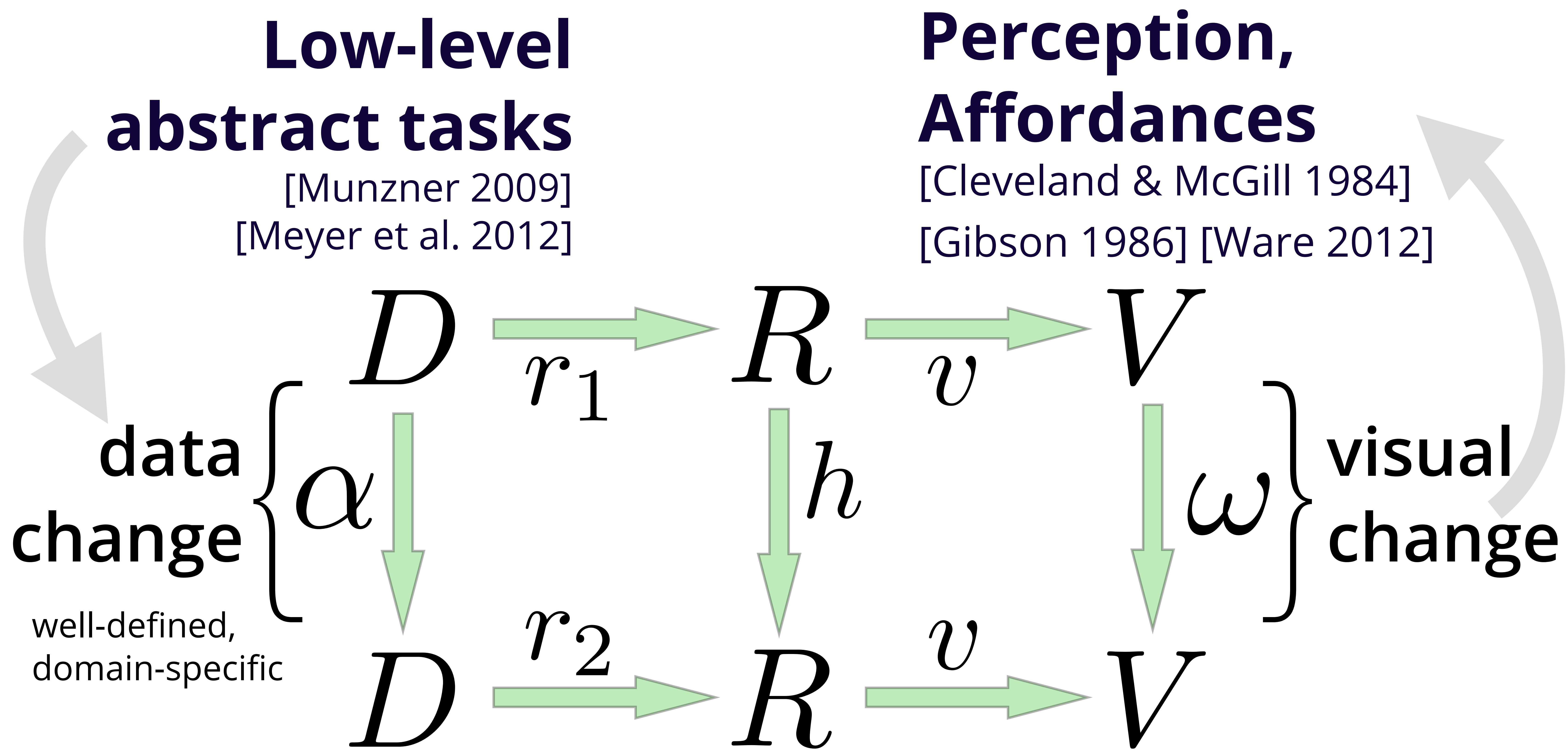
Not a taxonomy of tasks, data types, etc

Mathematical vocabulary for describing how a visualization does or doesn’t work

Underlying commutative diagram



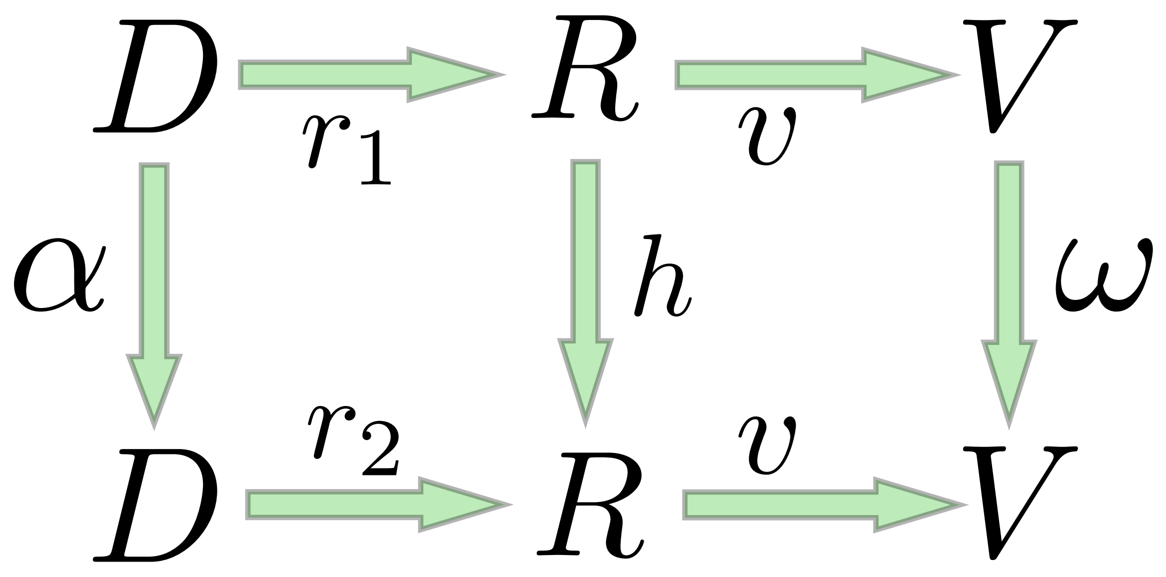
Design goal: Task → α, ω → affordance



Three Algebraic Design Principles

All derived from one diagram

Tools, not Rules



Does ω **make sense**, given α ?

→ 1. Principle of Visual-Data Correspondence

For all important α, is ω obvious?

→ 2. Principle of Unambiguous Data Depiction

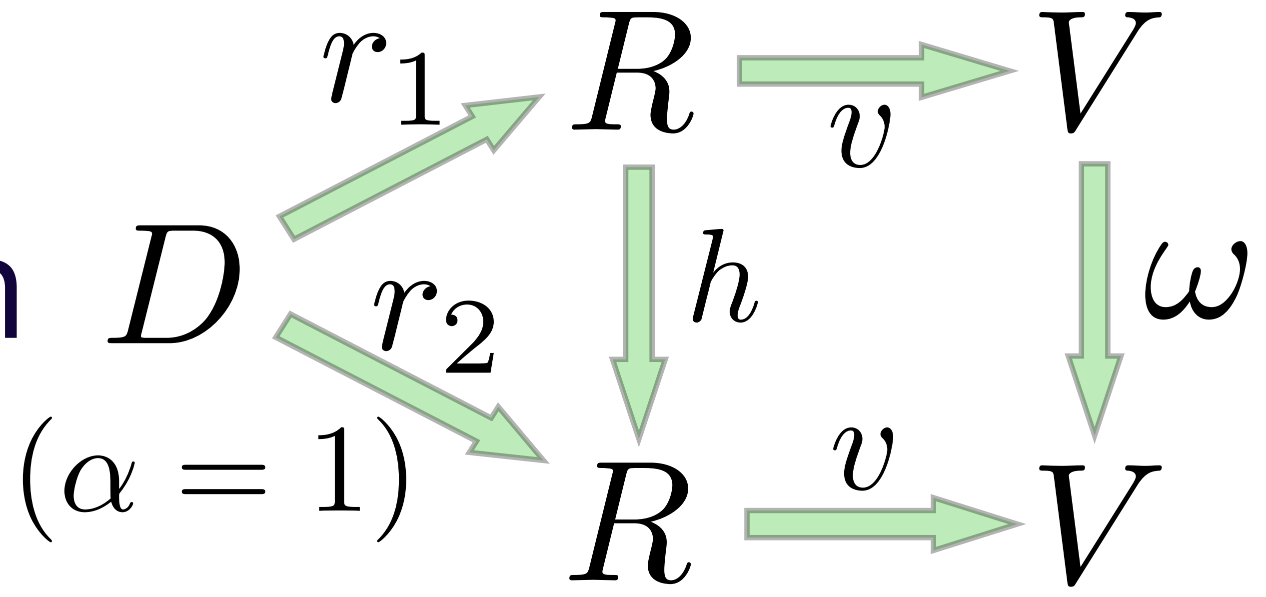
Can obvious ω arise without data change (α=1)?

→ 3. Principle of Representation Invariance

3. Principle of Representation Invariance

Visualization is invariant w.r.t changes in data representation

If α=1, then ω=1.



- Underlying **data** D ≠ **representation** R of data
  - e.g. sets as lists, eigenvectors as vectors
- **Invariantive**: Scale of measurement (nominal, ordinal, interval, ratio) limits permissible statistics [Stevens 1946]
- If change h in representation is visible (ω≠1), h is the "**hallucinator**"

Representation Invariance is old idea

[Stevens 1946]

SCIENCE

Vol. 103, No. 2684 Friday, June 7, 1946

On the Theory of Scales of Measurement

S. S. Stevens

Director, Psycho-Acoustic Laboratory, Harvard University

FOR SEVEN YEARS A COMMITTEE of the British Association for the Advancement of Science debated the problem of measurement. Appointed in 1932 to represent Section A (Mathematical and Physical Sciences) and Section J (Psychology), the committee was instructed to consider and report upon the possibility of "quantitative estimates of sensory events"—meaning simply: Is it possible to measure human sensation? Deliberation led only to disagreement, mainly about what is meant by the term measurement. An interim report in 1938 found one member complaining that his colleagues by the formal (mathematical) properties of the scales. Furthermore—and this is of great concern to several of the sciences—the statistical manipulations that can legitimately be applied to empirical data depend upon the type of scale against which the data are ordered.

A Classification of Scales of Measurement

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads

Representation Invariance is old idea

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
NOMINAL	Determination of equality	*Permutation group* $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode Contingency correlation
ORDINAL	Determination of greater or less	*Isotonic group* $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	*General linear group* $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	*Similarity group* $x' = ax$	Coefficient of variation

possible hallucinators:

e.g. taking median commutes w/ applying a monotonic function; taking the mean does not

Invariance example: Graph layout

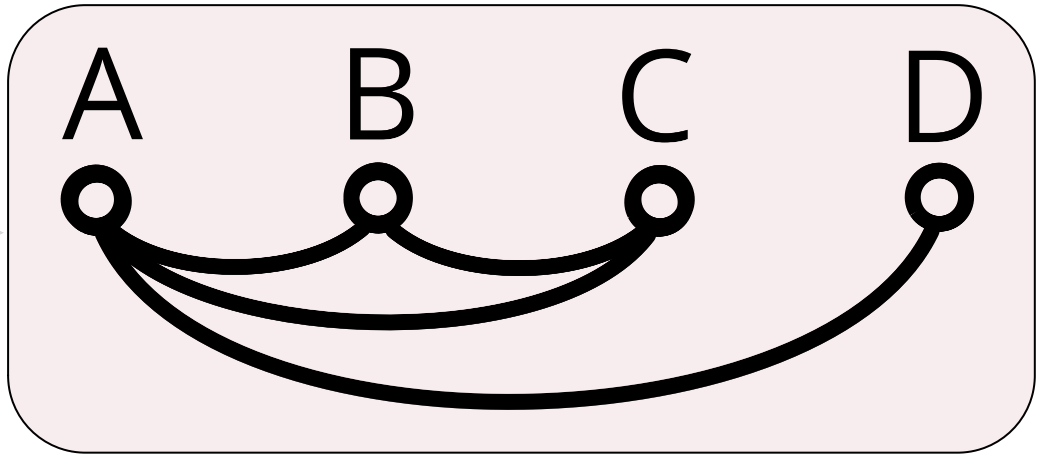
Representation: **lists**
of verts, edges

Data: a
graph on
4 vertices

D

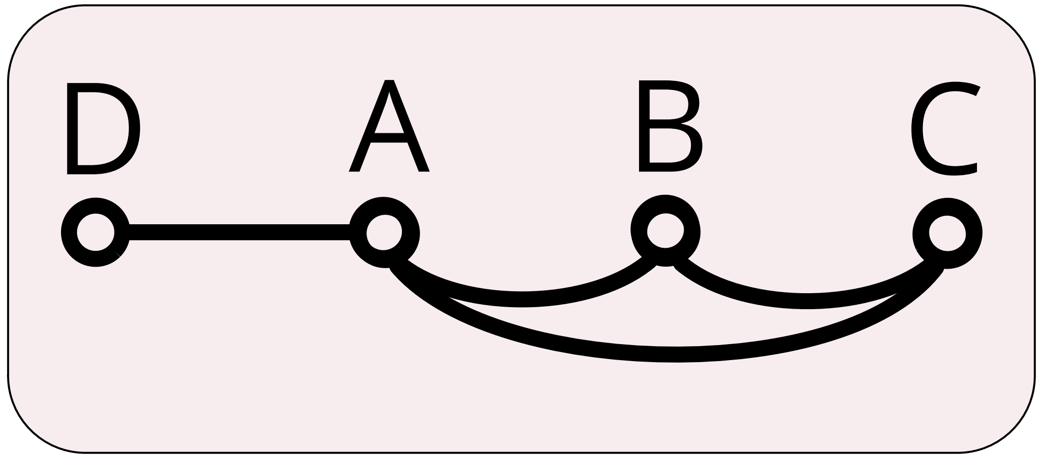
(α=1)

