The Future of Data: Too Much Visualization, Too Little Understanding?

An empirical investigation into the comprehension of data visualizations

RENÉ REITSMA¹ AND ANDREA MARKS²

1. Oregon State University, Corvallis, OR, USA
2. Oregon State University, Corvallis, OR, USA

SUGGESTED CITATION: Reitsma, R. & Marks, A. "The Future of Data: Too Much Visualization, To Little Understanding."
Dialectic, 2.2 (2019): pgs. 109-130. Published by the AIGA Design Educators Community (DEC) and Michigan Publishing.
DOI: <http://dx.doi.org/10.3998/dialectic.14932326.0002.207>

Abstract

Data visualization can be useful in enhancing and understanding the patterns and stories behind given data. For at least a decade, however, we have observed a disturbing trend of designers complicating and embellishing visualizations of essentially very simple data sets to the point of obscuring the essential information they were designed to convey. To gauge some of the effects of this practice on how well audiences comprehend the data underlying these types of visualizations, we conducted an experiment that entailed the selection of five examples from readily available and reputable published sources. These ranged from widely read newspapers and magazines to online news outlets to scholarly journals. We then developed simpler visualizations that utilized the identical data as those that we deemed to be overly complex, and randomly exposed subjects to either of these visualizations to assess their ability to effectively perceive the displayed data. The subjects were 72 junior and senior undergraduate students at a US land-grant university. We found that, on average, it took students *longer* to comprehend the complex versions, that they took *longer* to extract information from them, and that they made *larger* errors of interpretation as they did so. The experiment provided evidence that subjects eventually did learn to navigate the more complex versions of the data visualizations, but this took them significantly more time. This caused us to question whether the discipline of data visualization is trending toward a future that will challenge audiences to interpret too many complex visualizations of data that yield too little understanding.

The Future of Data: Too Much Visualization, Too Little Understanding?

An empirical investigation into the comprehension of data visualizations

RENÉ REITSMA & ANDREA MARKS

1

Lima, M. Visualizing Complexity. Mapping Patterns of Information. New York, NY, USA: Princeton Architectural Press, 2011.

2

Börner, K. Atlas of Science: Visualizing What We Know. Cambridge, MA, USA: The MIT Press, 2010.

3

Börner, K. Atlas of Knowledge; Anyone Can Map. Cambridge, MA, USA: The MIT Press, 2015.

4

Meirelles, I. Design for Information. An introduction to the histories, theories, and best practices behind effective information visualizations. Beverly, MA, USA: Rockport Publishers, 2013.

5

Börner, K. & Polley, D.E. VISUAL INSIGHTS. A Practical Guide to Making Sense of Data. Cambridge, MA, USA: The MIT Press, 2014.

6

SciAm Graphic Science. Scientific American. Online. Available at <https://www.scientificamerican.com/department/graphic-science/> (Accessed June 12, 2018).

7

Börner, K. Atlas of Knowledge; Anyone Can Map. Cambridge, MA, USA: The MIT Press, 2015.

Introduction

We live in a time in which we are bombarded with data and attempts to visualize it, from mainstream newspapers and magazines to academic journals and industry publications. These visualizations occur in both digital and printed media. A cursory look at compendia such as Lima's *Visualizing Complexity*,¹ Börner's *Atlas of Science*² and *Atlas of Knowledge*,³ information design textbooks such as those by Meirelles⁴ or Börner & Polley,⁵ or *Scientific American*'s monthly *Graphic Science*,⁶ shows the evolution and growth in the field of data visualization over the past decade. That some of these works are published as coffee table books is indicative of the 'wow' effect that many modern data visualizations evoke among broad readerships. They are indeed impressive in their presentation of often complex data. Moreover, as indicated by the subtitle of Börner's 2015 *Atlas of Knowledge—Anyone Can Map*⁷—data visualization software has become so powerful and easy to use that complex visualizations are no longer the explicit domain of professional designers. They have become part of the mass media or, as Peter Hall so poignantly states in his essay "*Bubbles, Lines and String: How Information Visualization Shapes Society*:" "data visualization has lately become an unlikely form of mass entertainment."⁸

Still, while marveling at the complexity and aesthetics of these visualizations, especially those of complex networks or mashups of large amounts

8

Hall, P. "Bubbles, Lines and String: How Information Visualization Shapes Society." In Graphic Design: Now in Production edited by A. Blauvelt and A. Lupton, pgs. 170-185. Minneapolis, MN, USA: Walker Art Center, 2011.

