

#### PRESENTED TO

#### UBC TABLEAU USERS GROUP

#### 25 JUL 2013

http://www.meetup.com/ubctug/events/122239352/

# analyze as a specialist

# present as a communicator

#### **DESIGN IS ALGORITHMICS FOR THE PAGE**

MARTIN KRZYWINSKI

CANADA'S MICHAEL SMITH GENOME SCIENCES CENTER BC CANCER RESEARCH CENTER

## YOUR AUDIENCE DOES NOT CARE WHAT TOOL YOU USED TO GENERATE A VISUALIZATION

The manufacture process of an object is irrelevant to its user.

## YOUR AUDIENCE DOES NOT CARE WHETHER YOUR VISUALIZATION IS CUTTING EDGE

Novelty and utility are distinct.

#### DATA VISUALIZATION IS VISUAL COMMUNICATION

You are talking with pictures.

But you're still talking.

## The Elements of Style

THE ORIGINAL EDITION

#### WILLIAM STRUNK, JR.

#### POINTS OF VIEW

All rights

2013

#### **Elements of visual style**

Translate the principles of effective writing to the process of figure design.

We all use words to communicate information—our ability to do so is extremely sophisticated. We have large vocabularies, understand a variety of written styles and effortlessly parse errors in real time. But when we need to present complex information visually, we may find ourselves 'at a loss for words', graphically speaking.

We can rationalize figure creation by applying principles of effective written communication. By leveraging our training and experience with words, we can turn graphical improvisation into a structured and reproducible process in which we assess and optimize each part of a figure just as we would each paragraph, sentence and word in a manuscript. Let's look at how Strunk and White's classic but stern *The Elements of Style*<sup>1</sup> can be applied to figures. (I encourage you to revisit your own favorite writing resources in the context of visual representation.)

Figure 1 | A flood of identical symbols triggers semantic satiation, a phenomenon in which overwhelming repetition results in loss of meaning. As an

accurate but visually unparsable representation of a breakpoint graph  $\!\!\!^5$  , the figure breaks Strunk and White's rule "Do not explain too much."  $\!\!\!^1$ 

A popular example of disregarding Strunk and White's dictum "Do not take shortcuts at the expense of clarity"<sup>1</sup> is the syntactically correct but incomprehensible sentence "Buffalo buffalo Buffalo buffalo buffalo Buffalo Buffalo Buffalo"<sup>2</sup>. Unfortunately, visual analogs of this construct appear all too frequently in the literature. If we cannot parse this eight-word sentence, how can we cope with the complexity of Figure 1?

Strunk and White also ask us to avoid overwriting because "rich, ornate prose is hard to digest, generally unwholesome, and



Figure 2 | Use the simplest visual representation<sup>6</sup> for objects and "omit needless words"<sup>1</sup>. (a) Visually garnished elements shout at the reader, who is at a loss to determine what is important. If you wouldn't write it this way, don't draw it either. (b) Simple shapes provide an elegant presentation. Complex shapes may carry unintended meaning (such as unduplicated versus duplicated chromosomes). In schematics, reserve the use of color for emphasis, where possible.



THIS MONTH

Figure 3 | Objects that interact or share common meaning should be formatted in a similar way that appeals to intuition. (a) Venn diagram colors should be selected to naturally communicate overlap. This can be automated by using blend modes in applications such as Illustrator or Inkscape. (b) Entity similarities in pathway diagrams are hard to identify when diverse icons are used. When only tone varies, *FOSL1* immediately stands out from the *FOX* gene family. (c) Symbols in a series should reflect the concept of progression as naturally as possible. For example, immune cells aren't actually a different shape, and it is not intuitive that pink cells should give rise to red cells.

sometimes nauseating"<sup>1</sup>. The visual equivalent is "chartjunk," a term coined by Tufte<sup>3</sup>. Examples are shimmering textures, gradients and a proliferation of shapes (Fig. 2), which all make interpreting the data more difficult, act as exclamation marks that make selective emphasis impossible, and "can never rescue a thin data set"<sup>3</sup>. If you cannot easily emphasize an element in your figure, chances are that it is overstated.

To reinforce the content and function of related ideas, use the visual equivalent of parallel construction and "express coordinate ideas in similar form"<sup>1</sup>. Choose shapes and colors that intuitively embody overlap, category hierarchy and importance (Fig. 3).

Keep in mind the needs and experience of your audience and "place yourself in the background"<sup>1</sup>: do not rely solely on your personal aesthetic (for example, black text overlaid on your favorite color may lack sufficient contrast to be legible). Instead, strive for simplicity and clarity. "Use definite, specific, concrete language"<sup>1</sup>. Be legible without shouting. Concise, but not opaque.