R $\xrightarrow{v}$



h: permute
vert list

R $\xrightarrow{v}$



ω≠1: layout
depends on
vertex
ordering

Invariance example: Graph layout

Representation: **lists** of verts, edges

Data: a graph on 4 vertices

R

v

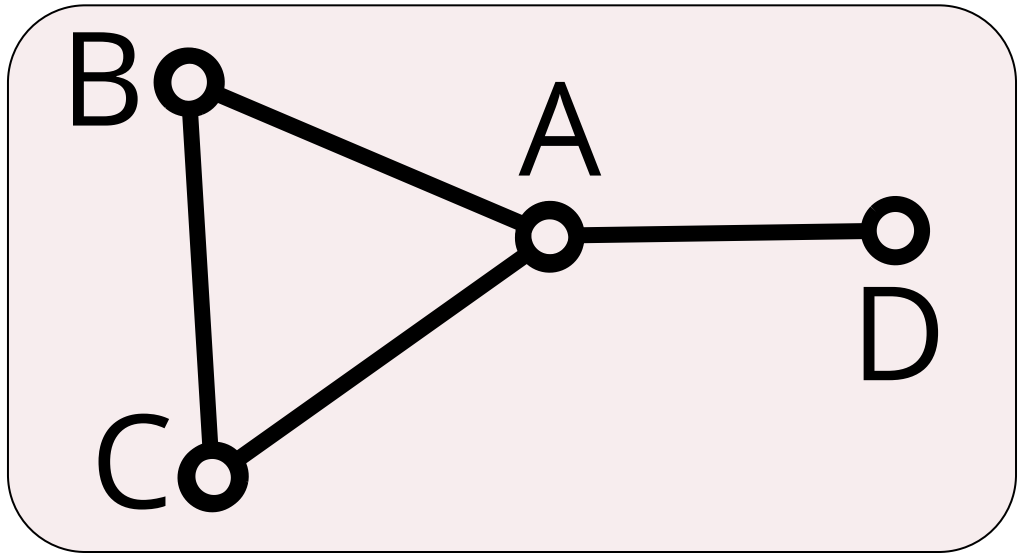
h: permute vert list

D

(α=1)

v

R



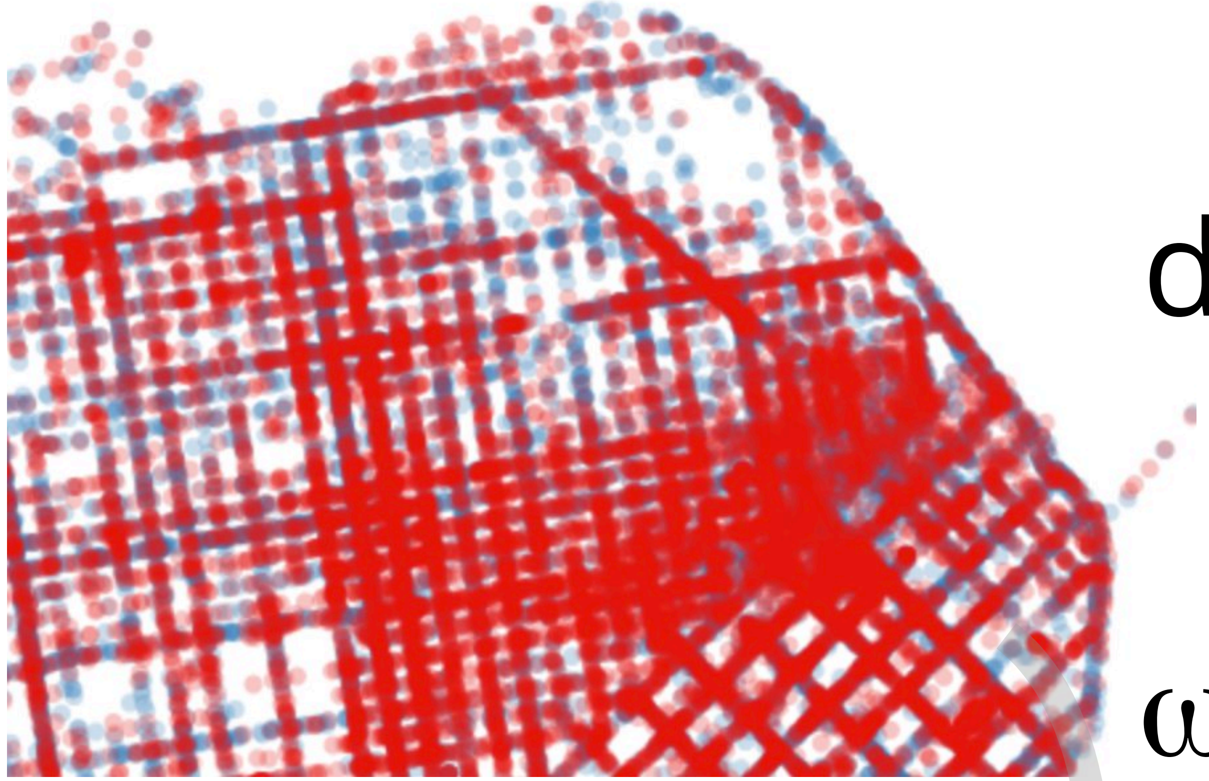
ω**=1**: with **order-invariant** layout

Invariance example: alpha-blended marks

Data: **set** of locations of taxi pickups & drop-offs

Representation: **list** of locations

$R \xrightarrow{v}$

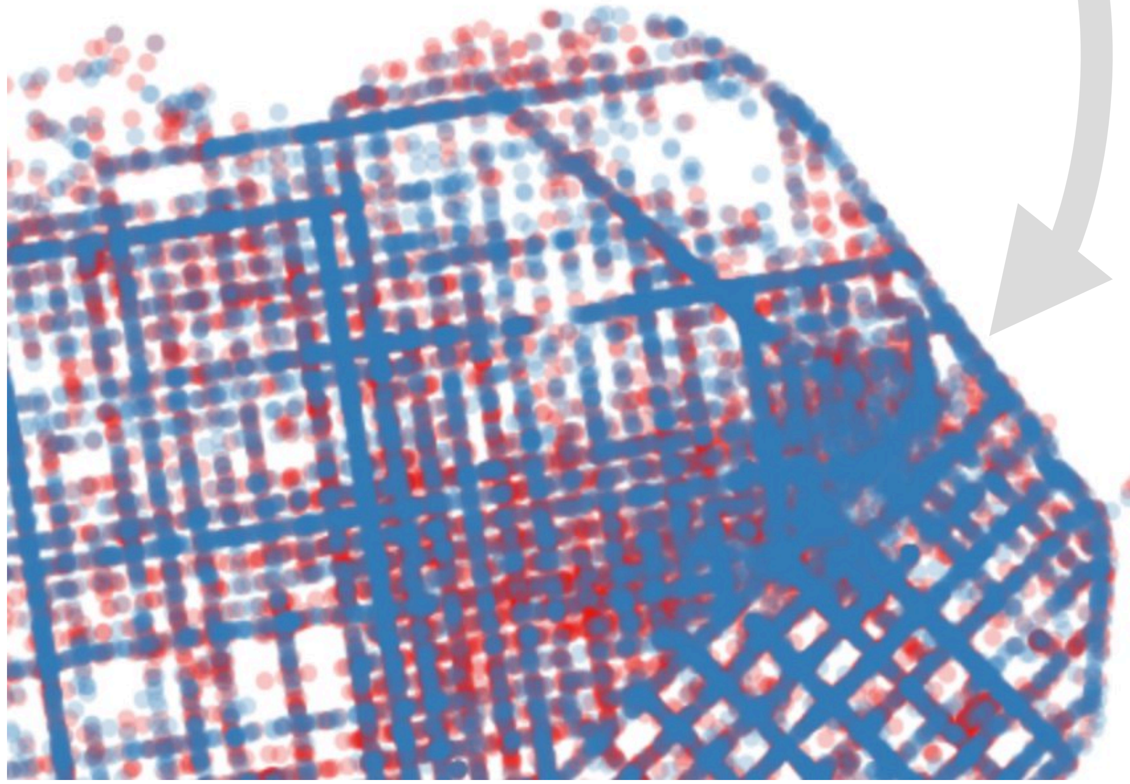


pick-up
drop-off

D (α=1)

h: permute list

$R \xrightarrow{v}$

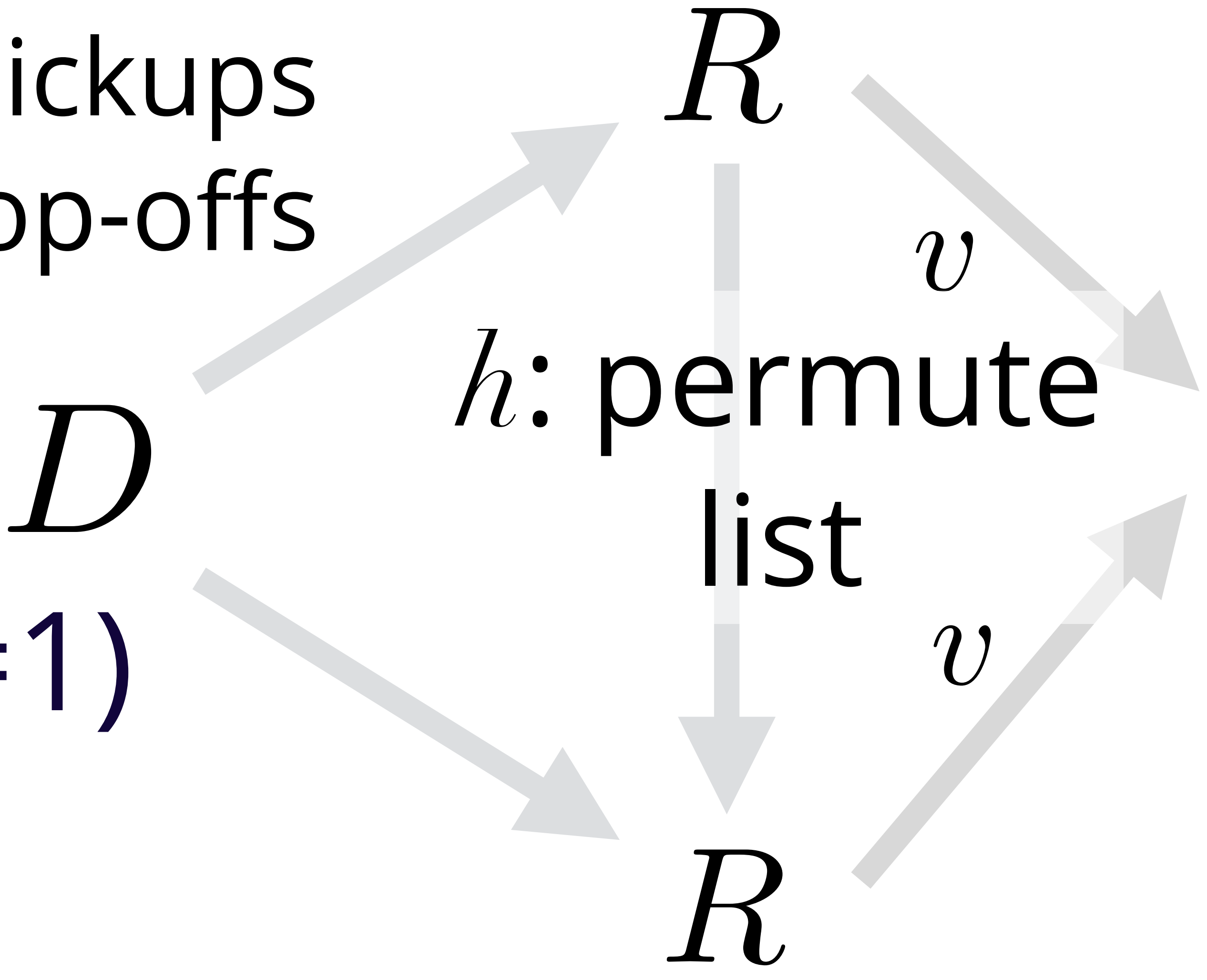


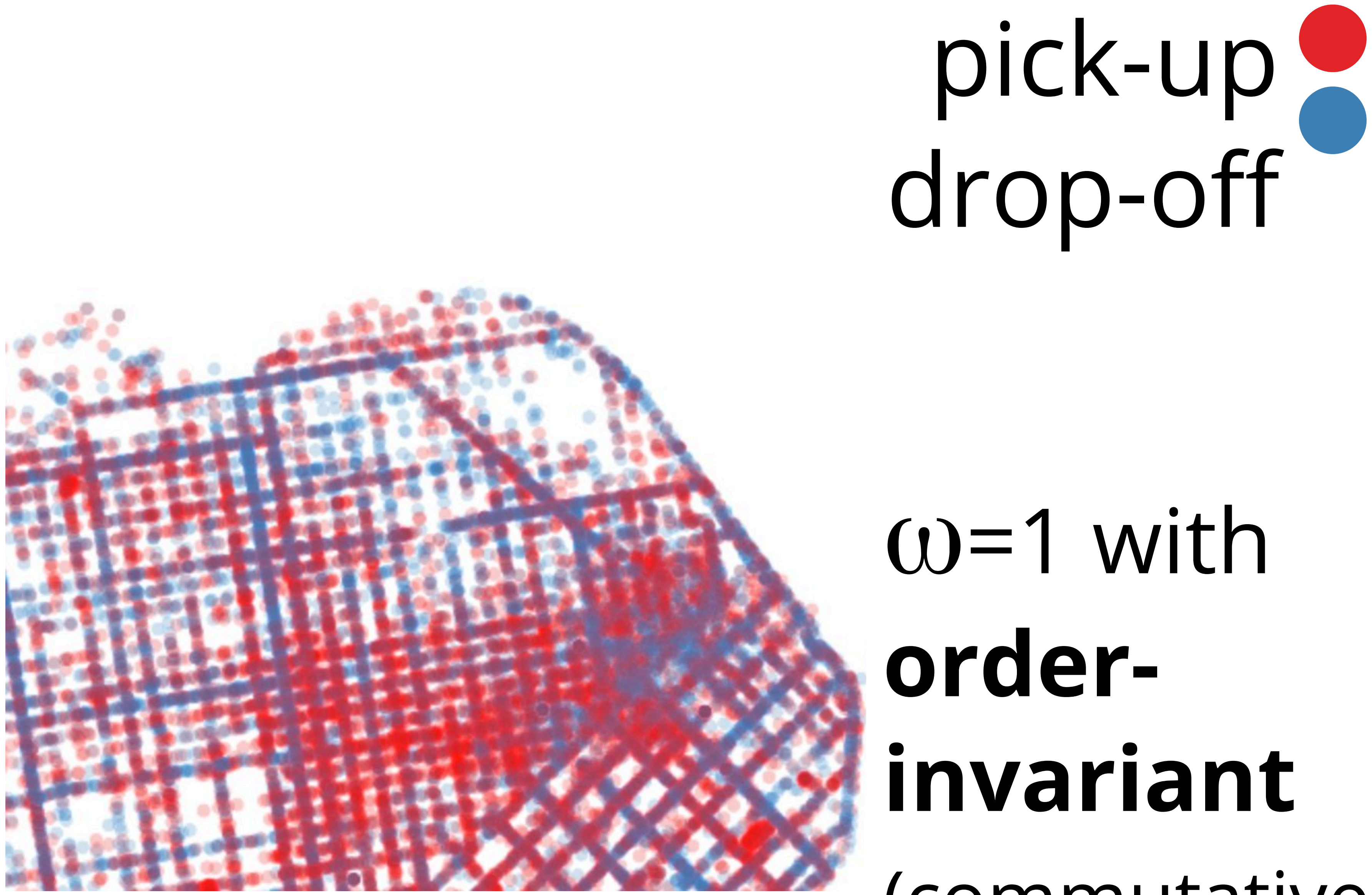
ω≠1 : "over" operator does not commute: permutation h is a **hallucinator**