9

Ibid.

10

Cheng, K. "Designing Data: Issue 31.4 Feature Introduction." ARCADE. 31.4 (2013). Online. Available at: <http://arcadenw.org/article/designing-data> (Accessed June 12, 2018).

11

Pettersson, R. Information Design. An Introduction. Amsterdam, NL and Philadelphia, PA, USA: John Benjamin's Publishing Company, 2002.

12

Reitsma R., Hsieh, P.-H., Diekema, A., Robson, R. and Zarsky, M. "Map- or List-based Recommender Agents? Does the Map Metaphor Fulfill its Promise?" Information Visualization, 16 (2016): pgs. 291-308.

13

Fabrikant, S.I., Montello, D.P. and Mark, D.M. "The Natural Landscape Metaphor in Information Visualization: The Role of Commonsense Geomorphology." Journal of the American Society of Information Science and Technology, 61 (2009): pgs. 233-270.

14

Kelleher, C. & Wagener, T. "Ten guidelines for effective data visualization in scientific publications." Environmental Modelling & Software, 26 (2011): pgs. 822-827.

of varying types of data, one may wonder how meaningfully, effectively and efficiently these visualizations can be interpreted (See Figure 1; Column 'Complex'). Do they, for instance, support the oft-made claim that they should help us see or discover patterns?⁹⁻¹¹ Even if they are merely meant to illustrate or exemplify, do they actually serve the purpose of telling the story that their creators want their audiences to distill or interpret from the data? One only has to review some of the many visualizations in Lima's and Börner's otherwise excellent compendia to develop a feeling of "great, but what can we really do with these?"

Interested in the usability of these visualizations, Reitsma *et al.*¹² reviewed more than 80 of the examples showcased in Lima's text and discovered that only a tiny percentage had been submitted to any type of usability testing to assess their efficacy. Moreover, if any testing was done, it was mostly of an anecdotal nature, was conducted with very small samples of subjects, and almost always involved only qualitative measures such as free-form opinions and perspectives. This lack of usability testing made Fabrikant *et al.* lament that "*most others who present information landscapes do not even bother to outline why they think that the displays they create might be effective.*"¹³ As for reasons of emphasizing aesthetics over functionality, Kelleher & Wagener offer: "*the visualization ability of software packages can tempt the user to use impressive rather than sensible plots.*"¹⁴ Along similar lines, Pettersson eloquently warns us against the dangers of form trumping function when he writes: "*At some point illustrations move from being engaging motivators to engaging distractors.*"¹⁵

Much empirical work has been done on assessing the efficacy of visual primitives such as lines, points, polygons, line width, color, symbol graduation, etc. For an overview, we refer to Schonlau and Peters.¹⁶ Modern data visualizations, however, continue to implement certain visual primitives which have repeatedly been shown in visual perception studies to be deficient. A typical example of this is graduated symbols inhibiting the accurate perception of corresponding metrics. Cartographic studies from as early as the 1950s have shown that people are not good at estimating relative magnitudes when they are presented with graduated circles, squares, rectangles and triangles.¹⁷⁻²⁵ Despite this evidence, in a wide variety of modern data visualizations, graduated circles are commonly used. Another example of the perceptual damage that can be wrought when visual primitives are used without caution is the now-all-too-common use of (stacked) area charts, about which Thudt *et al.* offer the following criticism: "*...despite the concerns over their readability, stacked graphs*

15

Pettersson, R. Information Design: An Introduction. Amsterdam, NL and Philadelphia, PA, USA: John Benjamins Publishing Company, 2002.