In his play *Horace*, Corneille wrote, "Un premier mouvement ne fut jamais un crime" ("A first impulse was never a crime")<sup>4</sup>. But in the process of making figures, it can be. Avoid the temptation of going with your first idea. Instead, use it as the starting point and then refine and clarify your message. A good figure, like good writing, doesn't simply happen—it is crafted. "Revise and rewrite"<sup>1</sup> becomes "revise and redraw."

#### COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests

#### Martin Krzywinski

- Strunk, W. Jr. & White, E.B. The Elements of Style 4th edn., Ch. 2, 21–26; Ch. 5, 70–75 (Longman, 1999.)
- 2. Pinker, S. The Language Instinct (W. Morrow, New York, 1994).
- Tufte, E.R. The Visual Display of Quantitative Information 2nd edn., 107–121 (Graphic Press, Cheshire, Connecticut, USA, 2001).
- Corneille, P. Horace (http://openlibrary.org/books/OL6939036M/Corneille's\_ Horace/) line 1648 (Heath, 1904).
- Alekseyev, M.A. & Pevzner, P.A. Genome Res. 19, 943–957 (2009).
   Wong, B. Nat. Methods 8, 611 (2011).
- wong, B. wat. Methods 8, 611 (2011).

Martin Krzywinski is a staff scientist at Canada's Michael Smith Genome Sciences Centre.

NATURE METHODS | VOL.10 NO.5 | MAY 2013 | 371

## **TOP-DOWN**

redundancy consistency conciseness clarity focus & emphasis salience & relevance truth, accuracy & detail



#### **BOTTOM-UP**

data encoding symbols color typeface arrows line weight alignment Software users ask

#### HOW DO I MAKE A BAR PLOT?

HOW DO I CHANGE THE BAR COLOR?

Designers ask

#### WHY DO I NEED A BAR PLOT?

#### WHY DID I CHOOSE THIS BAR COLOR?

Communicators ask

#### WHAT AM I TRYING TO SAY?

#### WHAT DOES THE READER WANT TO KNOW?

#### **ARE THE TWO COMPATIBLE?**

#### DO NOT BASE YOUR JUDGMENT ON YOUR SOFTWARE'S OPTIONS

#### USE YOUR JUDGMENT TO STEER YOUR SOFTWARE'S OPTIONS

#### TO EXPLORE DATA, USE EFFECTIVE ENCODING



#### TO COMMUNICATE CONCEPTS, USE EFFECTIVE DESIGN



## **OBJECTIVE ASPECTS OF DESIGN**

It's not all about taste.







## THE MUSEUM OF BAD ART

#### MASTERWORKS



MICHAEL FRANK and LOUISE REILLY SACCO









## Excellent organization and consistency. Vertical lines cue continuity. Good use of color.

Samollow, P.B., The opossum genome: insights and opportunities from an alternative mammal. Genome Res, 2008. 18(8): p. 1199-215.

## Chartjunk plentiful. Screaming ornamental and redundant elements. Text inconsistent and illegible.

Gentles, A.J., et al., Evolutionary dynamics of transposable elements in the short-tailed opossum Monodelphis domestica. Genome Res, 2007. 17(7): p. 992-1004.

#### **JUNK SCIENCE**

#### **REAL SCIENCE**







(left) Synastry chart. http://sasstrology.com/2011/03/the-astrology-of-marriage-in-the-royal-family-a-suitable-girl-and-the-bit-on-the-side.html (right) Shakhnovich, B.E. and E.V. Koonin, Origins and impact of constraints in evolution of gene families. Genome Res, 2006. 16(12): p. 1529-36.

#### **CREATE VISUALS WHEN NECESSARY**

Don't want something you don't need.



### ACCURACY DOES NOT ENSURE COMPREHENSION

Your audience is not a computer.

#### Buffalo buffalo Buffalo buffalo buffalo Buffalo buffalo buffalo

http://en.wikipedia.org/wiki/Buffalo\_buffalo\_Buffalo\_b



New York bison whom other New York bison bully, themselves bully New York bison.



Alekseyev, M.A. and P.A. Pevzner, Breakpoint graphs and ancestral genome reconstructions. Genome Res, 2009. 19(5): p. 943-57.

### **KNOW YOUR MESSAGE**

Stick to it.