Invariance example: alpha-blended marks

Data: **set** of locations of taxi pickups & drop-offs

Representation: **list** of locations



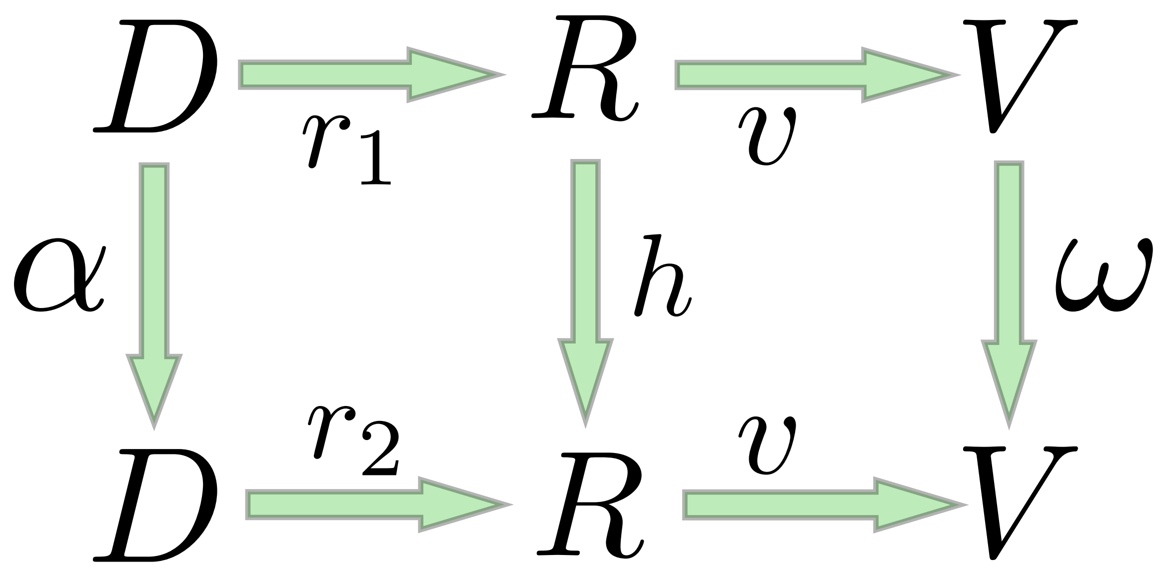


ω=1 with **order-invariant** (commutative) compositing

Three Algebraic Design Principles

All derived from one diagram

Tools, not Rules



Does ω **make sense**, given α ?

→ 1. Principle of Visual-Data Correspondence

For all important α, is ω obvious?

→ 2. Principle of Unambiguous Data Depiction

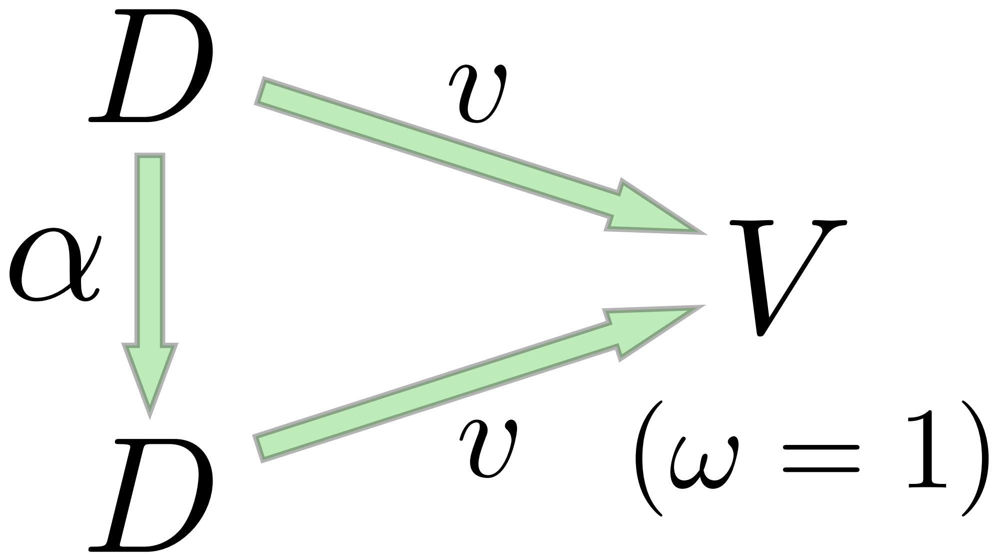
Can obvious ω arise without data change (α=1)?

→ 3. Principle of Representation Invariance

2. Principle of Unambiguous Data Depiction

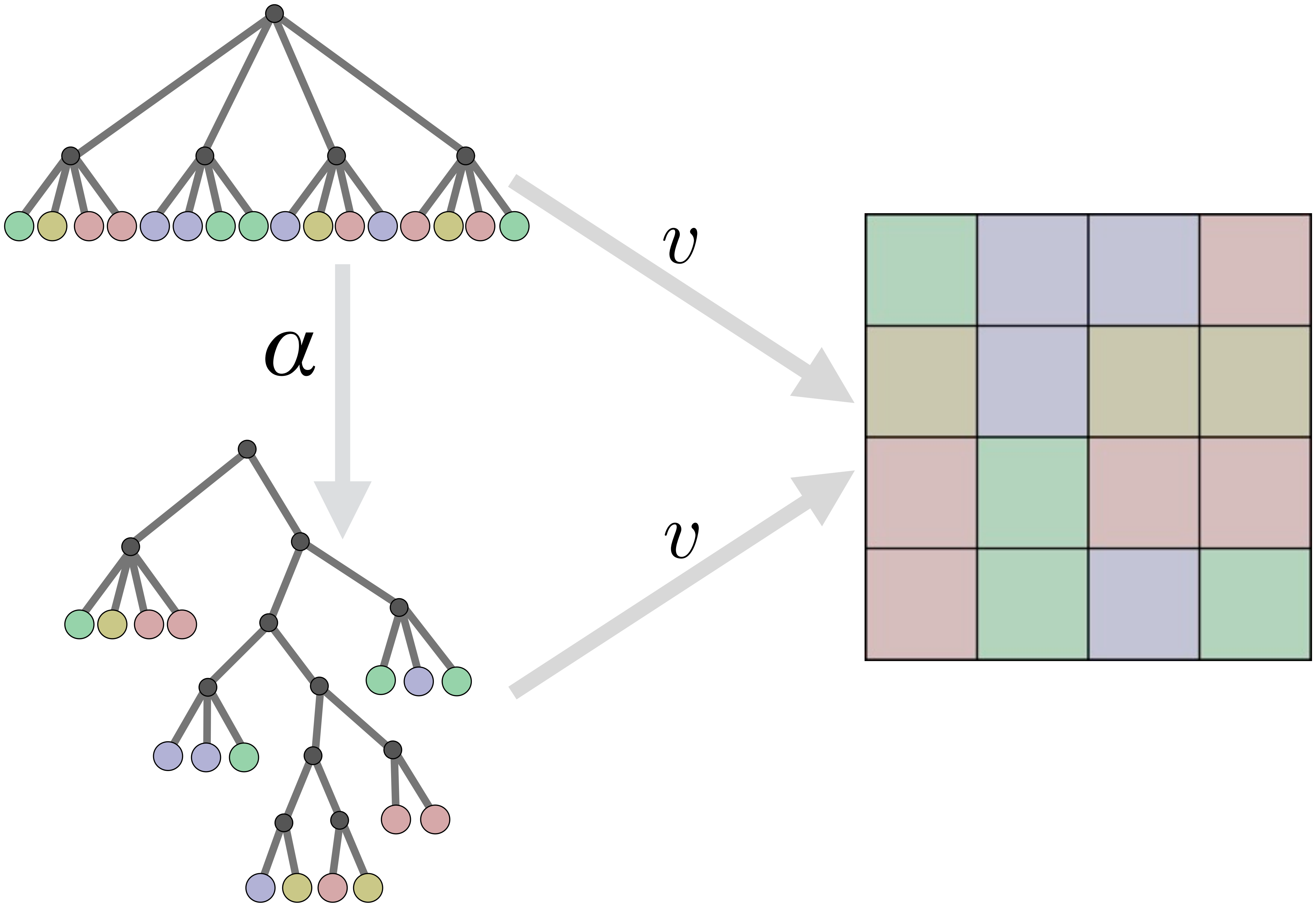
Important α map to obvious ω.

If ω=1, then α=1.



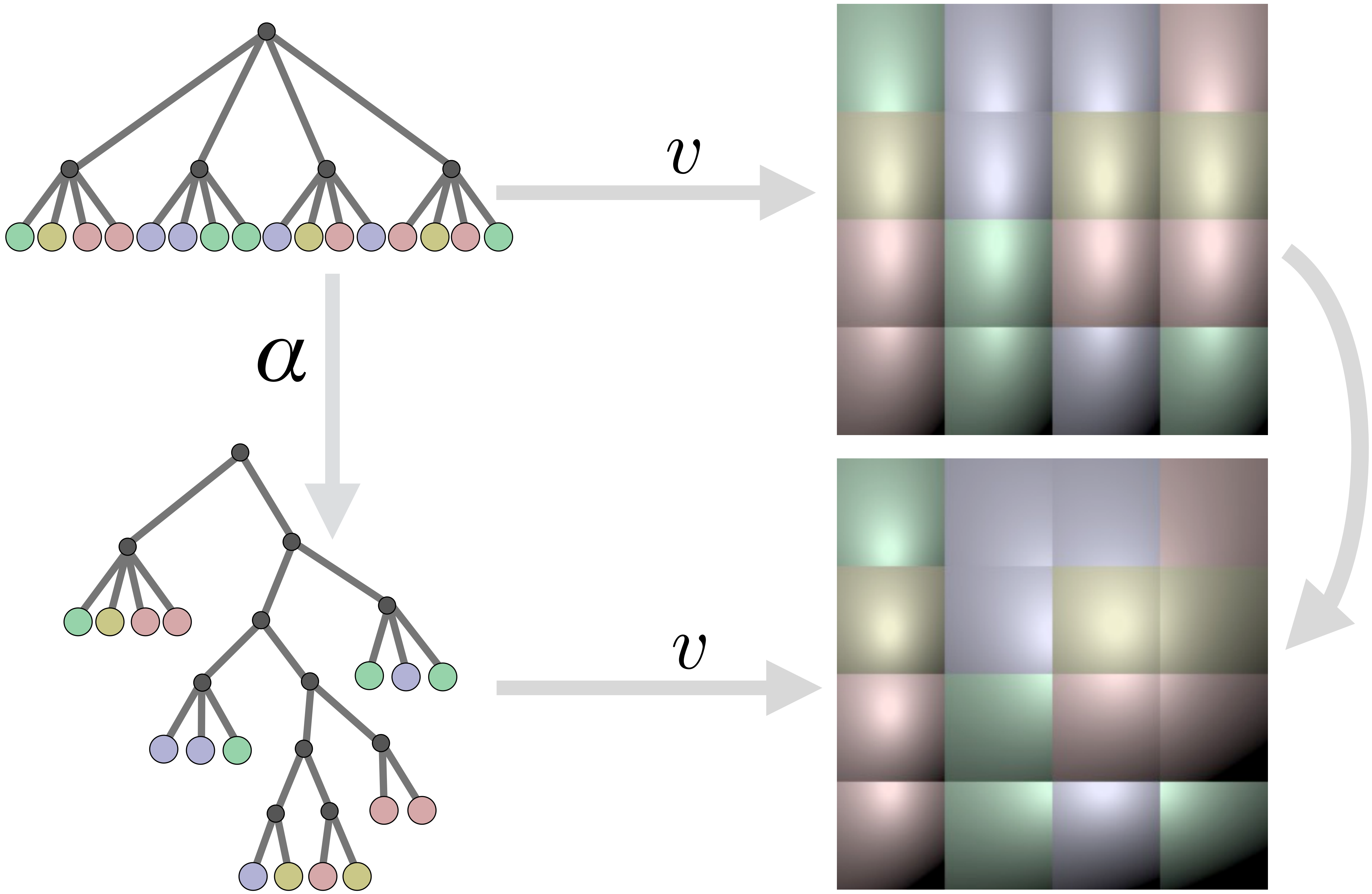
- **Expressiveness**: visualization shows all facts about data (and nothing more) [Mackinlay 1986]
- **Injectivity**: visualization preserves distinctness so viewer can invert it (read it) [Ziemkiewicz & Kosara 2009]
- If not v injective, α explicitly indicates the ambiguity; α is the "**confuser**"