16

Schonlau, M. & Peters, E. "Comprehension of Graphs and Tables Depends on the Task: Empirical Evidence from two web-based studies." Statistics, Politics and Policy, 3 (2012). Article 5

17

Williams, R.L. Statistical Symbols for Maps; Their Design and Relative Values. New Haven, CT, USA: Yale University, 1956.

18

Chang, K.T. "Visual Estimation of Graduated Circles." Canadian Cartographer, 14 (1977): pgs. 130-138.

19

Cox, C.W. Anchor "Effects and the Estimation of Graduated Circles and Squares." American Cartographer, 3 (1976): pgs. 65-74.

20

Crawford, P.V. "Perception of Grey-tone Symbols." Annals of the Association of American Geographers, 61 (1971): pgs. 721-735.

21

Crawford, P.V. "The Perception of Graduated Squares as Cartographic Symbols." Cartographic Journal, 10 (1973): pgs. 85-88.

22

Flannery, J.J. "The Relative Effectiveness of Some Common Graduated Point Symbols in the Presentation of Quantitative Data." Canadian Cartographer, 8 (1971): pgs. 96-109.

have aesthetic value, leading to widespread use on the web."²⁶ Schonlau and Peters provide an extensive overview of the research that studies the relationships between graphic primitives and data interpretation in work they published in 2012. They concluded that comprehension of data displayed in web-based graphic structures among given readerships was largely affected by the comprehension tasks they were being challenged to undertake, but that visual displays should not be used in situations that might cause them to impair comprehension.²⁷ While these examples from the extant body of research examine the relative efficacy of utilizing specific types of visual displays to enhance a given reader's understanding of comparable data, they tend not to account for how reader comprehension of visualized data, or data sets, is affected by more complex graphic presentations of imagery, typographic elements and the spaces between them.

The content of this article adds to the scholarly literature that calculates the cost and consequences of deviating from Pettersson's advice to "*Keep it simple!*,"²⁸ Tufte's recommendation to "*create the simplest graph that conveys the information you want to convey*,"²⁹ Tversky's advice of "*eliminating the information that clutters and distracts*,"³⁰ and Bigwood and Spore's conclusion that "*Less is often more, especially in communication*."³¹ Some examples of empirical work that assesses the effects of deviating from this principle can be found in Haemer,³² Wilkinson³³ and Schonlau & Peters.³⁴ Similarly, comparative research conducted by Bateman *et al.*³⁵ and Skau *et al.*³⁶ determined that embellishments not essential to a given readership's ability to compare data sets do not result in better comprehension.

Expected Effects of Unnecessarily Complicating the Graphic Display of Quantitative Visualizations

In hypothesizing the likely effects of excessively complex visualizations of data on reader comprehension, we follow the work of Javed *et al.*³⁷ in positing that the two most obvious variables that must be accounted for are the time necessary for readers to extract information from a given visualization (including the time needed for its initial comprehension), and the accuracy of the information-cum-knowledge readers were able to extract from it. To these we add a third effect, namely a reader's perceived (or stated) comprehension of the visualization. Considering the relationship between what we consider to be unnecessary graphic complication (such as the addition of extra layers of texture, profuse variations in line weights and gradients, decorative elements

^a Raluca Budiu of the Nielsen Norman Group offers the following explanation of a “between-subjects vs. within-subjects” study design. “When you want to compare several user interfaces in a single study, there are two ways of assigning your test participants to these multiple conditions:

- Between-subjects (or between-groups) study design: different people test each condition, so that each person is only exposed to a single user interface.
- Within-subjects (or repeated-measures) study design: the same person tests all the conditions (i.e., all the user interfaces).

(Note that here we use the word “design” to refer to the design of the experiment, and not to website design.)