#### **ACHIEVE FOCUS BY AGGREGATING**



#### **TELL A STORY**



#### WHERE THERE'S SMOKE—THERE'S CANCER

Cancer rates are up, but mortality is down. New diagnostics and treatments are responsible for part of this trend. But the greatest single contributing factor is the decline in smoking—rates are at their lowest level in 50 years.



## **TELL A STORY**

# <section-header><section-header><section-header>

#### 4 Decline in smoking

Since the 1964 first Surgeon General's report, smoking rates have been dropping. By 2010, the rate among males was down to 20%, from 50% at its peak. Among youth, rates have been on an even steeper decline since 1997.

Smoking prevalence (%)

#### 50adults 25youths in grade 12 0-1965 1990 2010

#### **5** Impact of smoking on cancer deaths

Smoking is a major risk factor for many types of cancer and significant contributor to cancer-related deaths. It remains the single largest preventable cause of disease and premature death in the US.

Percentage of cancer deaths attributable to smoking



## SATISFY YOUR AUDIENCE, NOT YOURSELF.

Be aware of bias in evaluating effectiveness of visual forms.



Linda S Elting, Charles G Martin, Scott B Cantor, Edward B Rubenstein

Response Fail



Poor prognosis

Good prognosis
Response

Eai



Linda S Elting, Charles G Martin, Scott B Cantor, Edward B Rubenstein





Linda S Elting, Charles G Martin, Scott B Cantor, Edward B Rubenstein

"...eight voiced considerable contempt for the [icon] display."



Linda S Elting, Charles G Martin, Scott B Cantor, Edward B Rubenstein

"... icon displays were often preferred by nurses, students, ... but were considered unacceptable by physicians."

#### WE ARE EASILY DECEIVED



#### **DO NOT BE CHARMED BY INEFFECTIVE FORMS**



#### **THE INCOMPARABLE NETWORK**



## **RESPECT HUMAN VISUAL LIMITATIONS**









#### VISUAL ANALYTICS vs VISUAL INTUITICS

Quantitative judgment, isn't.



# 

1 2 3 4 5 6 7







#### YOU ARE AN IMPERFECT MEASURING DEVICE



Heer J, Bostock M (2010) Crowdsourcing graphical perception: using mechanical turk to assess visualization design. Proceedings of the 28th international conference on Human factors in computing systems. Atlanta, Georgia, USA: ACM. pp. 203-212.



## REDUNDANCY

Don't repeat yourself. Don't repeat yourself.



Zinc finger exon analysis for ZNF493 and ZNF738, two divergent genes from the ZNF431 clade. Hamilton, A.T., et al., Evolutionary expansion and divergence in the ZNF91 subfamily of primate-specific zinc finger genes. Genome Res, 2006. 16(5): p. 584-94.

#### **REMOVE TO IMPROVE**

Use the fewest elements possible-keep data-to-ink ratio high.

Shelter your reader from unnecessary complexity.





#### **FOCUS & EMPHASIS**

Match the pertinence of an object with its visual salience.

Apply visual organization Gestalt principles.





#### **EVERYTHING IS EMPHASIZED**



#### **NOTHING IS EMPHASIZED**

![](_page_60_Figure_1.jpeg)

#### **MATCH VISUAL SALIENCE TO RELEVANCE**

#### **ATTENTIONAL CAPTURE**

![](_page_61_Figure_2.jpeg)

SALIENCE HIGH LOW

Fecteau JH, Munoz DP (2006) Salience, relevance, and firing: a priority map for target selection. Trends Cogn Sci 10: 382-390. Yantis S (2005) How visual salience wins the battle for awareness. Nat Neurosci 8: 975-977.

## **USE BREWER PALETTES**

![](_page_62_Figure_1.jpeg)

![](_page_62_Figure_2.jpeg)

Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, et al. (2003) Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Res 13: 2725-2735.

Bono H, Yagi K, Kasukawa T, Nikaido I, Tominaga N, et al. (2003) Systematic expression profiling of the mouse transcriptome using RIKEN cDNA microarrays. Genome Res 13: 1318-1323.

Tenney AE, Wu JQ, Langton L, Klueh P, Quatrano R, et al. (2007) A tale of two templates: automatically resolving double traces has many applications, including efficient PCR-based elucidation of alternative splices. Genome Res 17: 212-218.

## **HIERARCHY AND PRIORITY**

Use symbols that intuitively encode related concepts.

![](_page_64_Figure_0.jpeg)

![](_page_65_Figure_0.jpeg)

![](_page_66_Figure_0.jpeg)

![](_page_67_Figure_0.jpeg)

![](_page_68_Picture_0.jpeg)

![](_page_69_Picture_0.jpeg)