

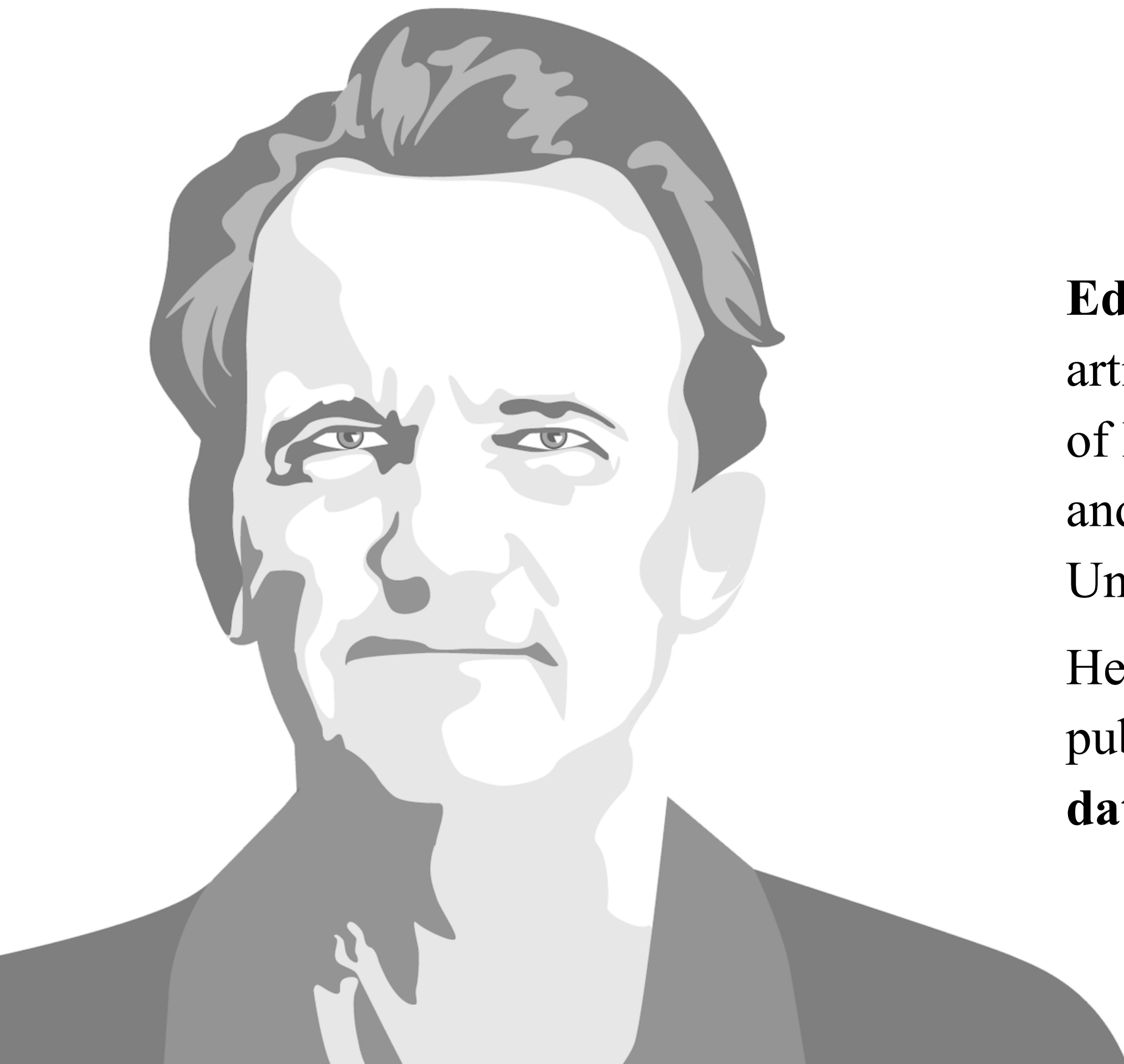
Visualization and Human Perception

A brief introduction of Tufte's Principles

Presented by Zhibang Jiang

THE VISUAL DISPLAY OF QUANTITATIVE INFORMATION



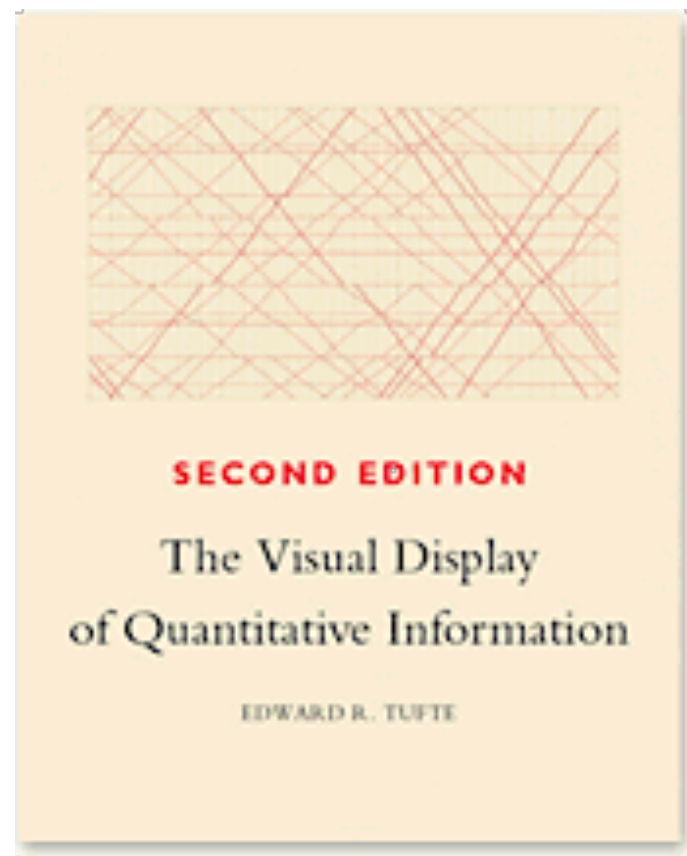


MY RESEARCH TOPIC

Edward Tufte is a statistician, artist and an Emeritus Professor of Political Science, Statistics, and Computer Science at Yale University.

He wrote, designed, and self-published **4 classic books on data visualization.**

MY RESEARCH TOPIC



This is a book all about design and statistics, about how to **communicate information** through the simultaneous presentation of words, numbers, and pictures.

"A landmark book, a wonderful book."

FREDERICK MOSTELLER

EDWARD'S AGENDA

The first part of this book reviews the graphical practice of the two centuries since **Playfair**.

The second part provides a language for discussing graphics and a practical theory of data graphics.

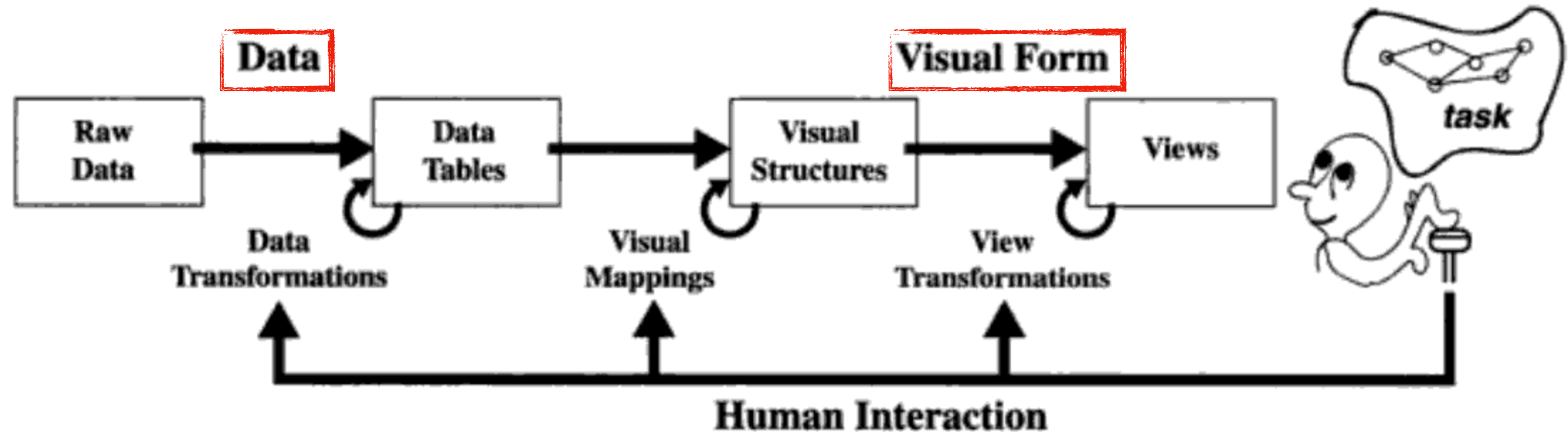
PART I GRAPHICAL PRACTICE

— [1	<i>Graphical Excellence</i>	13
	2	<i>Graphical Integrity</i>	53
	3	<i>Sources of Graphical Integrity and Sophistication</i>	79

PART II THEORY OF DATA GRAPHICS

— [4	<i>Data-Ink and Graphical Redesign</i>	91
	5	<i>Chartjunk: Vibrations, Grids, and Ducks</i>	107
	6	<i>Data-Ink Maximization and Graphical Design</i>	123
	7	<i>Multifunctioning Graphical Elements</i>	139
	8	<i>Data Density and Small Multiples</i>	161
	9	<i>Aesthetics and Technique in Data Graphical Design</i>	177
		<i>Epilogue: Designs for the Display of Information</i>	191

VISUALIZATION PROCESS MODEL PRESENTED BY CARD ET AL 1999



Raw Data: idiosyncratic formats

Visual Structures: spatial substrates + marks + graphical properties

Data Tables: relations (cases by variables) + metadata

Views: graphical parameters (position, scaling, clipping, ...)

Card, Mackinlay. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

MY(CARD'S) AGENDA

DATA TRANSFORMATION

VISUAL MAPPINGS

VIEW TRANSFORMATIONS

EDWARD'S AGENDA

PART I GRAPHICAL PRACTICE

- 1 *Graphical Excellence* 13
- 2 *Graphical Integrity* 53
- 3 *Sources of Graphical Integrity and Sophistication* 79

PART II THEORY OF DATA GRAPHICS

- 4 *Data-Ink and Graphical Redesign* 91
- 5 *Chartjunk: Vibrations, Grids, and Ducks* 107
- 6 *Data-Ink Maximization and Graphical Design* 123
- 7 *Multifunctioning Graphical Elements* 139
- 8 *Data Density and Small Multiples* 161
- 9 *Aesthetics and Technique in Data Graphical Design* 177
- Epilogue: Designs for the Display of Information* 191

Card, Mackinlay. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

TUFTE'S PRINCIPLES UNDER MY(CARD'S) AGENDA

DATA TRANSFORMATION

- 1. Nearly always multivariate
- 2. Graphics must not quote data out of context

VISUAL MAPPINGS

- 1. Requires telling the truth about the data.
- 2. The choice of design: the only worse design than a pie chart is several of them.
- 3. The representation of numbers should be directly proportional to numerical quantities represented. (we can do distortion, but not the same distortion.)
- 4. Show data variation, not design variation.
- 5. Standardized unites of monetary measurement are better than nominal unites.
- 6. The number of information-carrying dimensions depicted should not exceed the number of dimensions in the data.
- 7. Erase non- and redundant data-ink to maximize the data-ink ratio. (non-information overload)
- 8. Avoid content-free decoration, including chartjunk.