For example, if we wanted to compare two car-rental sites A and B by looking at how participants book cars on each site, our study could be designed in two different ways, both perfectly legitimate:

1. Between-subjects: Each participant could test a single car-rental site and book a car only on that site.
2. Within-subjects: Each participant could test both car-rental sites and book a car on each.

Any type of user research that involves more than a single test condition has to determine whether to be between-subjects or within-subjects. However, the distinction is particularly important for quantitative studies. This explanation can be found online at: <https://www.nngroup.com/articles/between-within-subjects/> (Accessed February 20, 2019.)

such as illustrations and symbols, the assertion of a wide variety of typographic styles, fonts, weights and sizes) and our three dependent variables, we hypothesize that unnecessarily complex visualizations require more time for readers to comprehend; that they require more of the reader’s time to extract information; and that they contribute to a higher level of erroneous information extraction. Whereas we consider it likely that these relationships will be confirmed in experimentation, we are also interested in the magnitude of these effects; i.e., just how “bad” are the reader comprehension errors that occur as a result of unnecessarily complicating data visualizations, and how much extra time do their readers spend trying to discern essential meaning from them?

An Overview of the Experiment We Conducted to Test Our Hypotheses

To assess these hypotheses, we devised an experiment with both between-subjects and within-subjects characteristics,^a in which subjects were shown either graphically complex visualizations or ones that were graphically more simplified; both of these depicting identical data sets. The data sets we utilized were derived from a renowned scholarly journal, a well-respected newspaper, a mainstream popular science magazine and a governmental agency of a G8 nation (see Figure 1). We designed this experiment as we did because we were interested in examining the aggregate effects these two types of visualizations would have on 1) reader comprehension and 2) time-necessary-to-comprehend. Designing our experiment in a manner that incorporated graphically complex diagrams that had been published in well-vetted publications enabled us to avoid having to manipulate the configuration of extant diagrams that were less graphically complex into ones that were more so.

Figure 1 depicts the five examples of information visualization selected for use in our experiment: two were published in the periodical *Scientific American*; one was published in *The New York Times*, one was published in *The New England Journal of Medicine*, and one was published in the book *Visualizing Complexity. Mapping Patterns of Information*³⁸ by Manuel Lima. All five visualize empirical data: weather data, CO2 emissions data, poverty incidence data, science citation data and obesity data. The data depicted in these examples are quite simple in nature: 2D time series data, 2D point data or 2D tabular data. However, rather than visualizing the data by means of more conventional means and techniques of depiction such as basic line graph, point graph and data tables, the designers of each of the graphically more complex examples chose to visualize the data in graphically more unique and visually appealing

23

Groop, R.E. & Cole, D. "Overlapping Graduated Circles, Magnitude Estimation and Method of Portrayal." Canadian Cartographer, 15 (1978): pgs. 114–122.

24

Reitsma, R. & Trubin, S. "Information Space Partitioning Using Adaptive Voronoi Diagrams." Information Visualization, 6 (2007): pgs. 123–138.

25

Heer, J. & Bostock, M. "Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design." in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2010. Association for Computing Machinery, 2010. pgs. 203–212. Online. Available at <http://vis.stanford.edu/files/2010-MTurk-CHI.pdf> (Accessed June 12, 2018).

26

Thudt, A., Walny, J., Perin, C., Rajabiyazdi, F., MacDonald, L., Vardeleon, R., Greenberg, S. and Carpendale, S. "Assessing the Readability of Stacked Graphs." in Proceedings of the 2016 Graphics Interface Conference. Victoria, BC, Canada. 2016. Online. Available at <http://innovis.cpsc.ucalgary.ca/innovis/uploads/Publications/Publications/ThudtStreamgraphs2016.pdf> (Accessed June 12, 2018).

27

Ibid.

28

Pettersson, R. Information Design: An Introduction. Amsterdam, NL and Philadelphia, PA, USA: John Benjamin's Publishing Company, 2002.