Unambiguity example: treemaps



ω=1: α is a **“confuser”** for this treemap

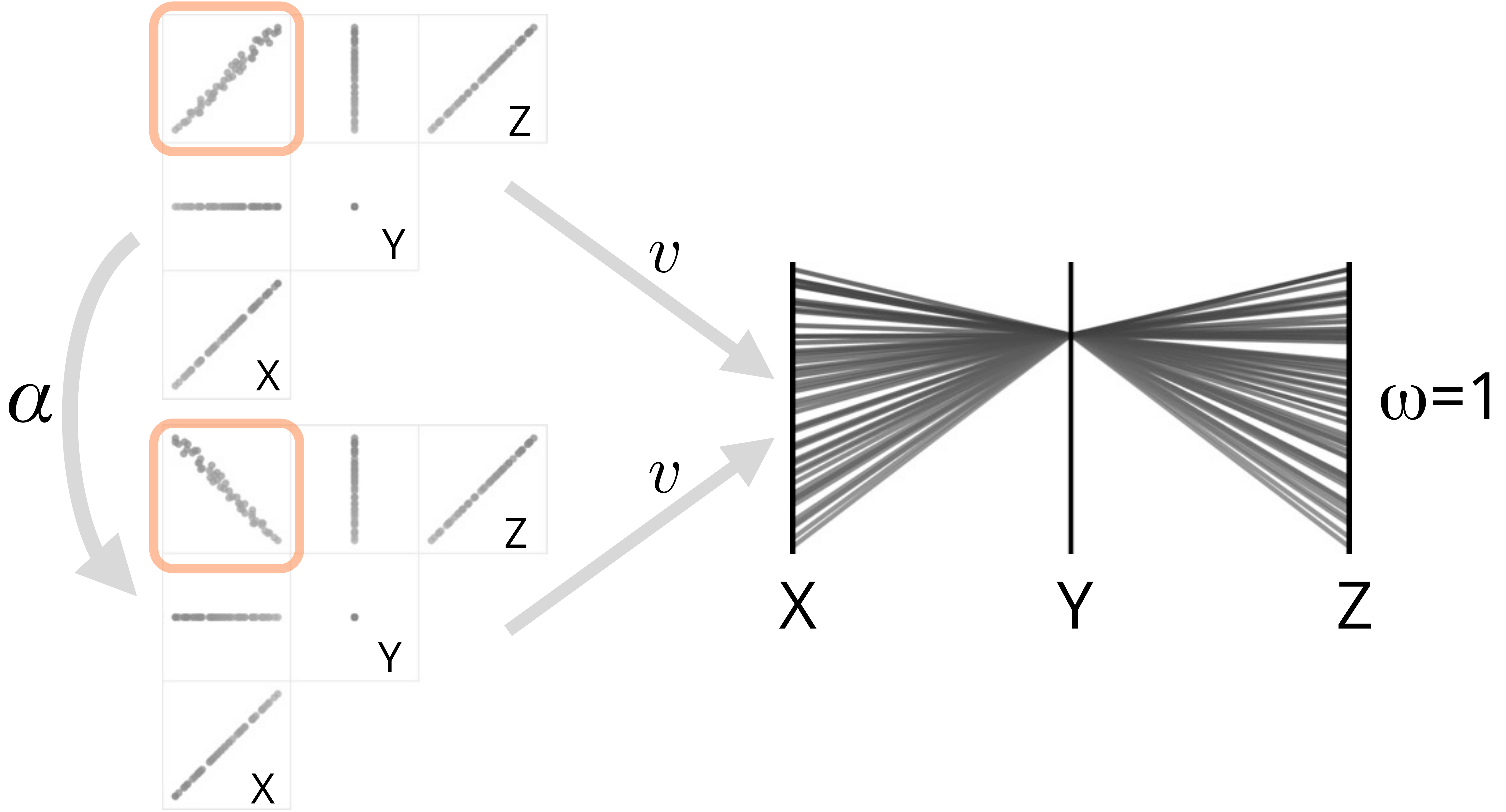
Unambiguity example: treemaps



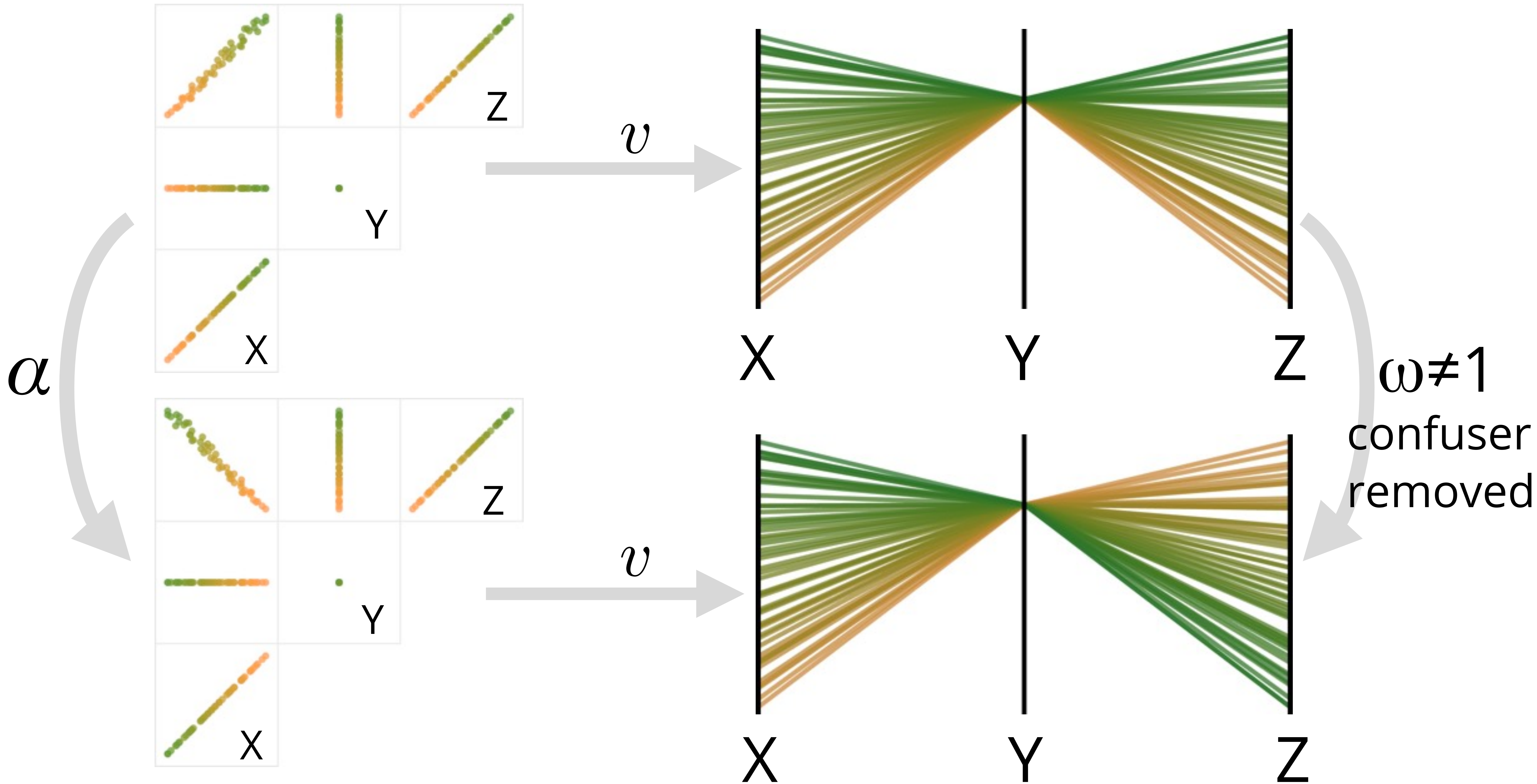
ω≠1: cushion treemaps removes **confuser**

[van Wijk & H. van de Wetering 1999]

Unambiguity example: parallel coordinates



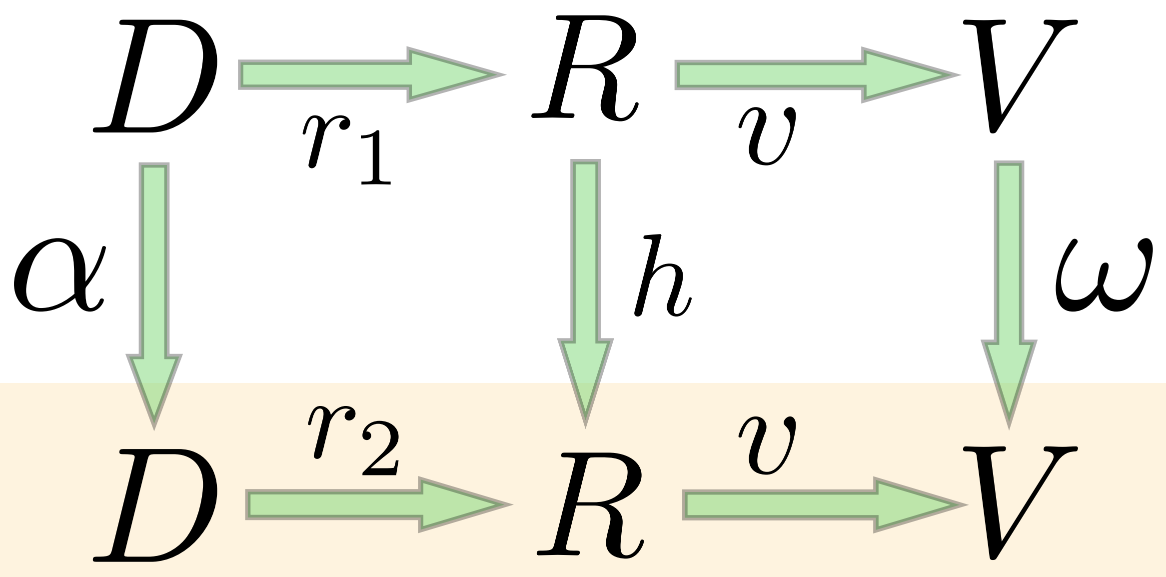
Unambiguity example: parallel coordinates



Three Algebraic Design Principles

All derived from one diagram

Tools, not Rules



Does ω **make sense**, given α ?

→ 1. Principle of Visual-Data Correspondence

For all important α, is ω obvious?

→ 2. Principle of Unambiguous Data Depiction

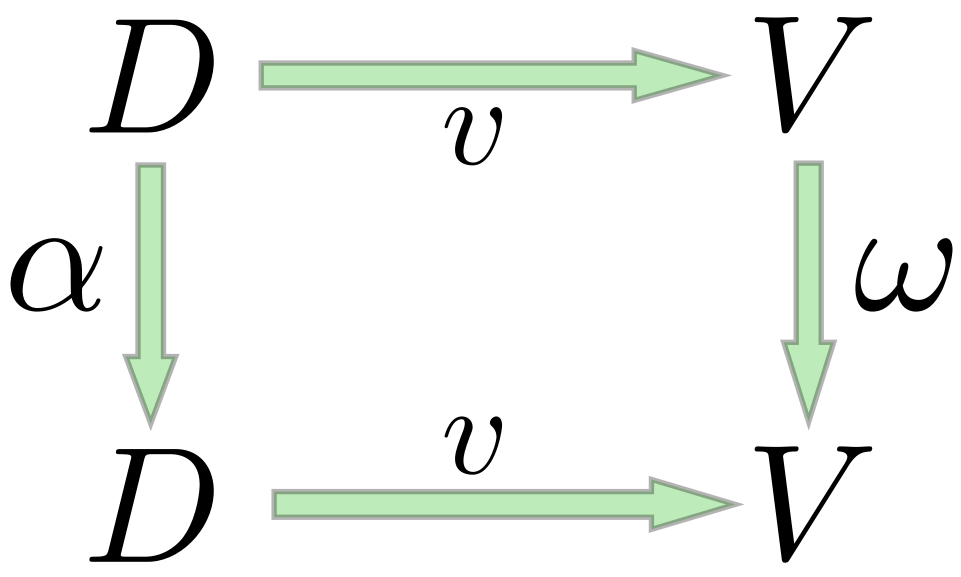
Can obvious ω arise without data change (α=1)?

→ 3. Principle of Representation Invariance

1. Principle of Visual-Data Correspondence

Important α produce obvious and meaningful ω

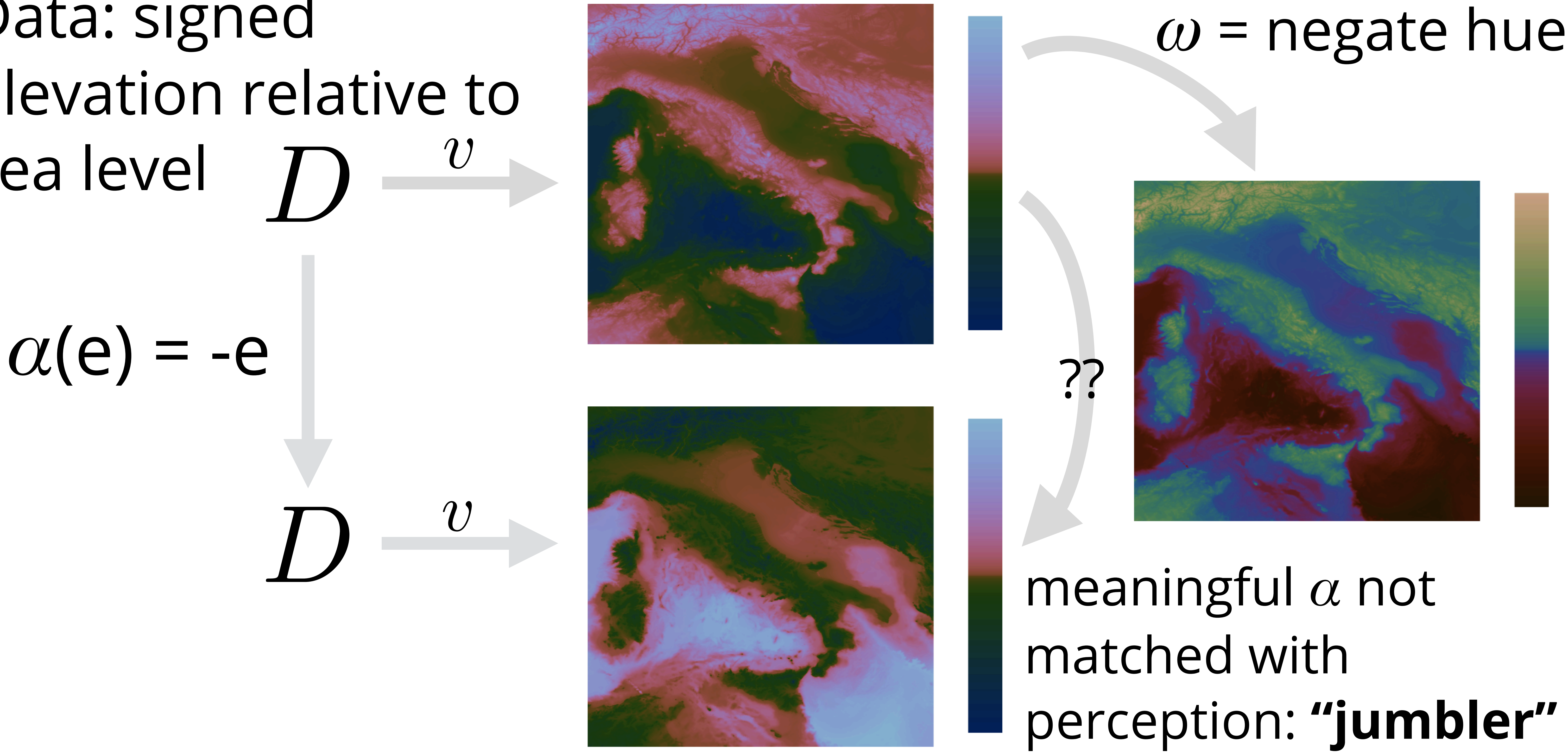
- α and ω well-matched, "$\alpha \cong \omega$"
- ω makes sense, given α