VIEW TRANSFORMATIONS

- 1. Small multiples
- 2. Proportion and scale: line weight and lettering
- 3. Proportion and scale: the shape of graphics

TUFTE'S PRINCIPLES UNDER 'VISUAL MAPPINGS'

VISUAL MAPPINGS

1. Requires telling the truth about the data.
- 2. The choice of design: the only worse design than a pie chart is several of them.**
- 3. The representation of numbers should be directly proportional to numerical quantities represented.** (we can do distortion, but not the same distortion.)
4. Show data variation, not design variation.
5. Standardized units of monetary measurement are better than nominal units.
- 6. The number of information-carrying dimensions depicted should not exceed the number of dimensions in the data.**
7. Erase non- and redundant data-ink to maximize the data-ink ratio. (non-information overload)
- 8. Avoid content-free decoration, including chartjunk.**

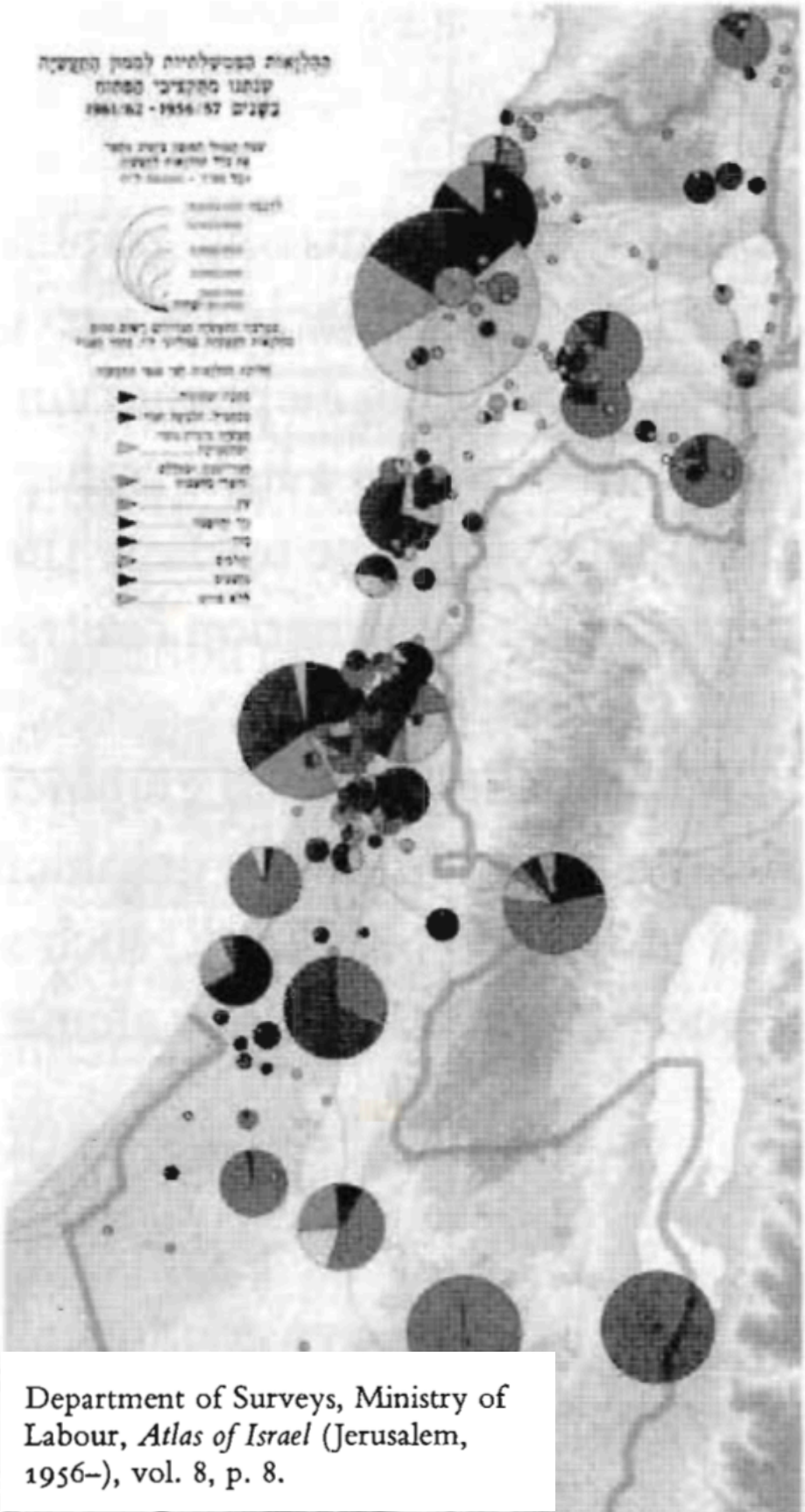
The Worst Chart in the World

The **only worse** design than a pie chart is **several of them**, for then the viewer is asked to compare quantities located in spatial disarray **both within and between pies**, as in this heavily encoded example from an atlas.¹

Given their low data-density and failure to order numbers along a visual dimension, pie charts should never be used.²

¹ Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 178)

² This point is made decisively in Jacques Bertin, *Graphics and Graphic Information Processing* (Bertin, 1981). Bertin describes multiple pie charts as “completely useless” (p. 111).



TUFTE'S PRINCIPLES UNDER 'VISUAL MAPPINGS'

THE CHOICE OF DESIGN

PIE CHART

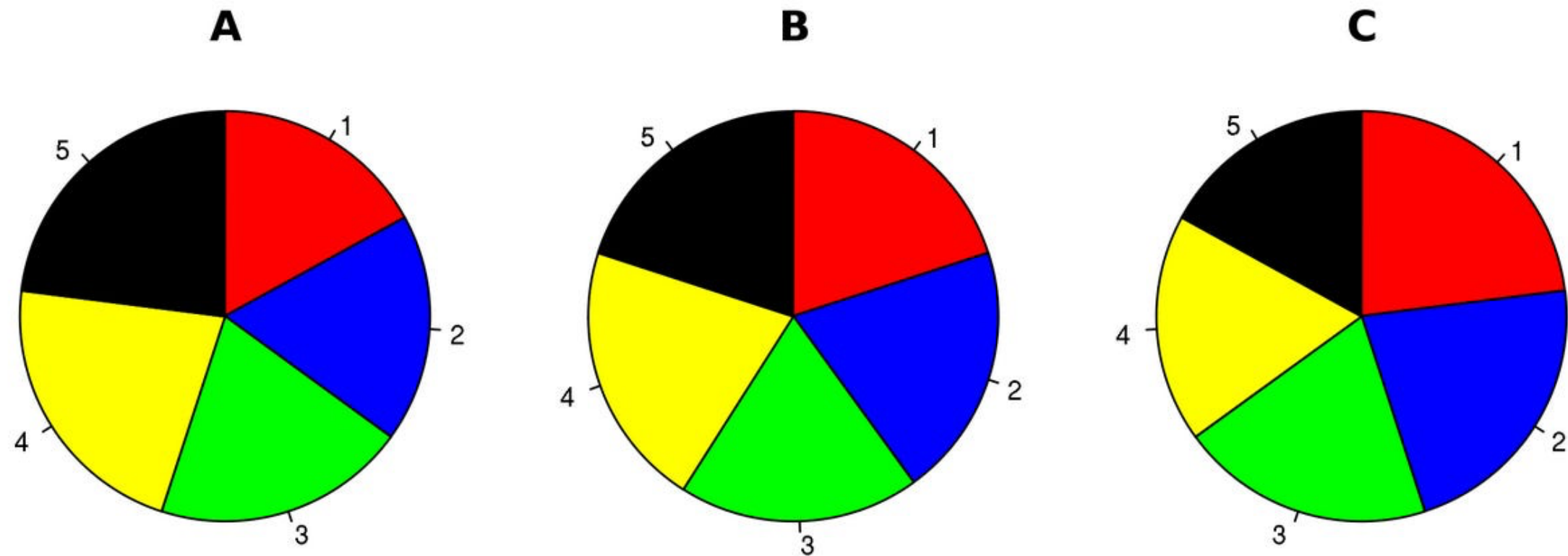
HUMAN PERCEPTION

THINKING TASKS

CHOICE OF CHART

MEDIUM DIMENSIONS

CHARTJUNK



In the first race, is candidate 5 doing better than candidate 3?

Who did better between time A and time B, candidate 2 or candidate 4?

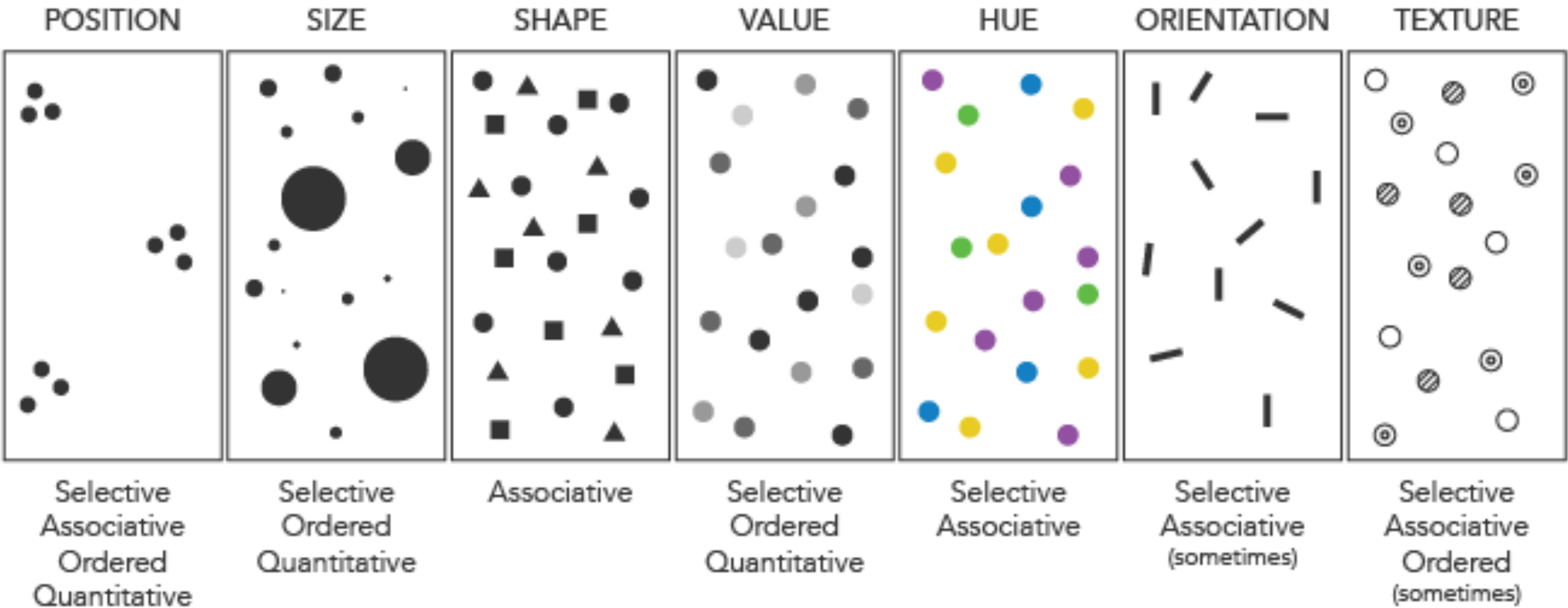
Who has the most momentum in the race?