29

Tufte, E.R. The Visual Display of Quantitative Information. Cheshire, CT, USA: Graphics Press, 1983.

a.

The intent of this information graphic by *The New York Times* was to visually communicate that 2016 was a warmer-than-normal year. The diagram shows six minimum and maximum daily temperature curves (actual, record and normal minimum and maximum) for New York City. (The original also depicted precipitation data, which was omitted as a means to make the perception of this diagram less difficult for our test subjects to comprehend).³⁹

b.

We replaced the six original daily minimum and maximum temperatures with a daily temperature differential (it communicates how much warmer or colder than normal the temperature was during a given span of time). This reduced the amount of plotted data by 83%, and also resulted in the bifurcation of the graph into two sections: one that depicts "warmer than normal" data (positive/red,) and another that depicts "colder than normal" data (negative/blue).

c.

This example depicts an alternative, simplified means of portraying this data. It plots the daily temperatures from 2016 against the set of "normal" daily temperatures recorded from previous years. A comparison between this diagram and the more graphically complex original reveals that the plotted data has been reduced by 67%.

d.

The designer of this visualization of data about Italian poverty took the original tabular data from the *Italian Statistics Bureau* and indicated geographic (dis)aggregation by means of an alluvial diagram.^{40, 41}

e.

Table 1. Poverty in Italy in 2006

	Percentage of households in poverty
North	5.2%
Central	6.9%
South	22.6%
Italy	11.1%

Table 2. Poverty in Italy in 2006 for different household sizes

	Household size				
	1 member	2 members	3 members	4 members	5 or more members
North	4.8%	5.4%	4.4%	6.2%	8.1%
Central	4.3%	7.5%	7.1%	7.2%	15.4%
South	17.1%	20.3%	20.9%	26.1%	37.5%
Italy	8.1%	9.9%	10.0%	14.8%	24.3%

Data from the original report on Italian poverty incidence by the *Italian Statistics Bureau*.⁴²

FIGURE 1: The examples of more and less graphically complex depictions of the same data sets we used during the operation of our experiment. These appear in the "Complex" and "Simple" columns, respectively. (Continues on pg 117)

30

Tversky, B. "Diagrams. Cognitive Foundation for Design." in *Information Design* edited by A. Black, P. Luna, P., O. Lund and S.E. Walker, pgs. 349–360. London, UK & New York, NY, USA: Routledge, 2017.

31

Bigwood, S. & Spore, M. "When to use numeric tables and why." in *Information Design* edited by A. Black, P. Luna, P., O. Lund and S.E. Walker, pgs. 503–507. London, UK & New York, NY, USA: Routledge, 2017.

32

Haemer, K. "The Pseudo Third Dimension." *The American Statistician*, 5.28 (1951).

33

Wilkinson, L. "Less is More: Two-and Three-Dimensional Graphics for Data Display." *Behavior Research Methods, Instruments and Computers*, 26 (1994): pgs. 172–176.

34

Ibid.

35

Bateman, S., Mandryk, R.L., Gutwin, C., Genest, A., McDine and D. and Brooks, C. "Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2010*. Association of Computing Machinery, 2010: pgs. 2573–2582. Online. Available at <http://dmrusell.net/CHI2010/docs/p2573.pdf> (Accessed June 12, 2018).

^b Association to Advance Collegiate Schools of Business

ways. This can be seen in Figure 1, which shows both the graphically more complex published version of each of these five information diagrams (in the "Complex" column), and the graphically simplified versions that were developed by us to depict the identical data (in the "Simple" column). The data we used to design the graphically simpler versions were either publicly available or were supplied to us by the owners of the data from which the complex versions were generated.