- **Congruence**: visual (external) structure ≅ viewer's mental (internal) structure [Tversky et al. 2002]
- **Effectiveness**: important data attributes mapped to readily perceived visual attributes [Mackinlay 1986]
- **Visual embedding**: visualization preserves distance (in spaces of data, perception) [Demiralp et al. 2014]

Correspondence example: elevation colormap

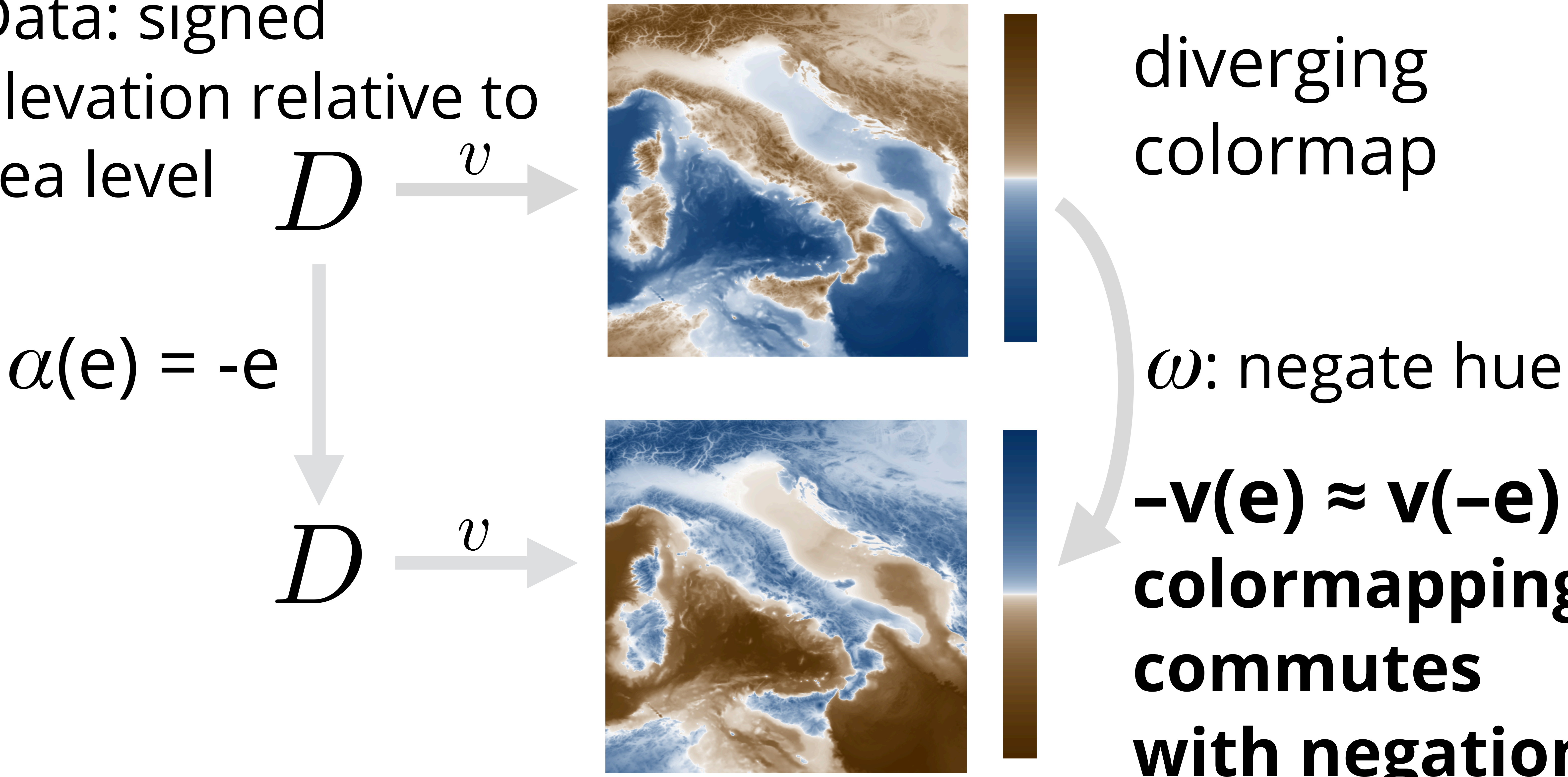
Data: signed elevation relative to sea level



meaningful α not matched with perception: **“jumbler”**

Correspondence example: elevation colormap

Data: signed elevation relative to sea level

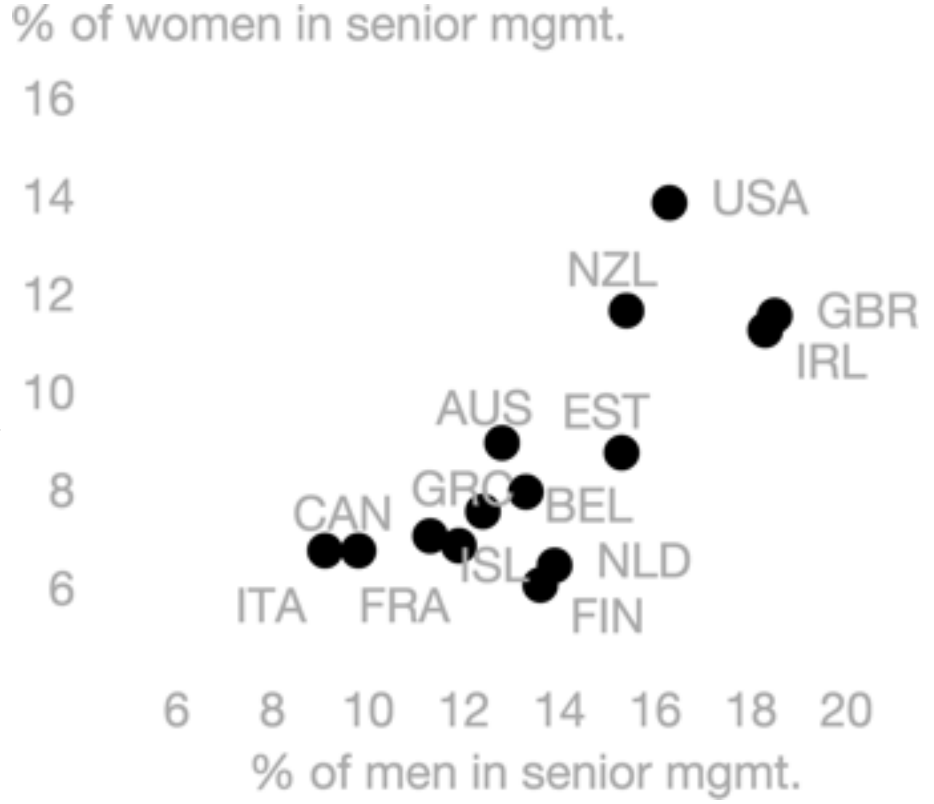


–v(e) ≈ v(–e)
colormapping commutes with negation

Correspondence example: scatterplots

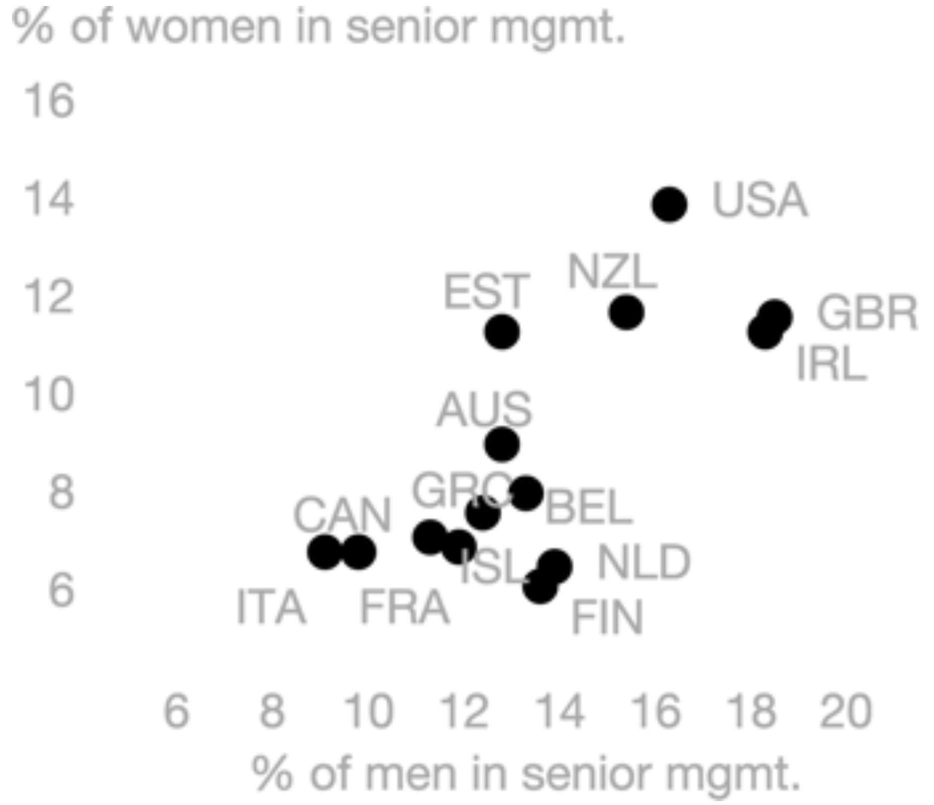
Data: % men vs women employed as senior managers in various countries

$D \xrightarrow{v}$



α: decrease gender gap for one country: EST

$D \xrightarrow{v}$



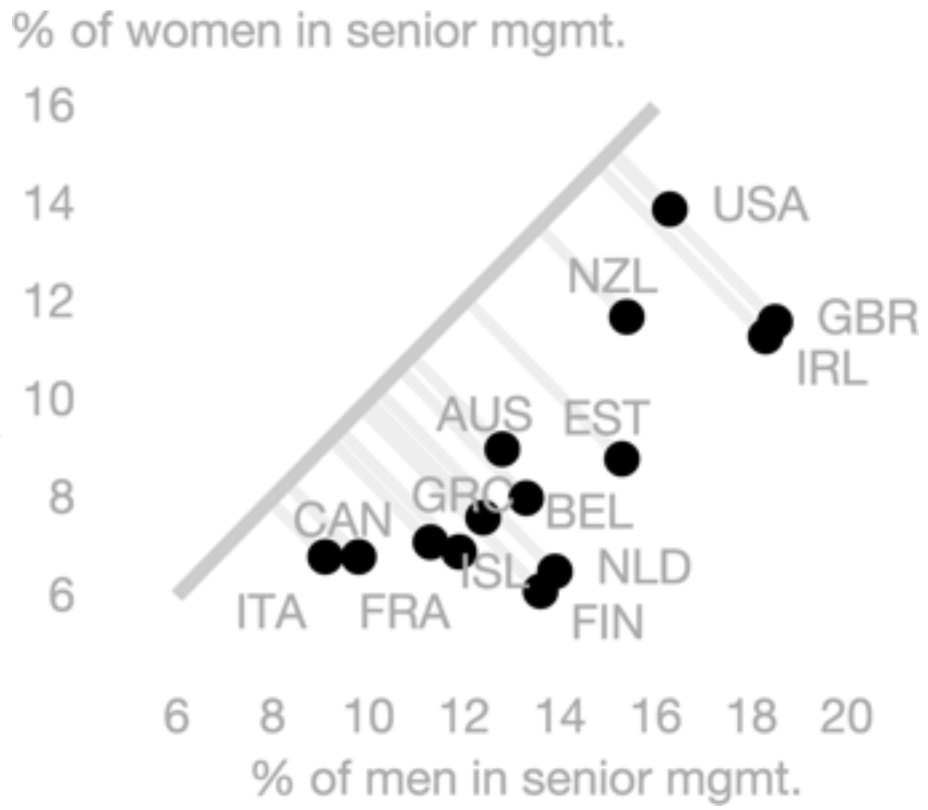
ω? Not clear how big that change was

http://economix.blogs.nytimes.com//2013/04/02/comparing-the-worlds-glass-ceilings/?_r=0