Source: <https://www.businessinsider.com/pie-charts-are-the-worst-2013-6>

TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

- THE CHOICE OF DESIGN
- PIE CHART
- HUMAN PERCEPTION
- THINKING TASKS
- CHOICE OF CHART
- MEDIUM DIMENSIONS
- CHARTJUNK

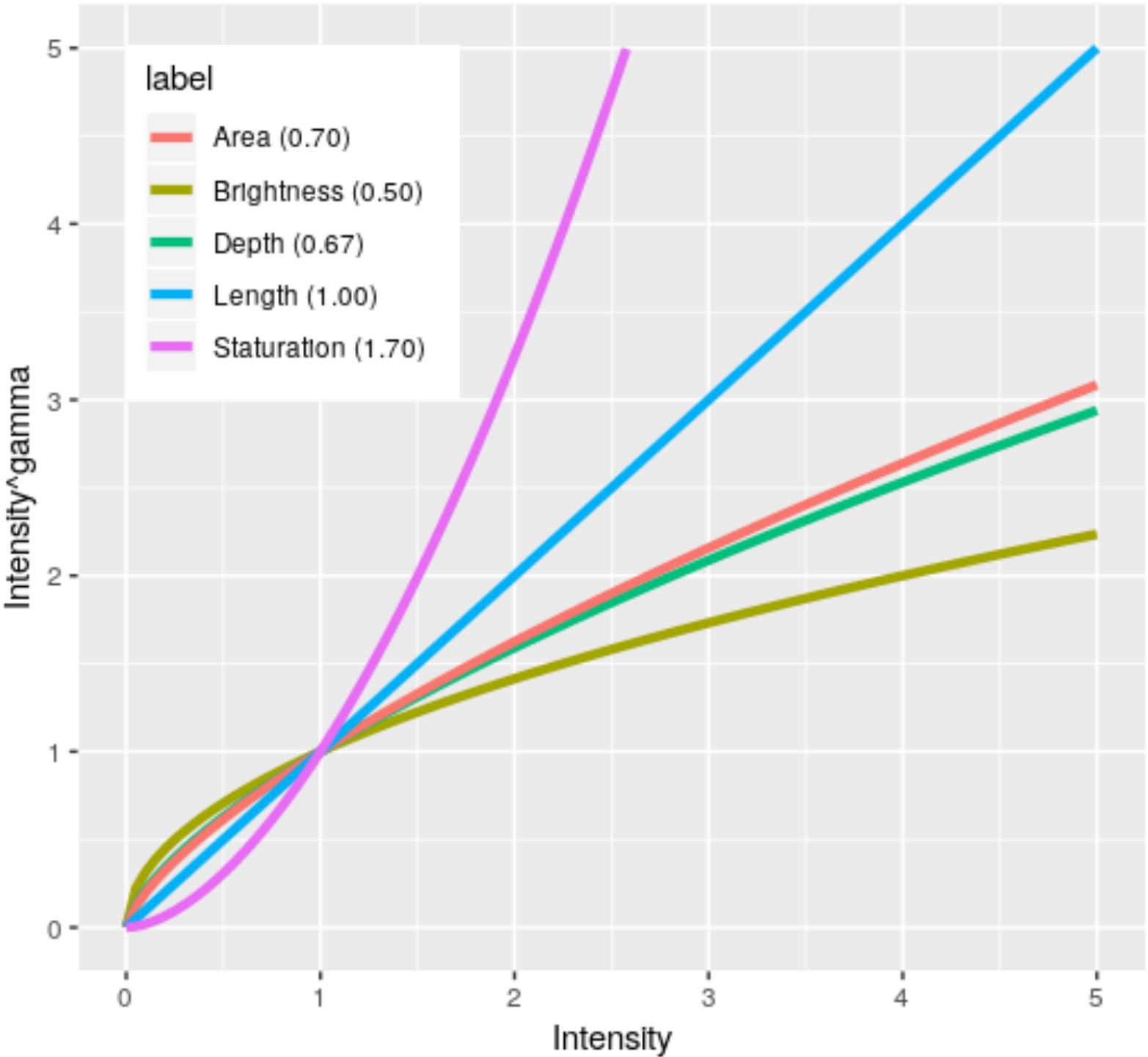
Bertin's Visual Variables



The Study of Channel Accuracy

Stevens' power law is a proposed relationship between the **magnitude** of a physical **stimulus** and the intensity or strength that people feel.

Stevens, S. S. (2017).
Psychophysics:
Introduction to its
perceptual, neural and
social prospects.
Routledge, 1975. (p. 15)



The Study of Graphical Perception

Cleveland, W. S., &
McGill, R. (1985).
*Graphical perception and
graphical methods for
analyzing scientific data.*
Science, 229(4716),
828-833.

Experiment 1

Position (Common)

Angle

Experiment 2

Position (Common)

Length

Experiment 3

Position (Common)

Position (Nonaligned)