Each subject who participated in this experiment was shown—via random selection—either the graphically complex or simple depictions of each of the five data sets. Prior to being prompted to extract information from each visualization, subjects were given a short explanation of the data it visualized. These explanations were identical regardless of whether a given subject was presented with graphically complex or simple version, although in some cases wording had to be slightly modified in order to properly refer to the diagram; e.g., 'diagram' vs. 'table.' Along with this introduction, subjects were asked to quantify how well they understood the example with which they were being presented by selecting one answer from among three: "Yes, I understand;" "I am not sure if I understand;" "No, I do not understand." Subjects were then asked to respond to several prompts presented in survey form about the data displayed in each of the diagrams. Table 1 lists these prompts and the associated response scales.

The experiment was administered as a web-based, *Qualtrics*TM survey, and the survey tool excluded the use of smartphones. Subjects were instructed to complete the survey at a location and time of their choice during a given span of time.

A Description of the Experimental Subject Pool

The subjects of this experiment were junior and senior (years three and four) undergraduate Business students majoring in Business Information Systems and/or Accounting and/or Finance at an AACSB-accredited ^b Business school in the United States. They were offered a small amount of extra class credit for participating in the survey. 72 students completed it. We acknowledge that although this pool of subjects is not necessarily representative of the readership of each of the visualizations tested here, undergraduate students are part of the normal readership of broadly published media such as *Scientific American* and the *New York Times*, as well as academic journal articles and governmental reports. University students are also used to and trained in the process of

36

Skau, D., Harrison, L. and Kosara, R. "An Evaluation of the Impact of Visual Embellishments in Bar Charts." in Proceedings Eurographics Conference on Visualization (EuroVis) 2015. Online. Available at <https://kosara.net/papers/2015/Skau-EuroVis-2015.pdf> (Accessed June 12, 2018).

37

Javed, W., McDonnell, B. and Elmqvist, N. "Graphical Perception of Multiple Time Series." IEEE Transactions on Visualization and Computer Graphics, 16 (2010): pgs. 927–934.

38

Lima, M. Visualizing Complexity. Mapping Patterns of Information. New York, NY, USA: Princeton Architectural Press, 2011.

39

Lai, KK Rebecca. "How Much Warmer Was Your City in 2016" New York Times online. January 18, 2017. Online. Available at <https://www.nytimes.com/interactive/2017/01/18/world/how-much-warmer-was-your-city-in-2016.html#nyc> (Accessed June 12, 2018).

40

Porpora, M. Italian Poverty Read Thread, 2009. Online. Available at <https://www.behance.net/gallery/463228/Italian-Poverty-Red-Thread>. (Accessed June 12, 2018).

41

Lima, M. Visualizing Complexity. Mapping Patterns of Information. New York, NY, USA: Princeton Architectural Press, 2011.

42

Istat. "La povertà relativa in Italia nel 2006." Statistiche in breve. Istituto Nazionale di Statistica. Roma. Online. Available at <http://www.istat.it/it/files/2011/03/poverta06.pdf> (Accessed June 12, 2018).

Complex

This graph published in *Scientific American* represents CO2 emissions by country (X-axis: total emissions; Y-axis, emissions per capita). All data points are plotted twice: once as a point in 2D space, and again as graduated and colored circles.⁴³

Simple

This simpler version of the CO2 emissions data plots each data point only once, thereby reducing the amount of plotted data by 50%. The data depicted in this example is gleaned from *The World Resources Institute*.⁴⁴

h.

Data on the historic citation frequencies of 15 of science papers are plotted as a vertically oriented time series of graduated circles in this diagram published by *Scientific American*.⁴⁵

i.

This simpler version of the citation data depicted in the diagram at left uses a conventional time series line graph with time depicted on the X-axis, and citation frequency on the Y-axis. Data from Ke et al.⁴⁶

j.

This is a *New England Journal of Medicine* diagram depicting a social network of 2,200 people. The size of a bubble represents each person's body mass index (BMI). The color of each bubble indicates whether or not a person is overweight. The colors of the lines indicate the type of relationships between individuals documented in this study.⁴⁷

k.