Correspondence example: scatterplots

Data: % men vs women employed as senior managers in various countries

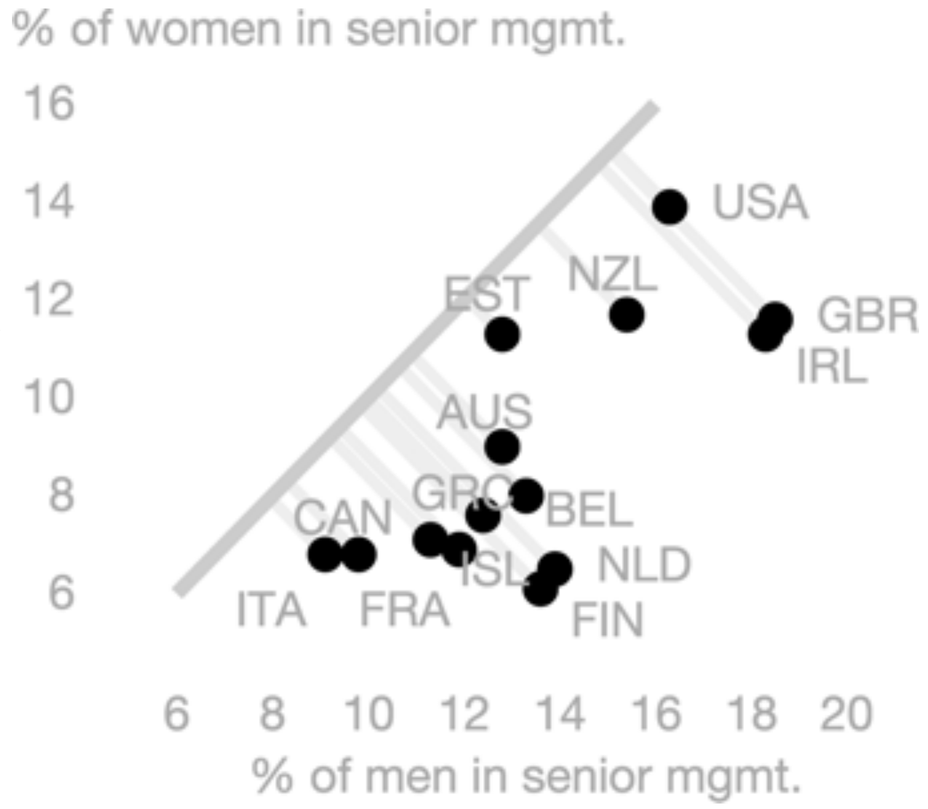
$D \xrightarrow{v}$



add diagonal line (%men = %women) and support lines

α: decrease gender gap for one country: EST

$D \xrightarrow{v}$



ω: change in position along a common scale [Cleveland & McGill 1984]

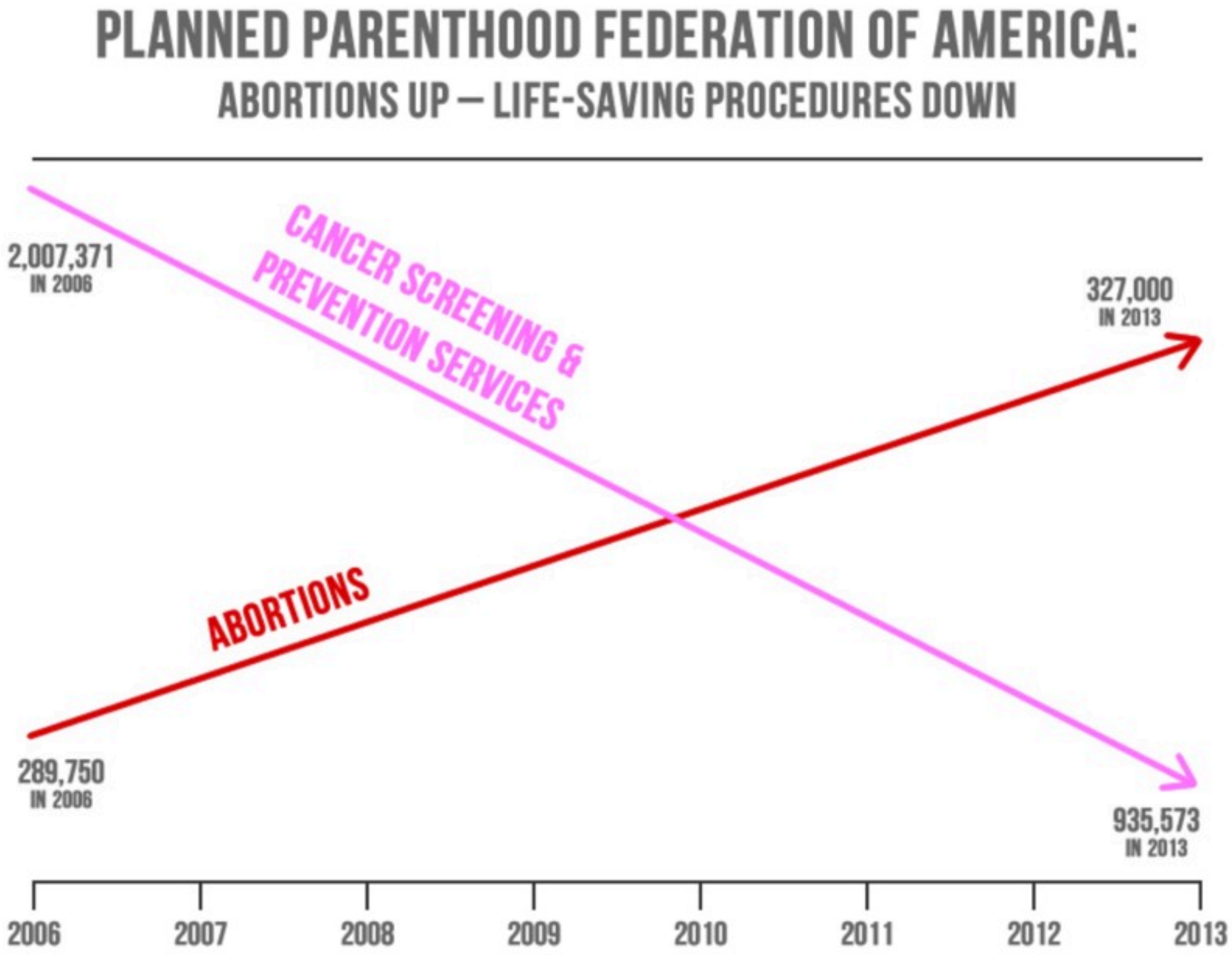
http://economix.blogs.nytimes.com//2013/04/02/comparing-the-worlds-glass-ceilings/?_r=0

Correspondence example: simple plots

29 Sept 2015 US Congressional hearing on Planned Parenthood

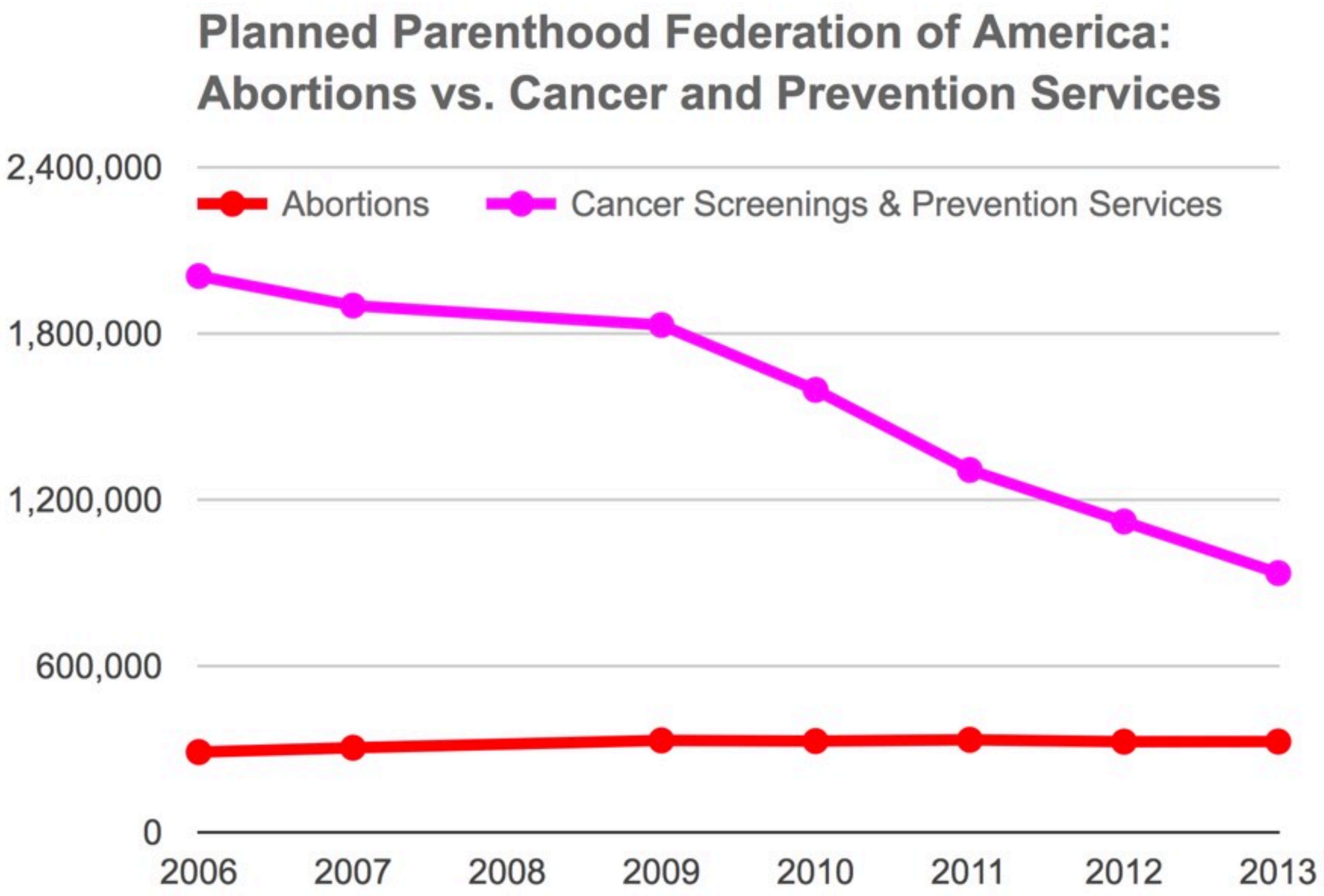
Visualization shown by Rep. Jason Chaffetz, (Republican-Utah)

Note two distinct vertical scalings!



Correspondence example: simple plots

29 Sept 2015 US Congressional hearing on Planned Parenthood

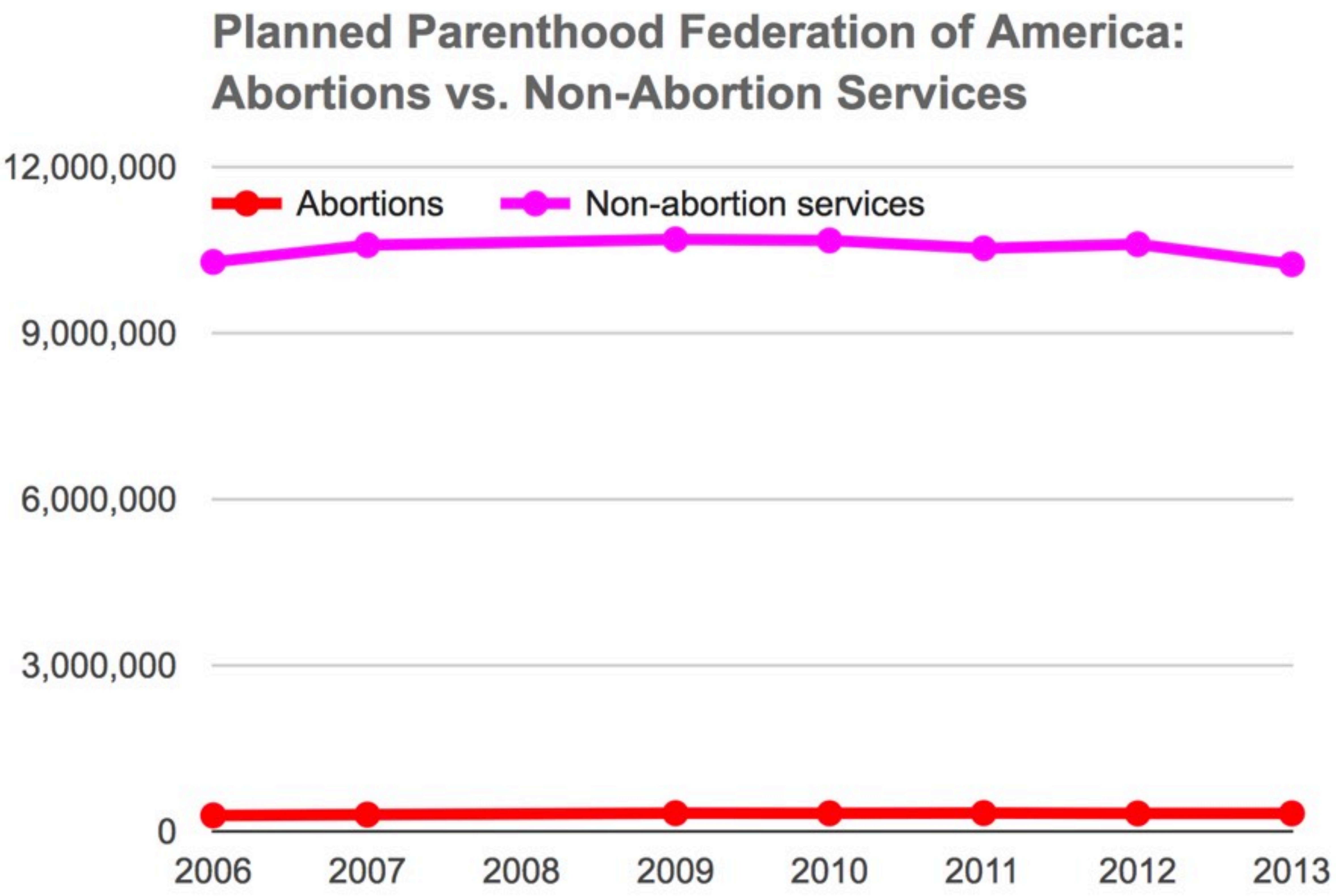


http://www.politifact.com/truth-o-meter/statements/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading-/