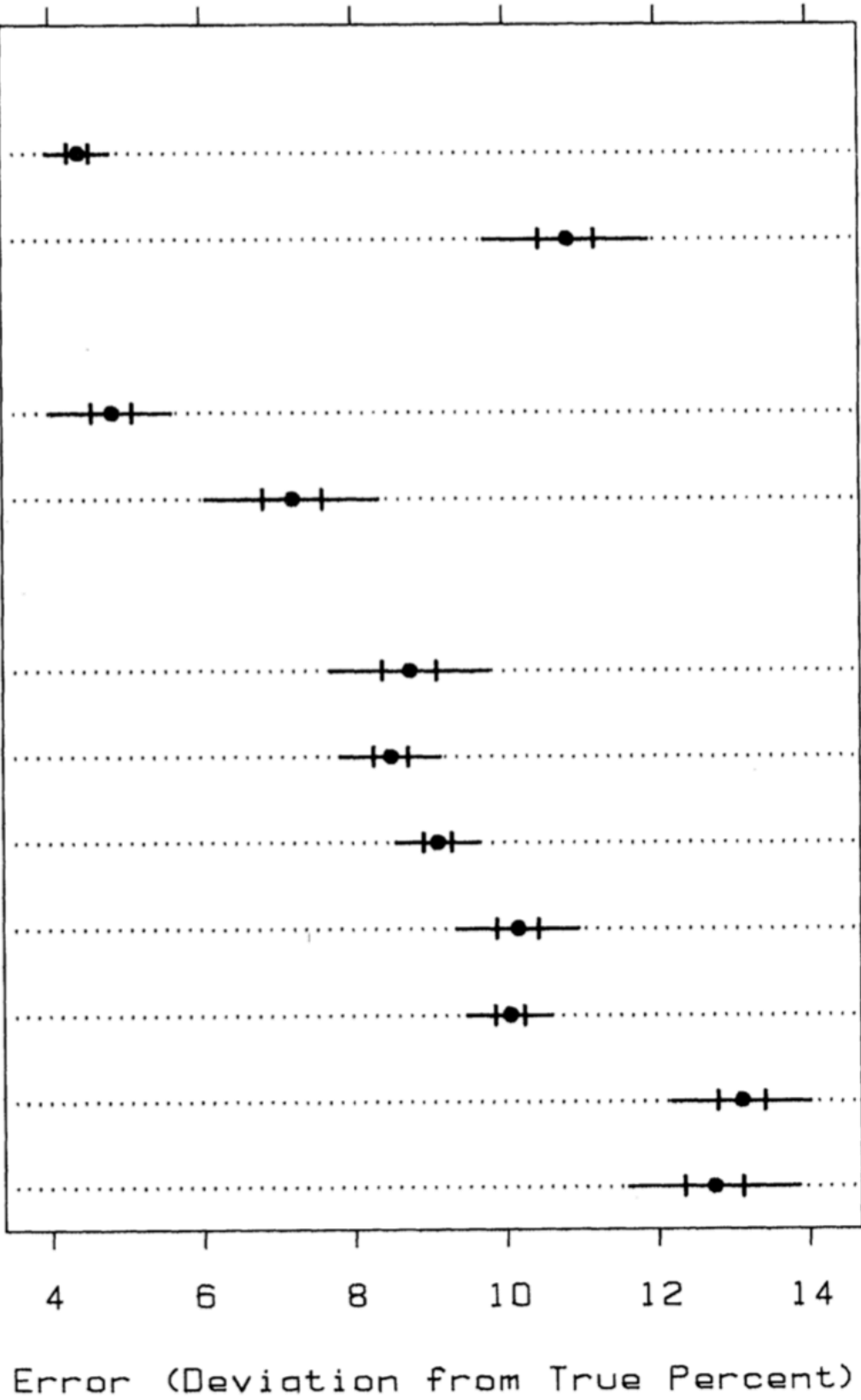
Length

Angle

Slope

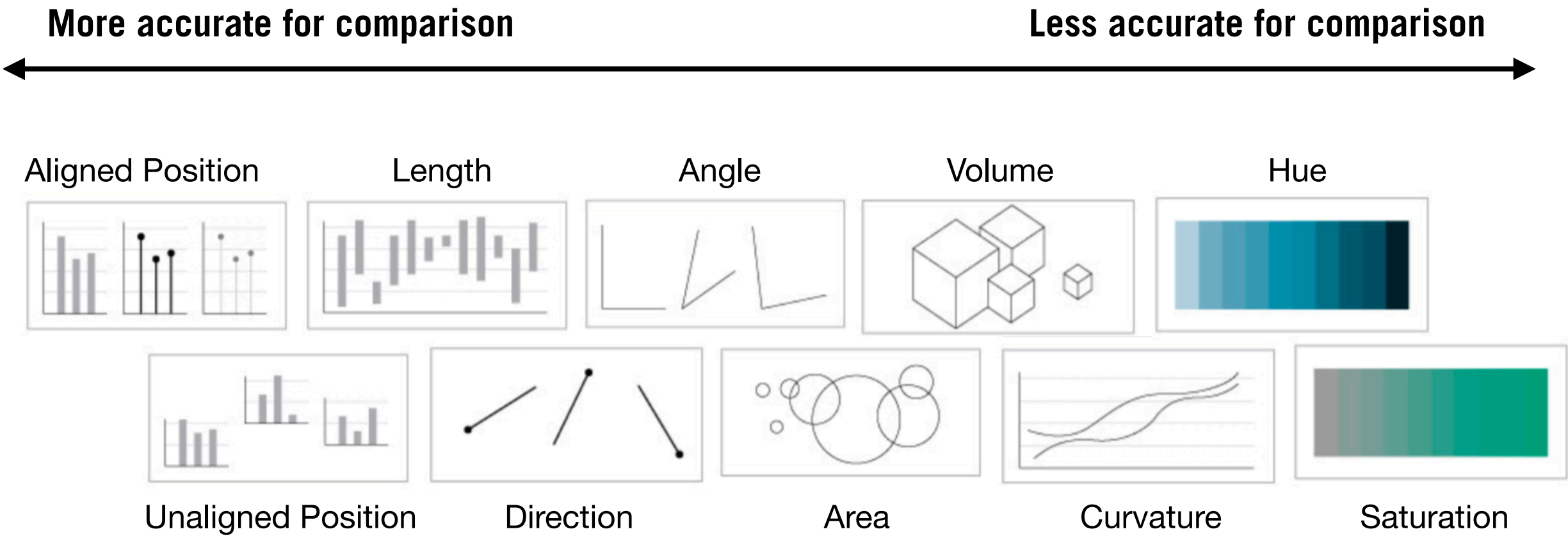
Circle Area

Blob Area



TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

- THE CHOICE OF DESIGN
- PIE CHART
- HUMAN PERCEPTION
- THINKING TASKS
- CHOICE OF CHART
- MEDIUM DIMENSIONS
- CHARTJUNK

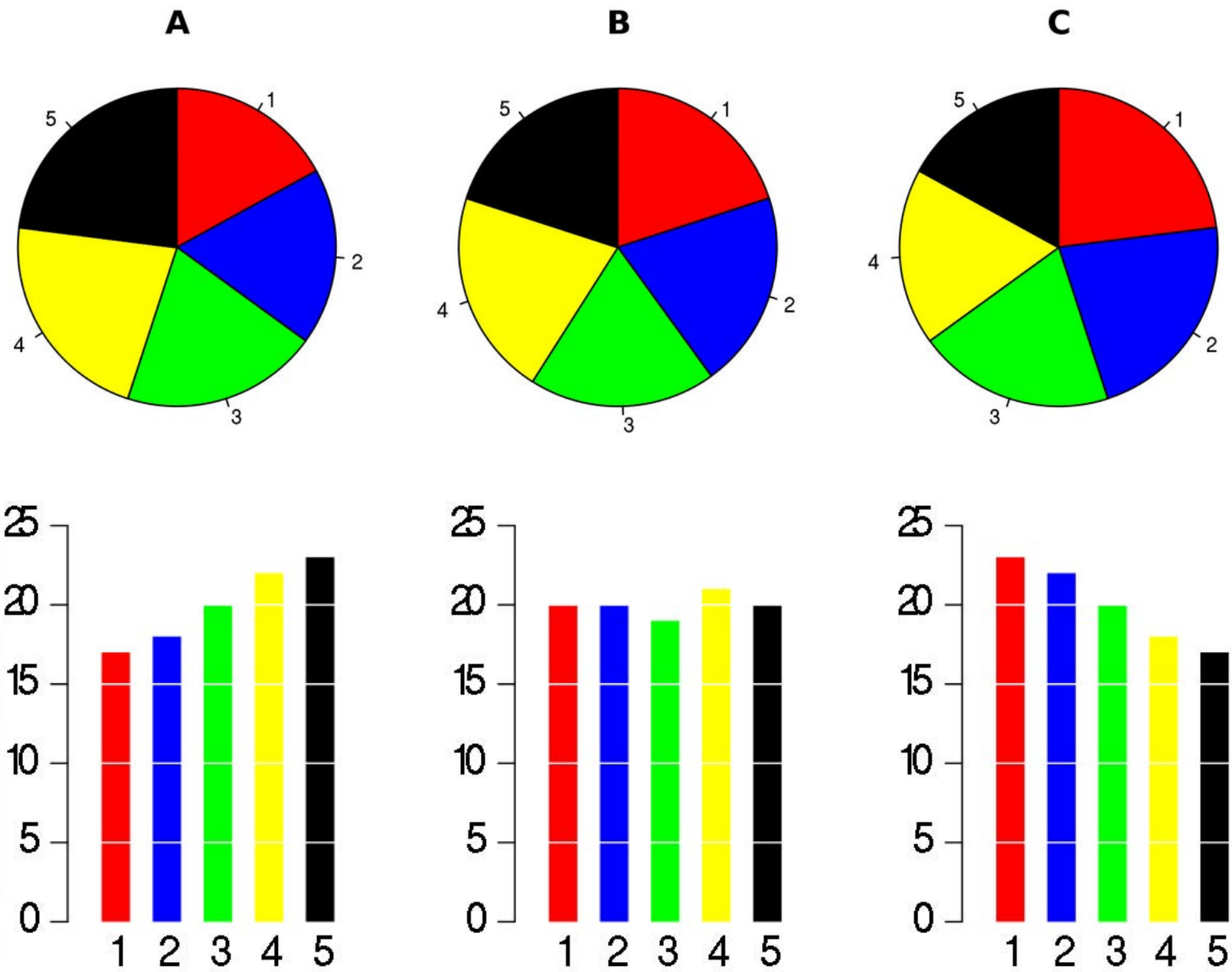


Ranking of perceptual tasks [Cleveland McGill 1985]

TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

The Worst Chart in the World

- THE CHOICE OF DESIGN
- PIE CHART
- HUMAN PERCEPTION
- THINKING TASKS
- CHOICE OF CHART
- MEDIUM DIMENSIONS
- CHARTJUNK



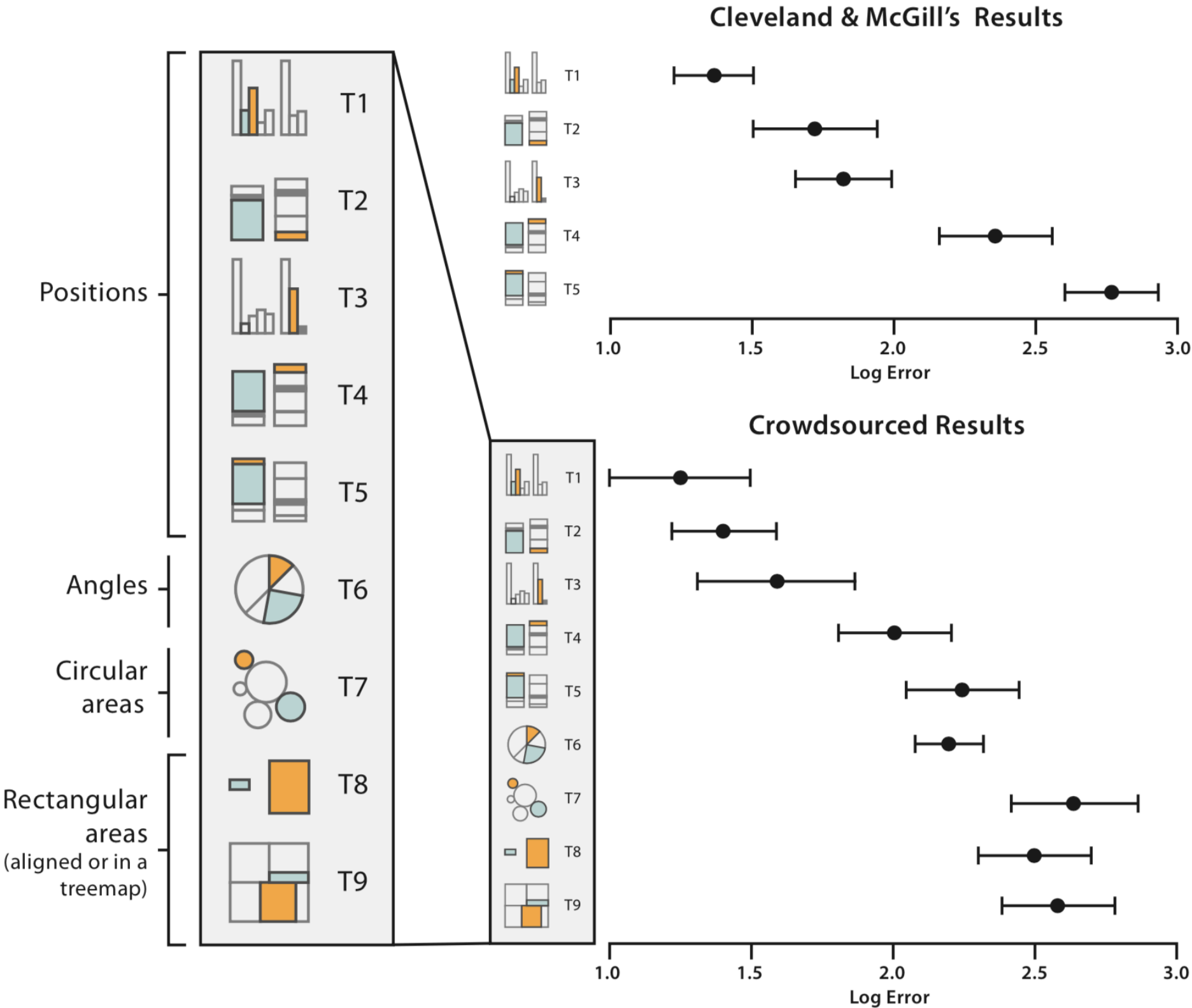
Source: <https://www.businessinsider.com/pie-charts-are-the-worst-2013-6>

TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

- THE CHOICE OF DESIGN
- PIE CHART
- HUMAN PERCEPTION
- THINKING TASKS
- CHOICE OF CHART
- MEDIUM DIMENSIONS
- CHARTJUNK

The Study of
Graphical
Perception

Jeffrey Heer and Michael Bostock (2010)
“Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design,”
Proceedings of the SIGCHI, 203-212.

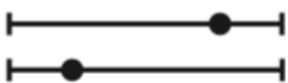
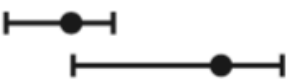

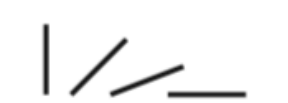

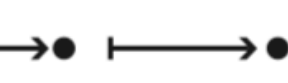






TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

- THE CHOICE OF DESIGN
- PIE CHART
- HUMAN PERCEPTION
- THINKING TASKS
- CHOICE OF CHART
- MEDIUM DIMENSIONS
- CHARTJUNK

Channels: Expressiveness Types and Effectiveness Ranks

➔ **Magnitude Channels: Ordered Attributes**

Position on common scale		Effectiveness
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	

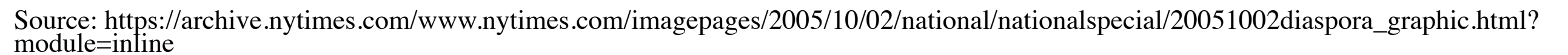
Munzner, Tamara. Visualization Analysis and Design, (2014). (p. 102)

Summary

“A graphical form that involves elementary perceptual tasks that lead to more accurate judgements than another graphical form will result in a better organization and increase the chances of a correct perception of patterns and behavior.”