	Table 1: Likelihood of becoming obese based on personal and relationship characteristics		
	Relationship is a friend	Relationship is a spouse	Relationship is a parent, child, or sibling
Person's relation is obese	+12%	+37%	+40%
Person's age	-1% per year of age	-1% per year of age	0% per year of age
Person's years of education	-1% per year of education	-2% per year of education	-4% per year of education

This table depicts associations between people's relationships with specific types of others and their chances of becoming obese. We are unsure why the network diagram (j) was included in the original paper from which this data table was derived, as it is essentially impossible to gauge these associations from it.

Table 1: Survey questions and answers

Case	Question*	Answer Scale
Was 2016 a warm year?	Indicate your level of agreement with the following statement: <i>2016 was a warmer than normal year in New York City.</i>	Five-point Likert scale (strongly disagree, ..., strongly agree)
	Fill in the blank in the following statement: _____ percent of days in 2016 in New York City were warmer than normal.	Fill in the blank
	Indicate your level of agreement with the following statement: <i>May 2016 in New York City had more warmer-than-normal days than colder-than-normal days.</i>	Five-point Likert scale (strongly disagree, ..., strongly agree)
Italian poverty	Fill in the blank in the following statement: _____ percent of households in the Center of Italy are poor.	Fill in the blank
	Fill in the blank in the following statement: _____ percent of three-member households in the North of Italy are poor.	Fill in the blank
	Fill in the blank in the following statement: _____ percent of households with five or more members in Italy as a whole are poor.	Fill in the blank
	Indicate your level of agreement with the following statement: <i>In Italy, in 2006, households with more members have a higher likelihood of being poor than households with fewer members.</i>	Five-point Likert scale (strongly disagree, ..., strongly agree)
CO ₂ emissions by country	Fill in the blank in the following statement: <i>China's 2012 total emissions were about _____ times the size of those by India.</i>	Fill in the blank
	Fill in the blank in the following statement: <i>China's 2012 total emissions were about _____ times the size of those by the USA.</i>	Fill in the blank
Sleeping beauties	Which paper accumulated the most citations during its lifetime?	Multiple choice; list of all 15 papers
	Which paper enjoyed the fastest growth in number of citations?	Multiple choice; list of all 15 papers
	Fill in the blank in the following statement: Complex: In 2011 paper No 2 received approximately _____ times the number of citations of paper No 13. Simple: In 2011 the paper by Hummers & Offeman received approximately _____ times the number of citations as the paper by Pugh (1954).	Fill in the blank
Obesity network		Multiple choice: likelihood increases / decreases / is not affected / cannot tell)
	Is a person's likelihood of becoming obese affected by with whom (s)he is connected in his/her social network?	Multiple choice: likelihood increases / decreases / is not affected / cannot tell)
	Indicate your level of agreement with the following statement: <i>Regardless of people's social network, the likelihood that people become obese increases as they grow older.</i>	Five-point Likert scale (strongly disagree, ..., strongly agree)

* Each question was preceded by the phrase: "Based solely on the data in the diagram."

43

Fischetti, M. China and the U.S. Are Not the Biggest Carbon Emitters on a Per-Person Basis. Scientific American, Graphic Science. October, 2015. Online. Available at <https://www.scientificamerican.com/article/china-and-the-u-s-are-not-the-biggest-carbon-emitters-on-a-per-person-basis/> (Accessed June 12, 2018).

44

World Resources Institute–CAIT Climate Data Explorer Total CO2 Emissions Excluding Land-UseChange and Forestry–2012. Online. Available at <https://bit.ly/35TMjJI> (Accessed June 12, 2018).

45

Williams, A. “Some of the Best Science Can Slumber for Years.” Scientific American, Graphic Science. January, 2016. Online. Available at <https://www.scientificamerican.com/article/graphic-science-some-of-the-best-science-can-slumber-for-years/> (Accessed June 12, 2018).