Correspondence example: simple plots

29 Sept 2015 US Congressional hearing on Planned Parenthood

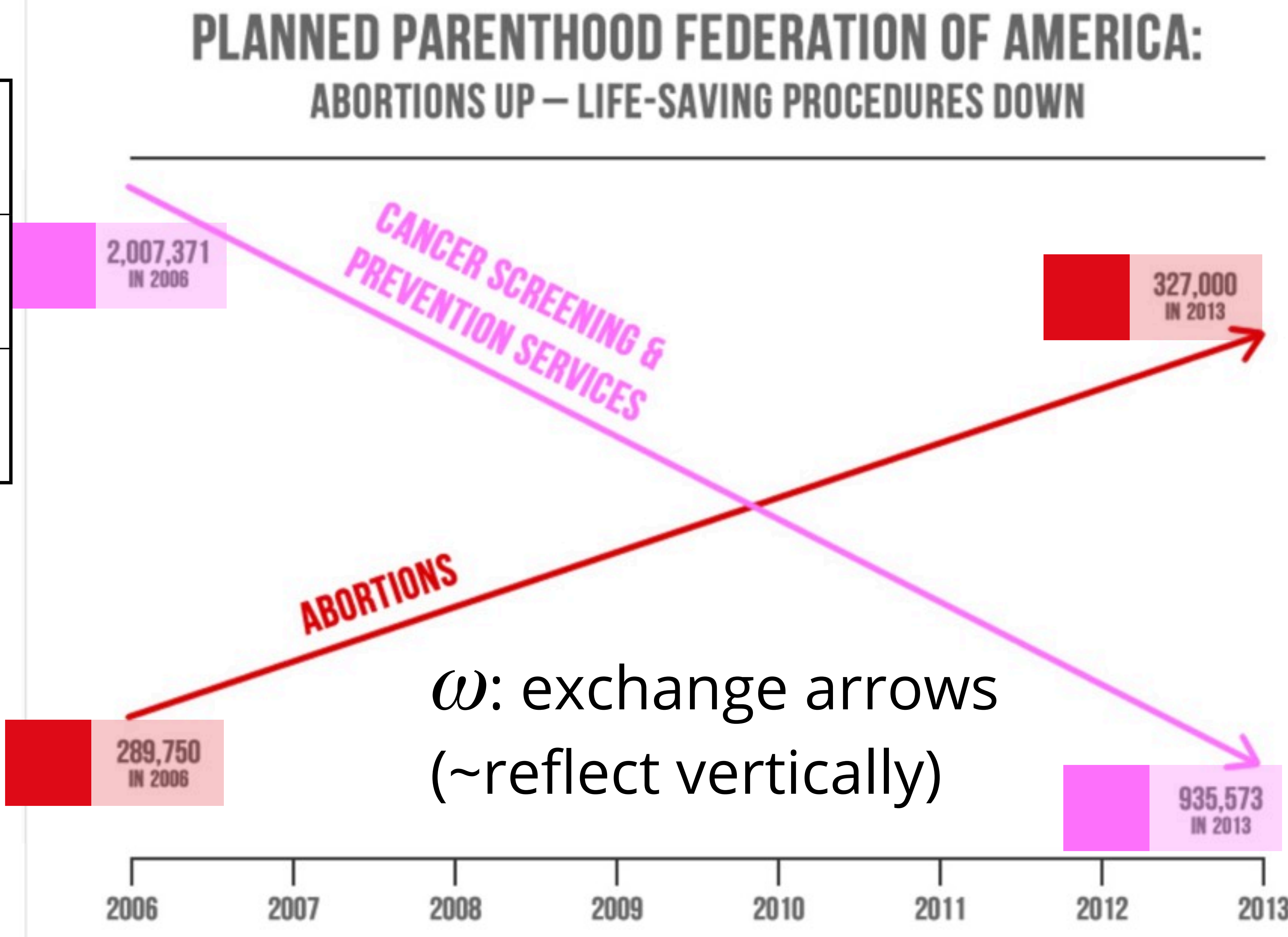
http://www.politifact.com/truth-o-meter/statements/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading-/



So what is misleading, exactly?

Original data values:

	2006	2013
Abortions	0.29M	0.33M
Cancer Scrns & PSs	2.0M	0.94M

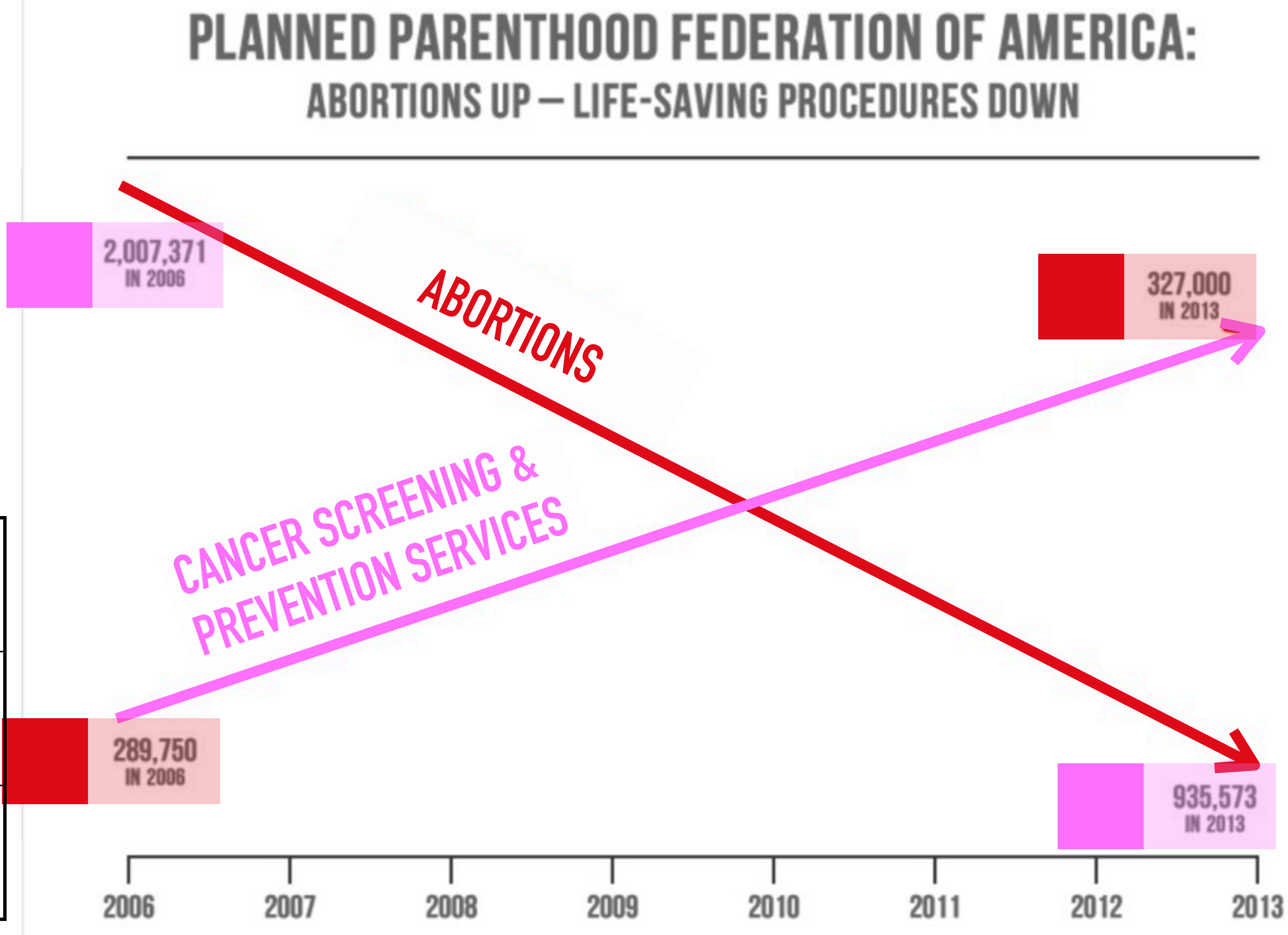


ω: exchange arrows (~reflect vertically)

So what is misleading, exactly?

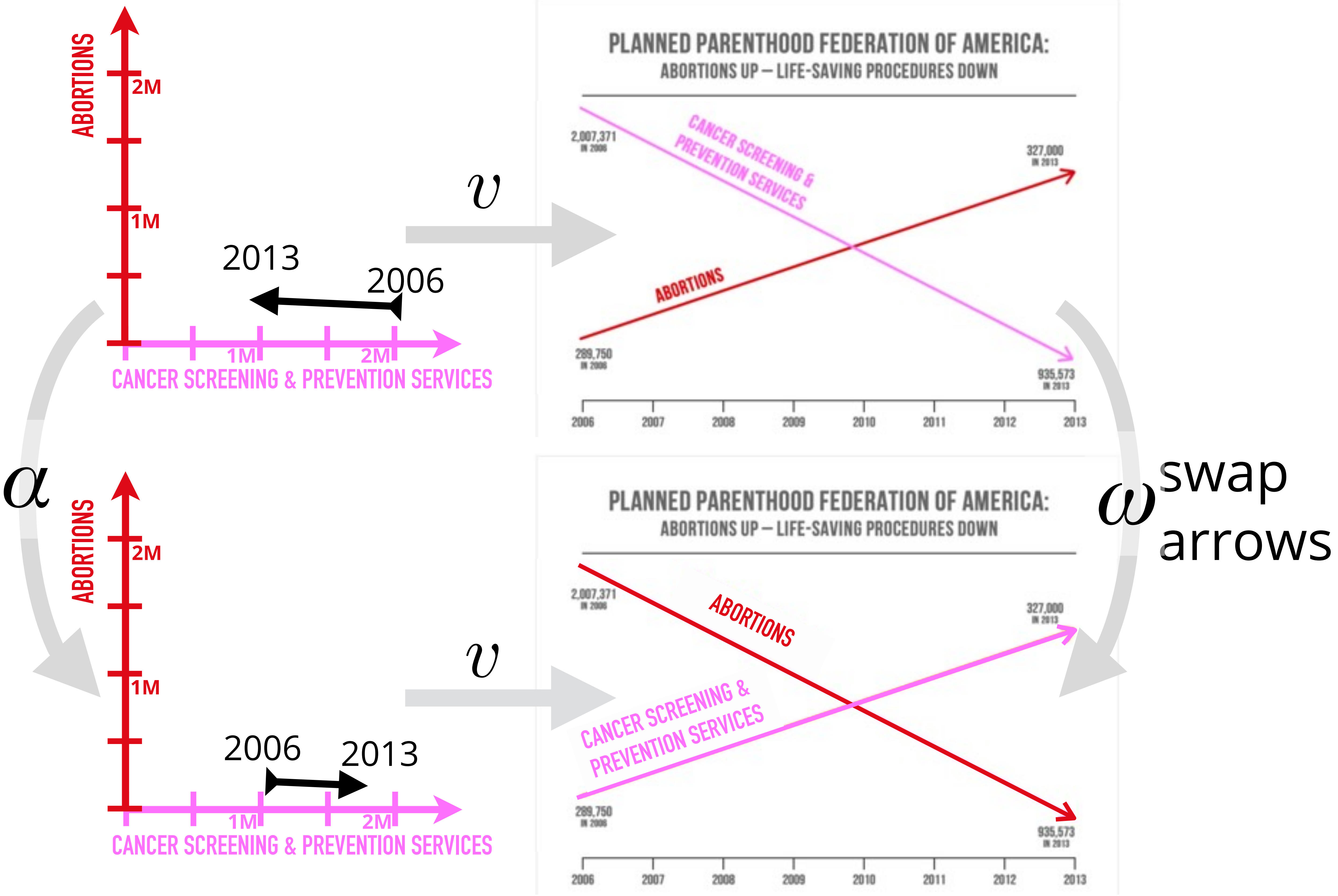
Reading off values (of swapped lines) implied by two distinct vertical scales:

	2006	2013
Abortions	0.34M	0.29M
Cancer Scrns & PSs	1.0M	1.7M



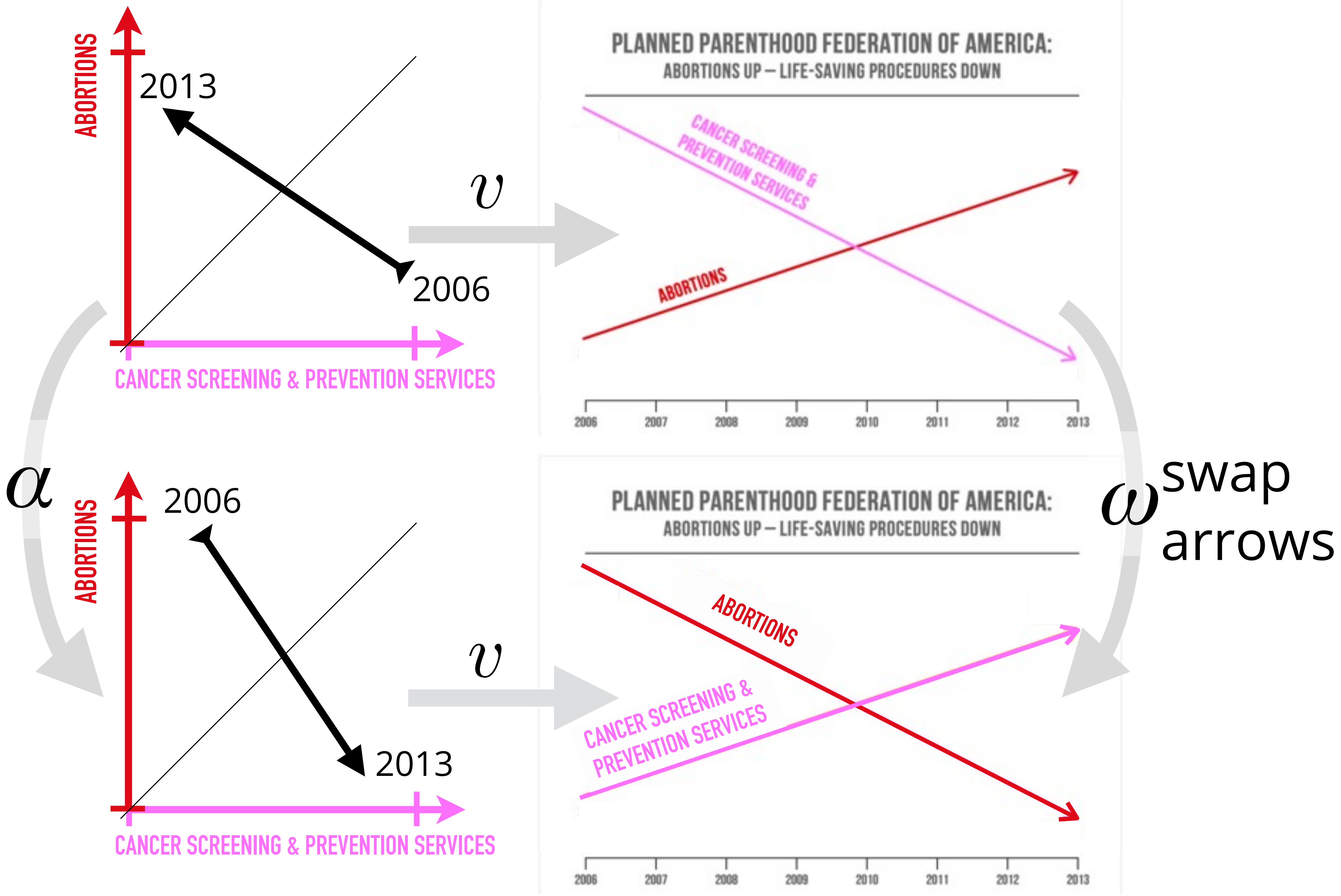
Correspondence example: simple plots

The different vertical scales mean that a clear and obvious ω corresponds to an α that is not especially important
$\Rightarrow$ **ω is a misleader**