——Cleveland and McGill

CHARTJUNK

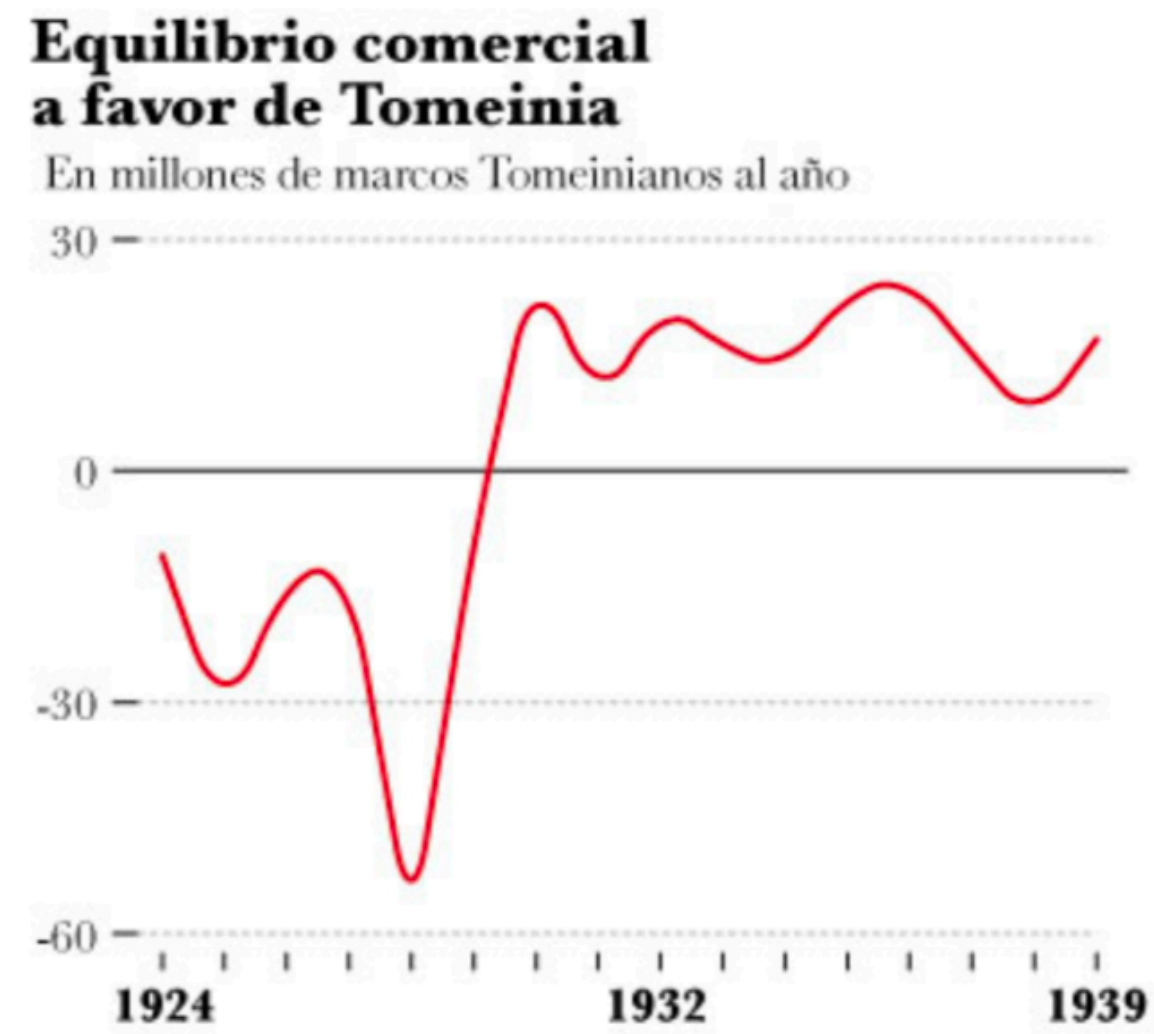


*“What are the thinking tasks
that these displays are
supposed to serve?”*

——Edward R. Tufte

Difference between two variables?

- THE CHOICE OF DESIGN
- PIE CHART
- HUMAN PERCEPTION
- THINKING TASKS
- CHOICE OF CHART
- MEDIUM DIMENSIONS
- CHARTJUNK



TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

THE CHOICE OF DESIGN

PIE CHART

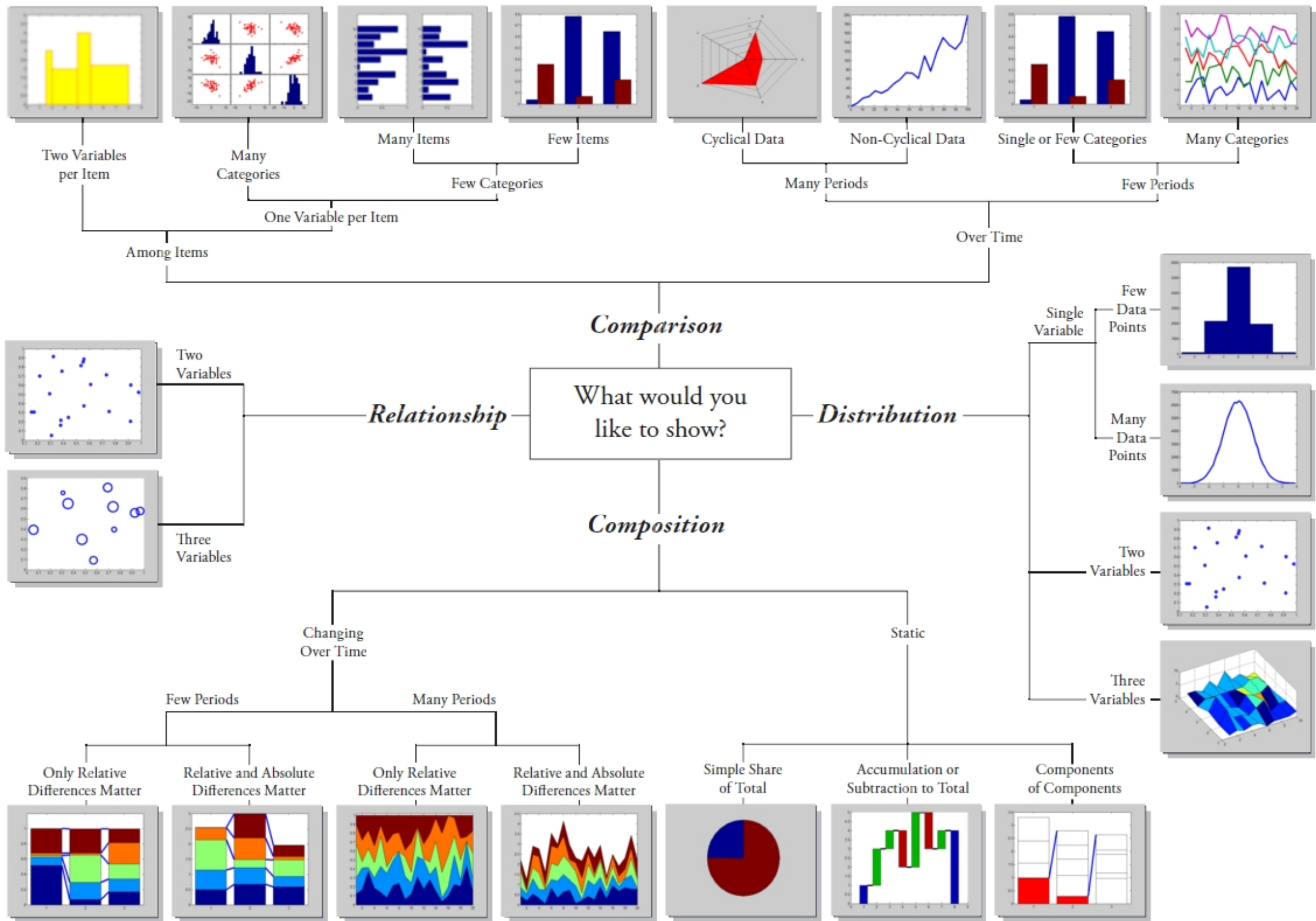
HUMAN PERCEPTION

THINKING TASKS

CHOICE OF CHART

MEDIUM DIMENSIONS

CHARTJUNK



Source: <http://apandre.wordpress.com/dataviews/choiceofchart/>

To Defeat Distortion in a Data Graphic

Two Principles:

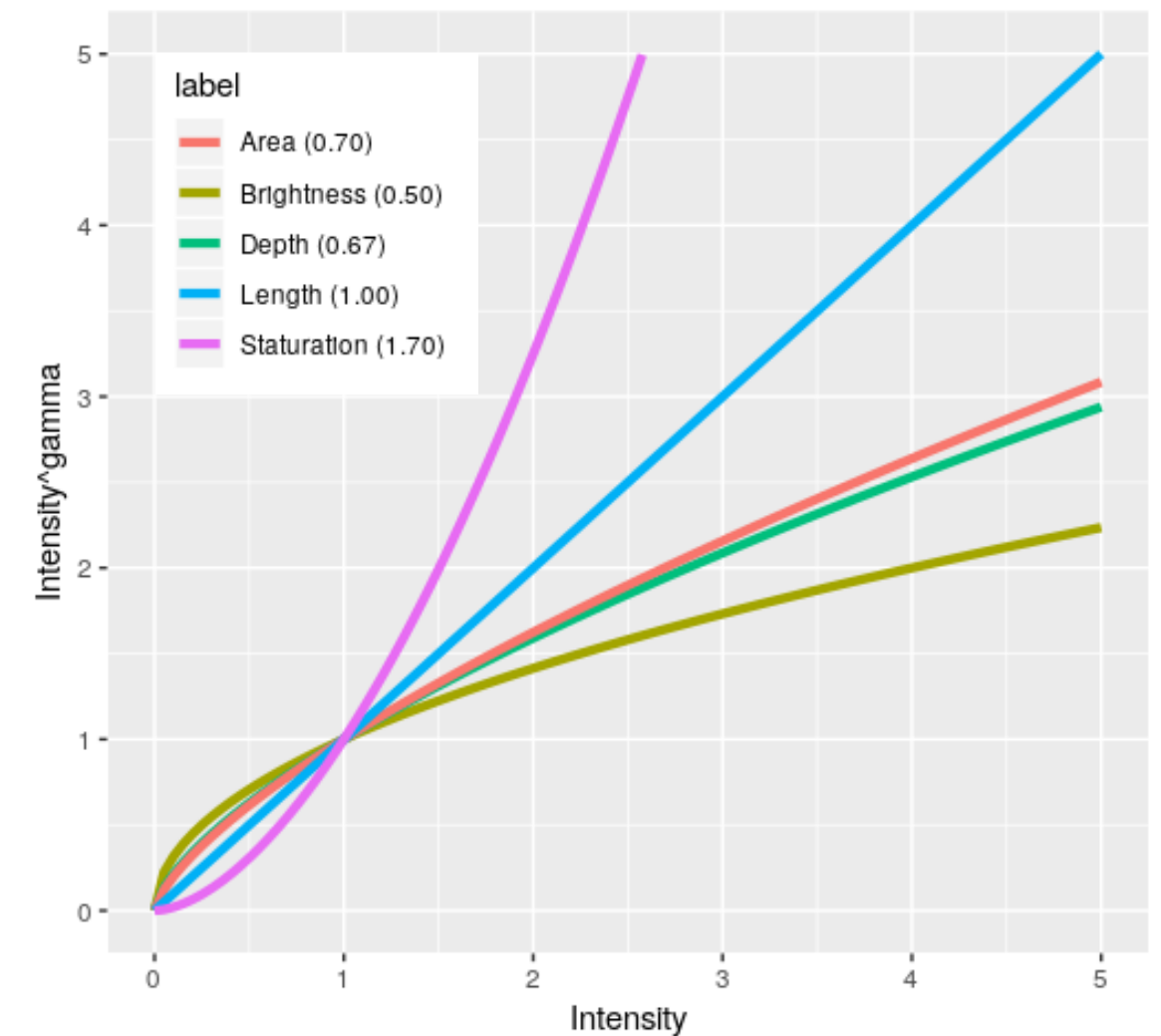
The representation of numbers should be **directly proportional to numerical quantities** represented.

Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity.

Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 55)

Stevens, S. S. (2017). *Psychophysics: Introduction to its perceptual, neural and social prospects*. Routledge, 1975. (p. 15)

I think I see that area B
is 3.14 times bigger than
area A. Is that correct?



Lie Factor

Violations of the first principle (The representation of numbers should be directly proportional to numerical quantities represented) to what extent constitute one form of graphic misrepresentation, measured by the

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

If $LF = 1$, accurate representation

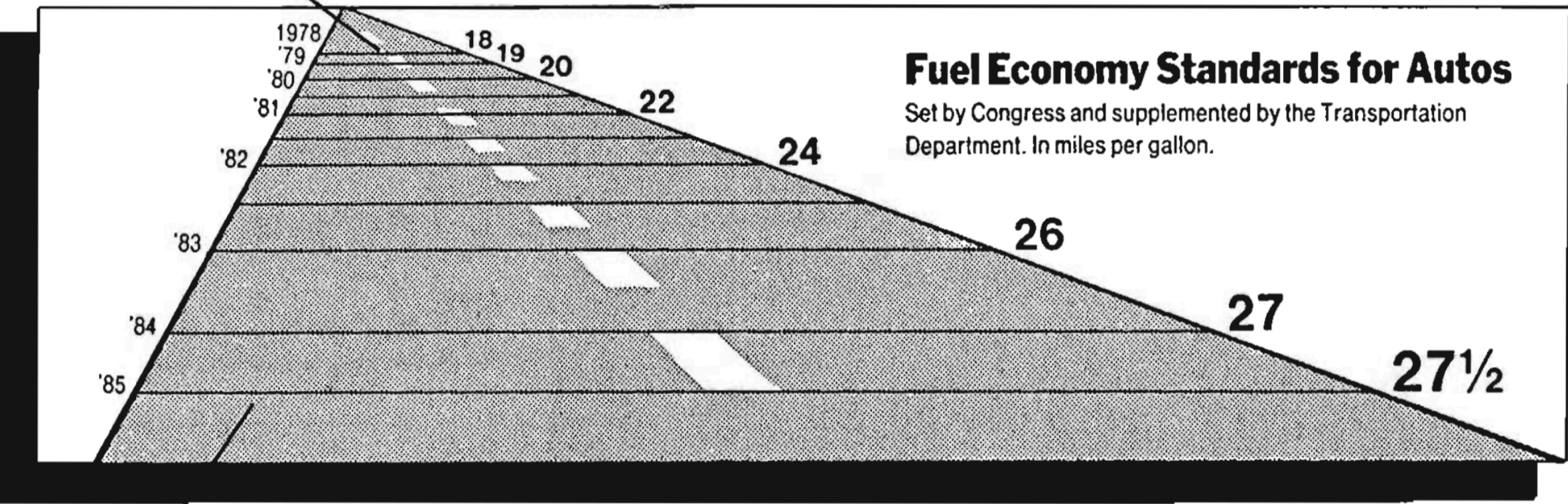
If $LF > 1.05$ or $< .95$, substantial distortion

Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 57)

An Extreme Example

Lie Factor = $\frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

New York Times, August 9, 1978, p. D-2.

$\frac{27.5 - 18.0}{18.0} \times 100 = 53\%$

$\frac{5.3 - 0.6}{0.6} \times 100 = 783\%$

Lie Factor = $\frac{783}{53} = 14.8$

Visual Area and Numerical Measure

Another way to confuse data variation with design variation is **to use areas to show one-dimensional data.**

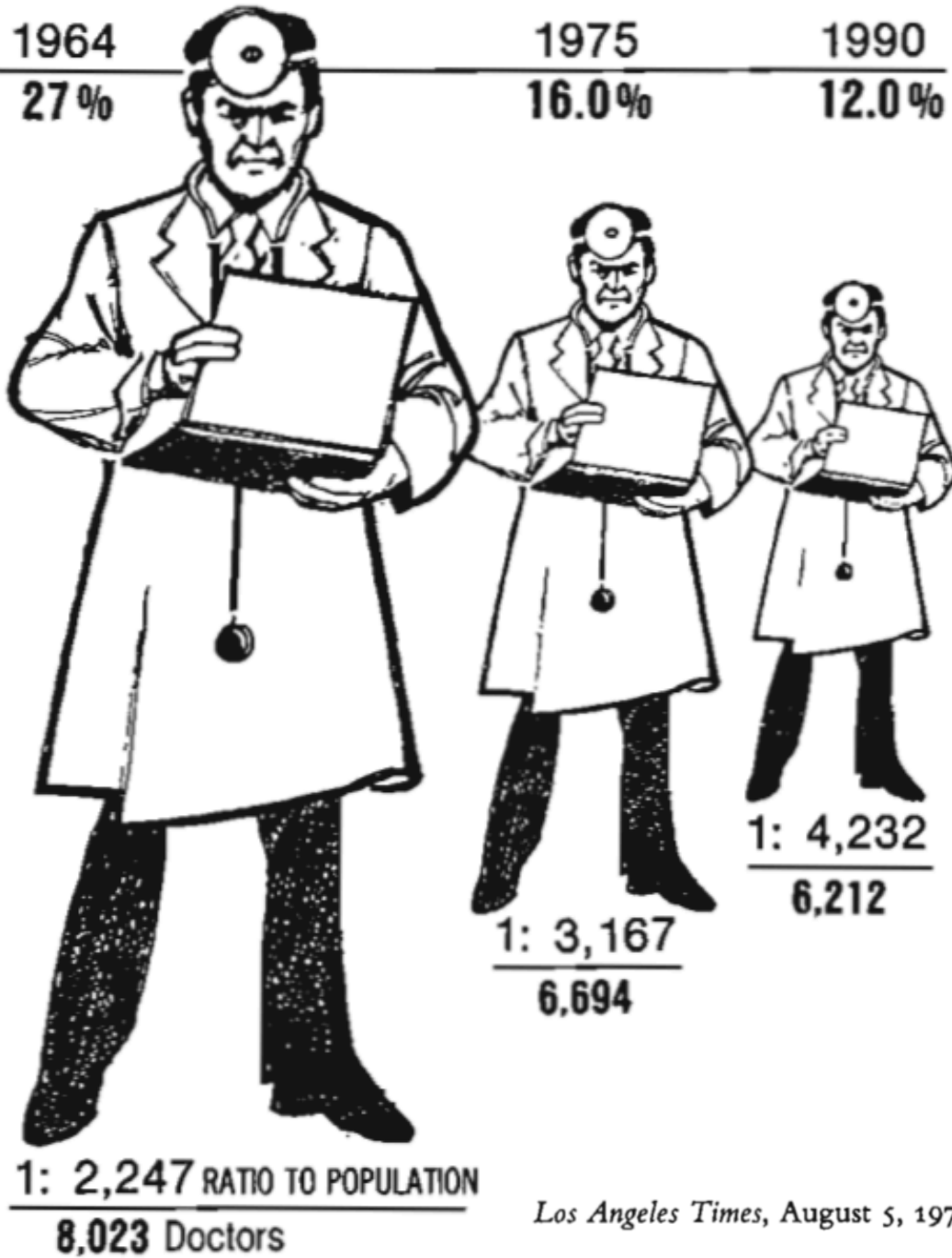
And here is the incredible shrinking doctor, with a lie factor of 2.8, not counting the exaggeration from the overlaid perspective and the incorrect horizontal spacing of data.

Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 69)

THE SHRINKING FAMILY DOCTOR
In California

Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27 %	16.0 %	12.0 %

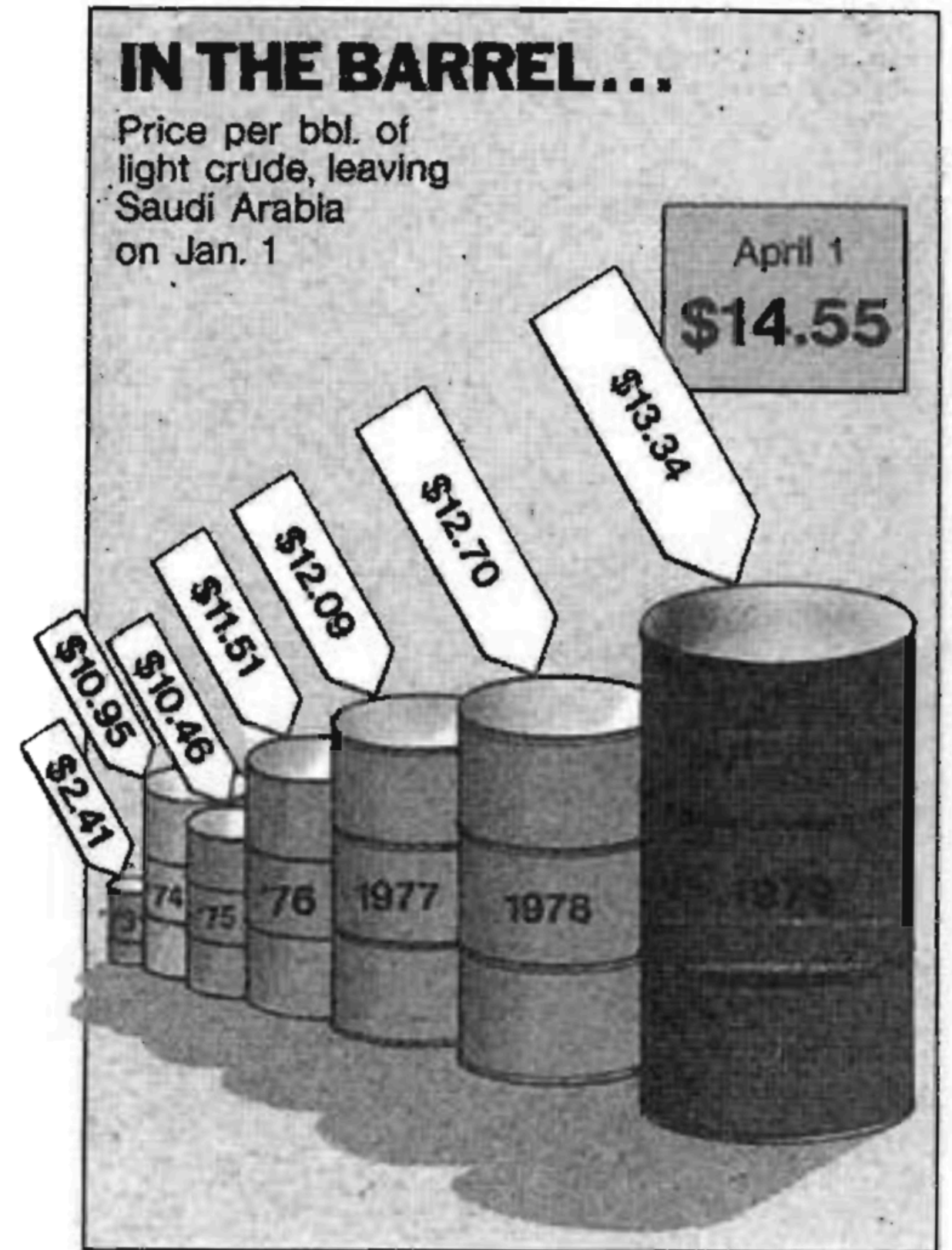


Visual Area and Numerical Measure

There are considerable ambiguities in how people perceive a 2D surface and then convert that perception into a 1D number.

The use of two or three varying dimensions to show 1D data is **a weak and inefficient technique**, often with error in design and ambiguity in perception.

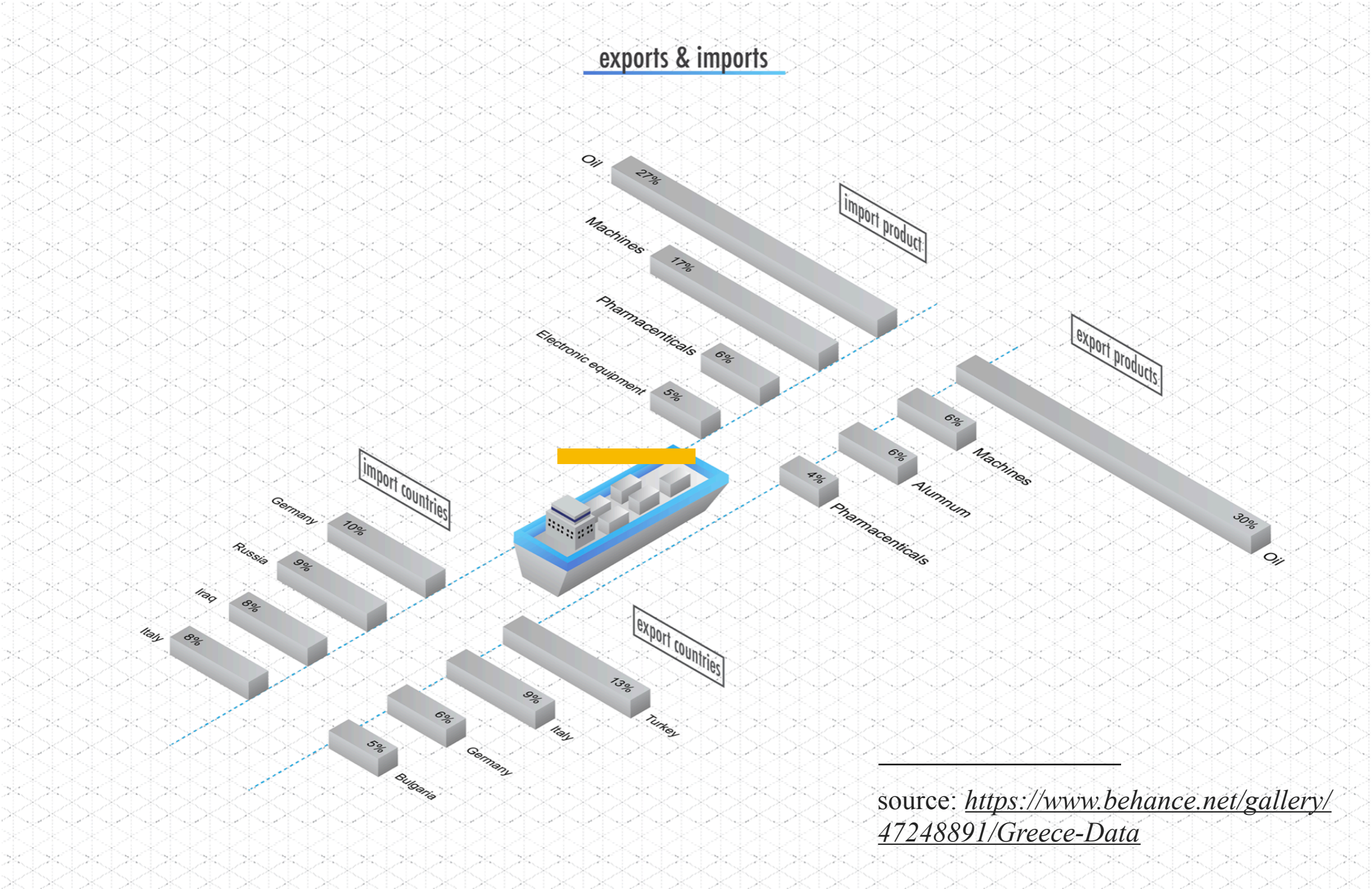
Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 71)



TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

- THE CHOICE OF DESIGN
- MEDIUM DIMENSIONS
- DISTORTION
- LIE FACTOR
- MEDIUM DIMENSIONS
- CHARTJUNK

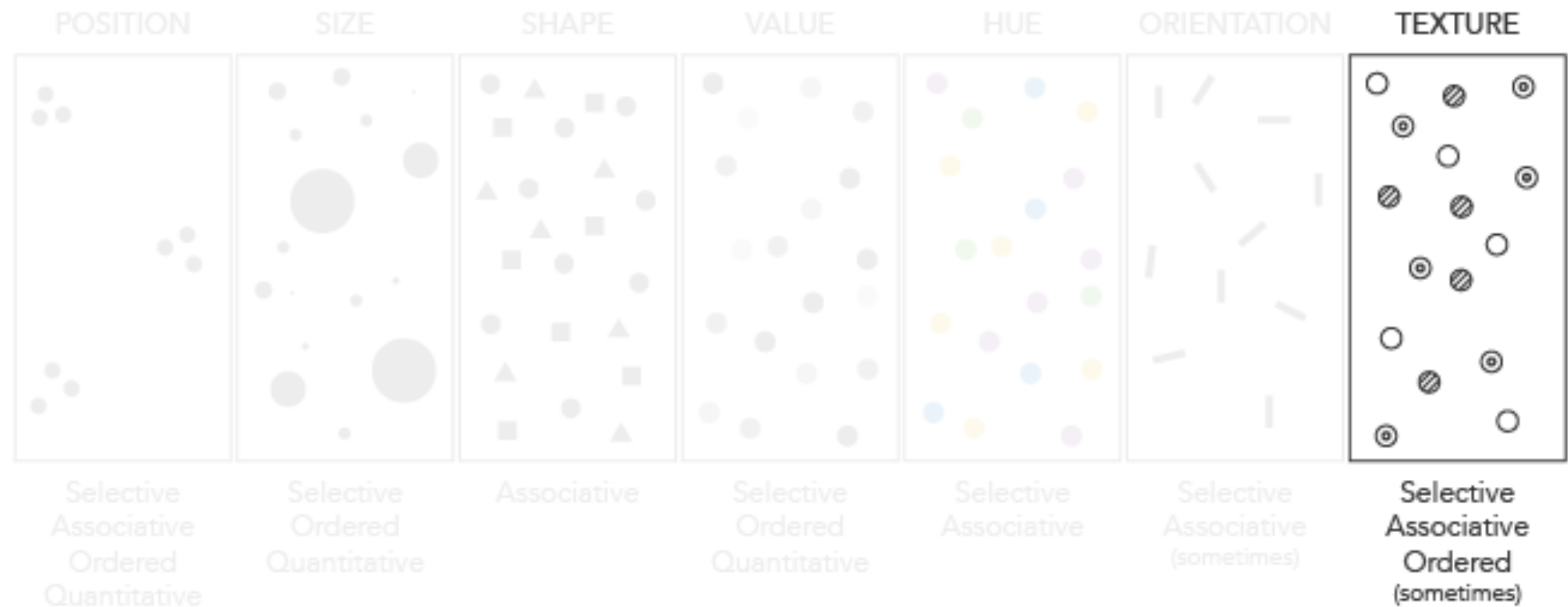
Visual Area and Numerical Measure



source: <https://www.behance.net/gallery/47248891/Greece-Data>

Discussion

Do you think TEXTURE, as one of the visual variables, is somewhat useless or not easy to use?

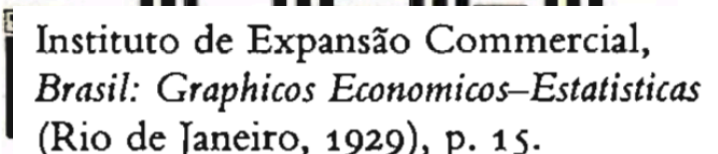


THE CHOICE OF DESIGN

MEDIUM DIMENSIONS

CHARTJUNK

DISCUSSION



TUFTE'S
PRINCIPLES
UNDER 'VISUAL
MAPPINGS'