Correspondence example: simple plots

With single vertical scale: same ω would correspond to meaningful α: swapping values, or reflecting across x=y (preserving the implied negative correlation)



Colormaps and color ordering

Categorical data: no ordering

http://colorbrewer2.org/

Color ordering: (primarily) luminance

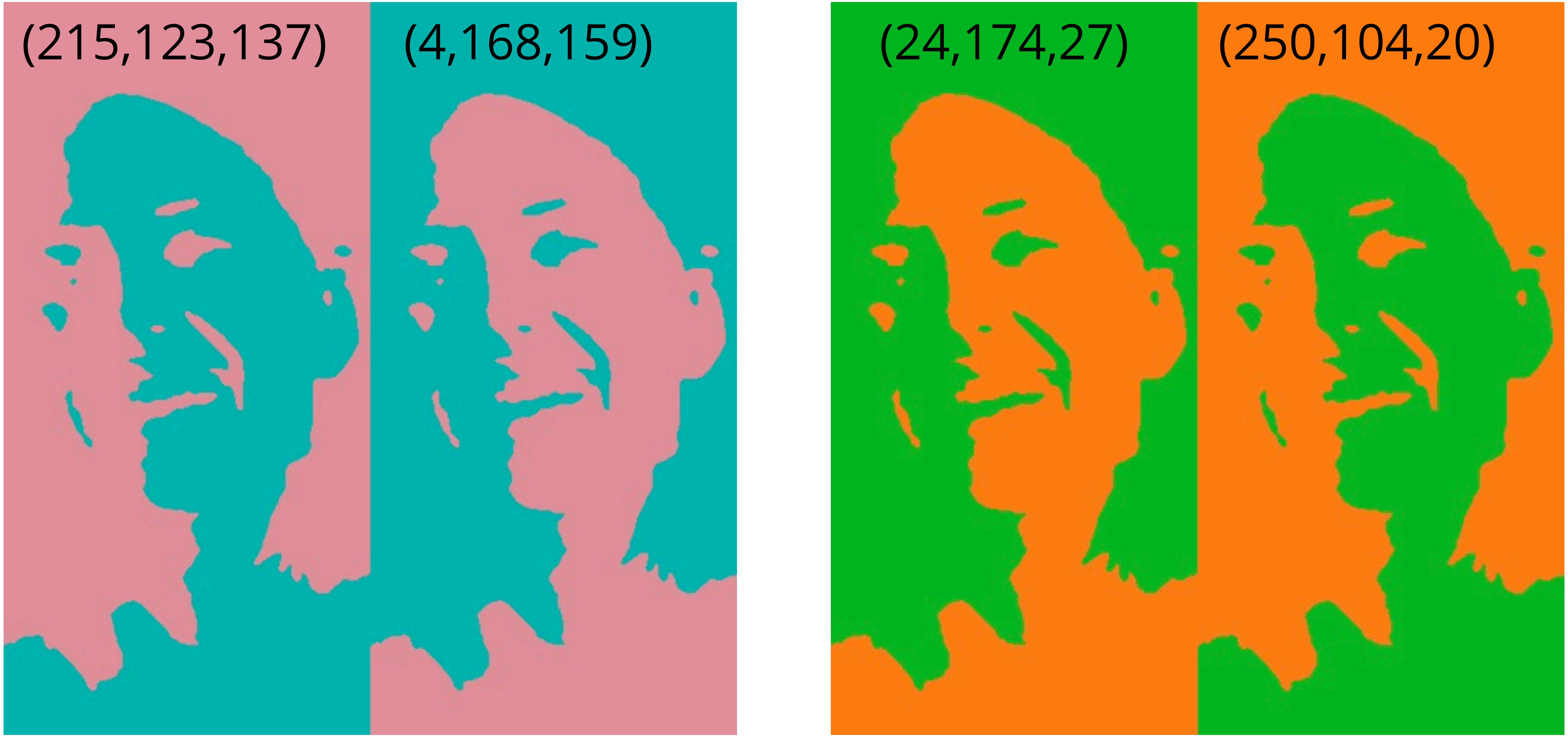
Can trust “L” of LAB or HCL colorspace, or can experimentally compare luminances

- Students more empowered

Face-based luminance matching [Kindlmann et al. 2002]

- web demo by Kai Li

All colors: L=62 in HCL space



Permuting categories will be a jumbler

Visualizing Principle Comp. Analysis (PCA)

Students tasked with creating colormaps to visualize principle components:

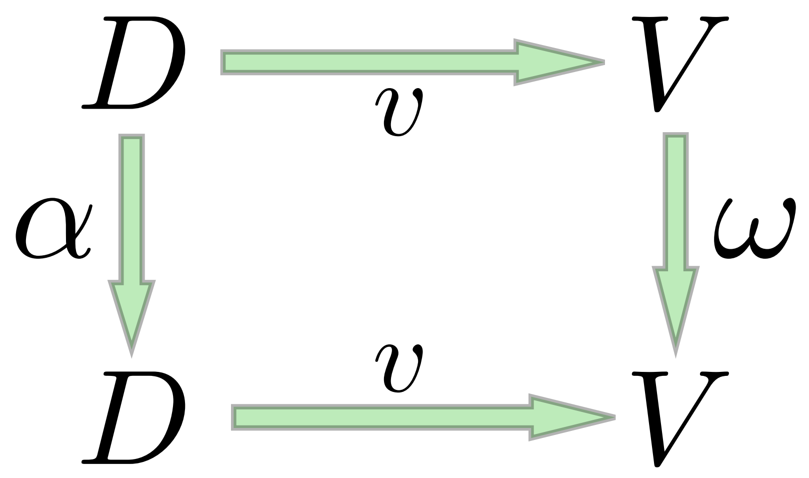
- Maximize correspondence
- Minimize hallucinators

(web demo)

Summary of 3 Principles

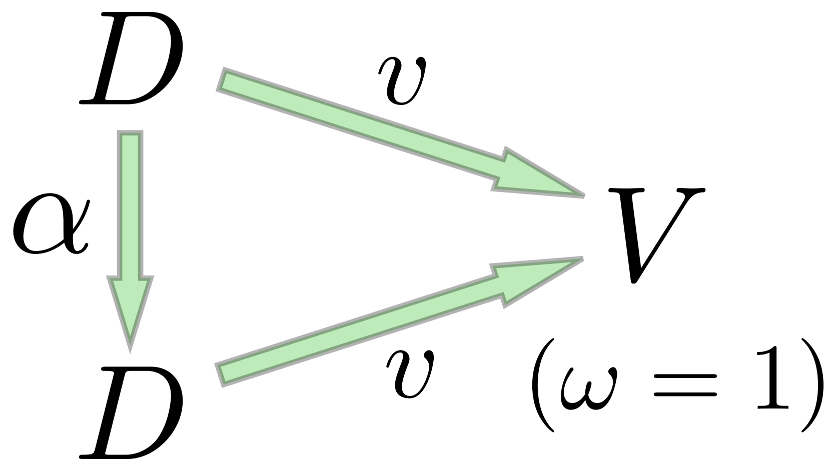
Visual-Data Correspondence

or else a **jumbled** α, or **misleading** ω



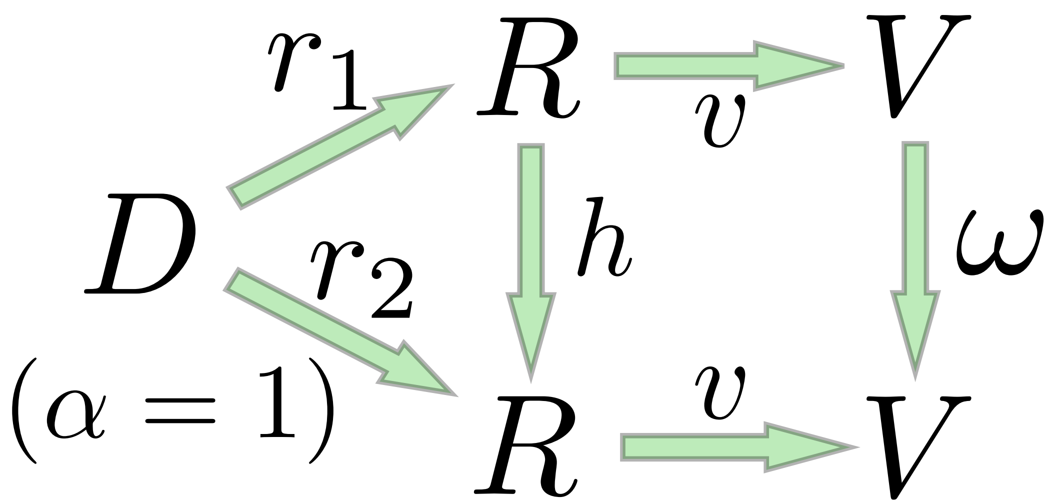
Unambiguous Data Depiction

or else a **confuser** α



Representation Invariance

or else a **hallucinator** h



Questions to ask of a visualization

- If the data were different, would the vis be different (Unambiguous), and different in an informative way? (Correspondence)
- If ambiguous: what are the data changes am I blind to? (Confuser) Is that a problem?
- If not informative: is there another way to lay out or encode the data to create a better correspondence? (removing Jumblers)
- Are there apparent properties in the vis that are not actually in the data (Misleader)
- Could the vis have ended up appearing differently, in a way that is not determined by the data? (Invariance)
- What are changes in the computational/numerical representation, or the execution of algorithm, that should be inconsequential, but are not? (Hallucinator)

References

[Brehmer & Munzner 2013]: M Brehmer and T Munzner. A multi-level typology of abstract visualization tasks. IEEE TVCG., 19(12):2376–2385, 2013

[Card 1999]: SK Card, JD Mackinlay, and B Shneiderman. Information Visualization. In Readings in Information Visualization: Using Vision to Think, chapter 1, pages 1–34. Morgan Kaufmann, 1999

[Cleveland & McGill 1984]: WS Cleveland and R McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. J. American Statistical Association, 79(387):531–554, 1984.

[Demiralp et al. 2014]: Ç Demiralp, CE Scheidegger, GL Kindlmann, DH Laidlaw, and J Heer. Visual embedding: A model for visualization. IEEE CG&A., 34(1):10–15, 2014.

[Gibson 1986]: JJ Gibson. The Ecological Approach To Visual Perception, chapter 8: The Theory of Affordances. Lawrence Erlbaum Associates, 1986.

[Kindlmann et al. 2002]: G Kindlmann, E Reinhard, and S Creem. Face-based Luminance Matching for Perceptual Colormap Generation. In Proceedings of IEEE Visualization, pages 299–306, October 2002.

[Kindlmann & Scheidegger 2014]: G Kindlmann and C Scheidegger. An Algebraic Process for Visualization Design. IEEE Transactions on Visualization and Computer Graphics (Proceedings VIS 2014), 20(12):2181–2190, November 2014.

[Mackinlay 1986]: J Mackinlay. Automating the design of graphical presentations of relational information. ACM Trans. Graph., 5(2):110–141, 1986.

[Meyer et al. 2012]: M Meyer, M Sedlmair, T Munzner. The four-level nested model revisited: blocks and guidelines. In Proc. 2012 BELIV Workshop, pages 11:1–11:6, 2012.

[Munzner 2009]: T Munzner. A nested model for visualization design and validation. IEEE TVCG., 15(6):921–928, 2009.

[Stevens 1946]: SS Stevens. On the theory of scales of measurement. Science, 103(2684):677–680, 1946.

[Tufte 1983]: ER Tufte. The visual display of quantitative information. Graphics Press, 1983

[Tversky et al. 2002]: B Tversky, JB Morrison, and M Betrancourt. Animation: can it facilitate? Intl. J. Hum.-Comp. Stud., 57(4):247–262, 2002.

[Ware 2012]: C Ware. Information visualization: perception for design. Elsevier, 2012.

[van Wijk & H. van de Wetering 1999]: JJ Van Wijk, H van de Wetering. Cushion treemaps: Visualization of hierarchical information. In Proc. Info. Vis., pp 73–78, 1999.

[Ziemkiewicz & Kosara 2009]: C Ziemkiewicz and R Kosara. Embedding information visualization within visual representation. In Z. W. Ras and W. Ribarsky, editors, Advances in Information and Intelligent Systems, volume 251 of Studies in Computational Intelligence, pages 307–326. Springer, 2009.

Acknowledgements

Planned Parenthood plots: http://www.politifact.com/truth-o-meter/statements/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading-/

Conversations with: Tamara Munzner, Stephen Ingram, Hadley Wickham, Çağatay Demiralp, Xavier Tricoche, and Thomas Schultz

2009 Dagstuhl Scientific Visualization Seminar 09251

Questions?