- THE CHOICE OF DESIGN
- MEDIUM DIMENSIONS
- CHARTJUNK
- DISCUSSION

Discussion

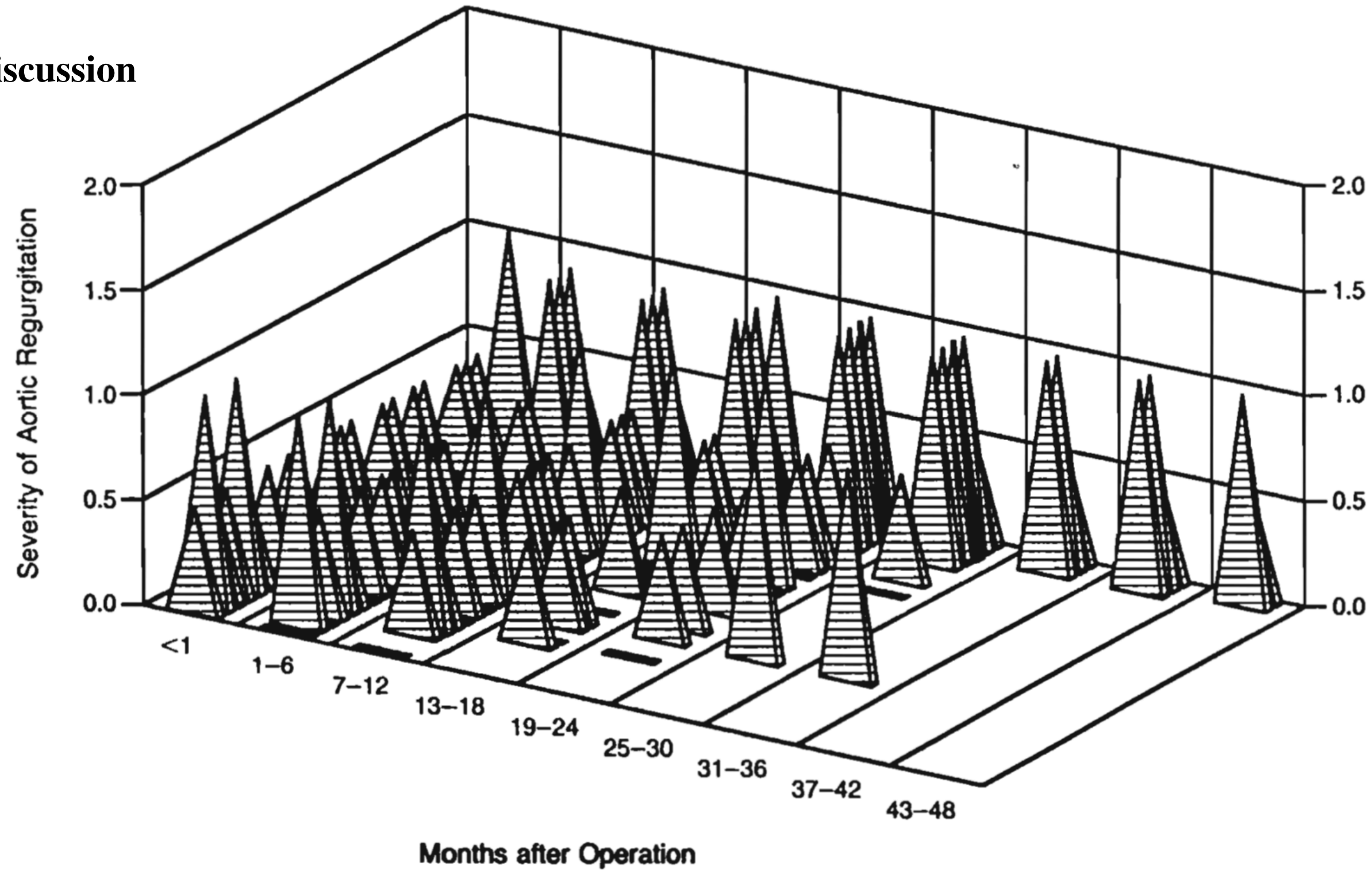
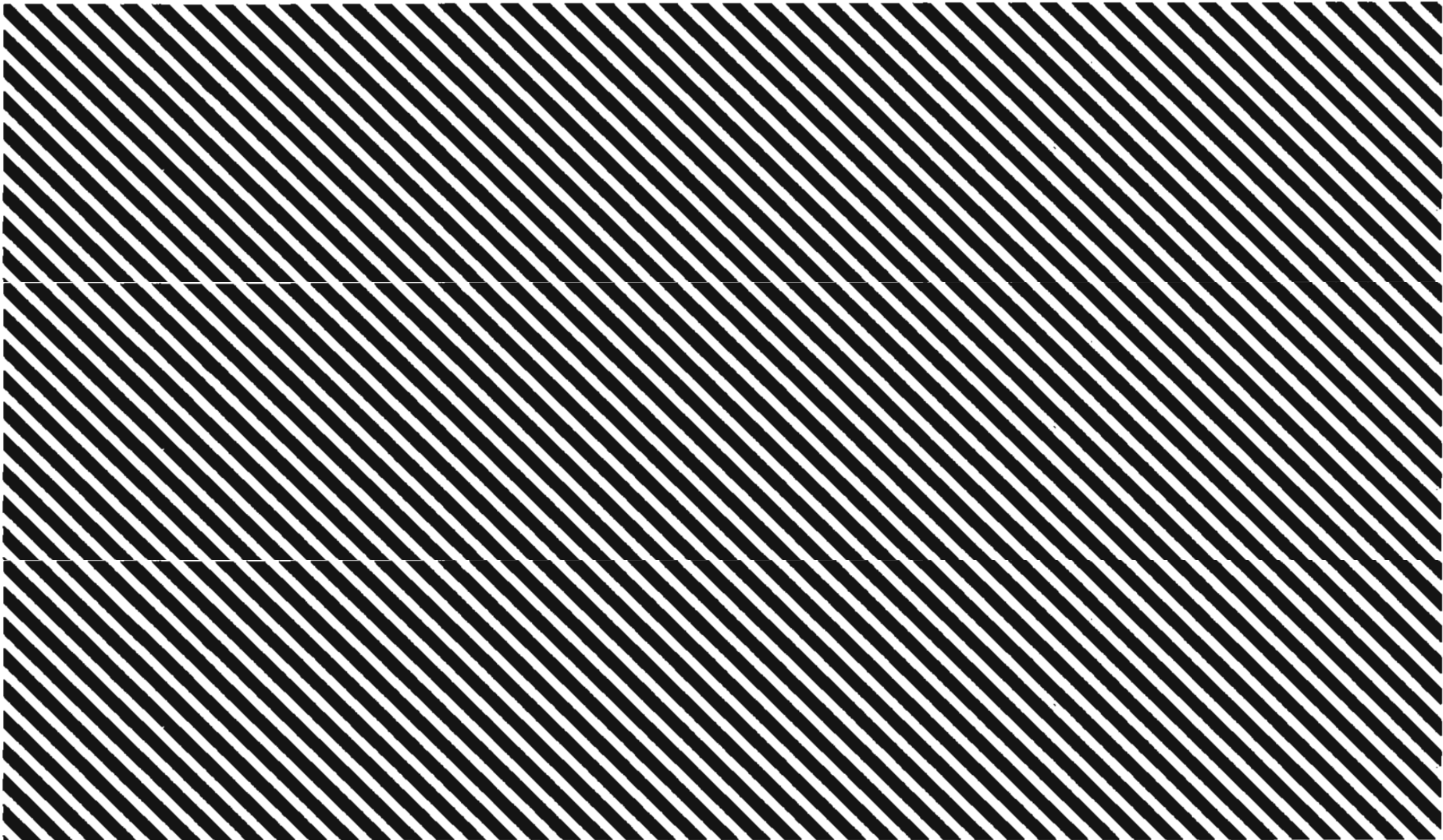


Figure 2. Serial Echocardiographic Assessments of the Severity of Regurgitation in the Pulmonary Autograft in 31 Patients. The numerical grades were assigned according to the severity of regurgitation, as follows: 0, none; 0.5, trivial; 1.0 to 1.5, mild; 2.0, moderate; and 3.0, severe.

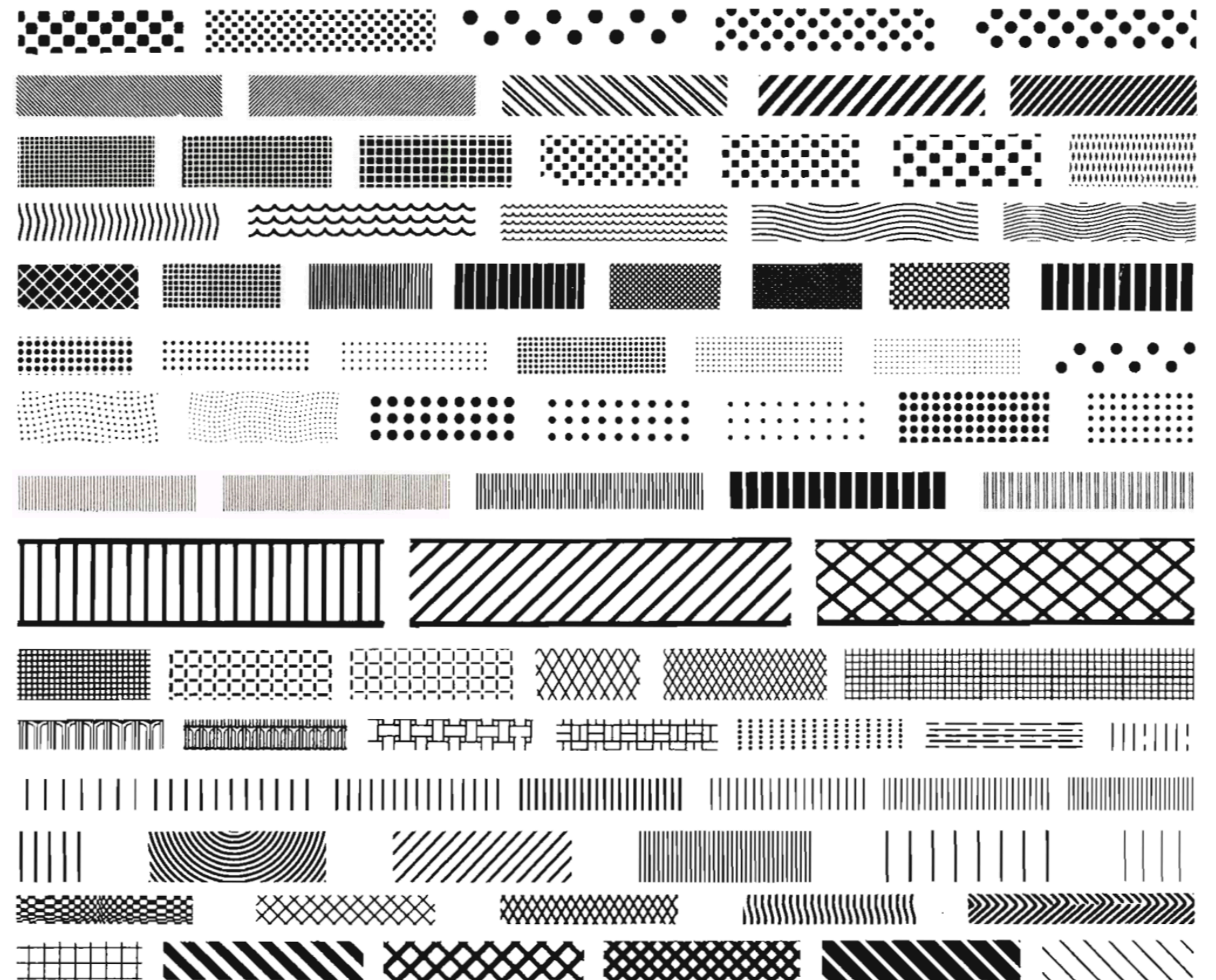
Discussion: Moiré Effects



Discussion

Moire effects have proliferated with computer graphics. Such **unfortunate** patterns were once generated by means of thin plastic transfer sheets; now the computer produces instant **chartjunk**.

Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 111)



Discussion

Can optical art effects ever produce a better graphic? Bertin exhorts: “It is the designer’s duty to make the most of this variation; to obtain the resonance [of moiré vibration] without provoking an uncomfortable sensation: to flirt with ambiguity without succumbing to it.”¹ But can statistical graphics successfully “flirt with ambiguity”? It is a clever idea, but no good examples are to be found. The key difficulty remains: moiré vibration is an *undisciplined* ambiguity, with an illusive, eye-straining quality that contaminates the entire graphic. It has no place in data graphical design.

¹ Bertin, Jacques. *Semiology of graphics; diagrams networks maps*. (Madison, Wisconsin, 1983, translated by William J. Berg), p. 80; this book is the English translation of Bertin’s important work, *Sémiologie graphique* (Paris, 1967).

Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.(p. 112)