46

Ke, Q., Ferrara, E., Radicchi, F. and Flammini, A. “Defining and Identifying Sleeping Beauties In Science.” Proceedings of the National Academy of Sciences (PNAS), 112 (2014): pgs. 7426–7431

47

Christakis, N.A., Fowler, J.H. “The Spread of Obesity in a Large Social Network over 32 Years.” New England Journal of Medicine. 357 (2007): pgs. 370–379. Available at <https://www.nejm.org/doi/full/10.1056/NEJMsa066082> (Accessed Sep. 12, 2019)

Table 2: Task completion time statistics (in seconds) after removal of extremes

	Task completion times (seconds)						
	MIN.	1ST QUART.	MEDIAN	MEAN	3RD QUART.	MAX	STD. DEY.
a. Ex data for subjects with unusually short task completion times	240	543	706	911	1013	5,634	739
b. Ex data for subjects with unusually short completion times and ex unusually long and short response times per question	240	489	672	716	890	1,690	325

extracting quantitative information from graphs and diagrams (their textbooks are filled with them). As such, most of them can reasonably be expected to interpret information from complex diagrams.

*A Description of Our Data Cleansing Process and Its Intent:
The Removal of Extremes*

Prior to engaging in the analysis of the data we collected, two data-cleansing operations were performed.

1. Removal of all data provided by subjects who exhibited unusually short task completion times.
2. Removal of all single responses gleaned from subjects who exhibited unusually short or unusually long response times.

In each of these instances, ‘unusual’ was defined as the logged values of time taken to complete the survey that lay within the 5% tail(s) of the normal distribution.

Table 2 shows the task completion time statistics after we affected these removals. They show a large correction on the long completion-time side of the distribution. These long completion times were most likely the effect of the survey not being administered under controlled, timed conditions, as well as allowing subjects to start contemplating a prompt (begin to have it be displayed in their browser), and not urging them to respond to it for some time. Removing these instances reduced the width of the task completion time distribution by 73% and its standard deviation by 56%. Mean survey completion

time after this data cleansing was 716 seconds (13 minutes), with a standard deviation of 325 seconds (5.5 minutes)

An Analysis of How Well Test Subjects Understood the Data-cum-Information Visualized in The Five Examples that Comprised the Experiment

Table 3 contains the proportions and response times of the ‘stated understanding’ prompt for each of the five visualization examples utilized in our experiment. Posited in the form of a basic question, stated understanding was articulated as follows: “Do you understand the data in the diagram?” We observed the following:

- For the “2016 Warm Year” (Figure 1-a/b/c), “Italian Poverty” (Figure 1-d/e) and “Sleeping Beauties” (Figure 1-h/i) examples, subjects took more time on average to respond to the prompt if they understood the information that was visualized in the complex version of the diagram, and fewer stated that they actually understood the information that was visualized in that diagram.
- Mean response times for the “CO₂ Emissions” example (Figure 1-f/g.) ‘stated understanding’ prompt is not significantly different for the small sample size of 72. However, if the pattern holds for a larger sample size, only tripling it (n=207) would render the same difference in means highly significant. We note that presenting the data in two different formats—as was done in Figure 1-f—appears to have had no beneficial effect on stated comprehension among test subjects; *i.e.*, the graduated circles and coloring scheme appear to have no beneficial effect on either time needed by test subjects to gain fundamental understanding of the data depicted, or on the stated degree of their understanding.
- With regard to how test subjects responded to the “Obesity Network” example, mean response times were also not significantly different for the small sample size of 72. With that stated, it must be noted that the relationship between time taken by subjects to effectively discern information visualized in the Complex example versus that of the Simple example measured in the opposite direction of expectation (Complex: 33.34 seconds vs. Simple: 40.04 seconds). We surmise that this inverse relationship is the result

Table 3: Stated understanding and answer times

Case	Proportion “I understand the data in the diagram”			Mean response time (seconds)		
	COMPLEX	SIMPLE	Z-TEST P-VALUE	COMPLEX	SIMPLE	T-TEST P-VALUE
2016 warm year	.792	(a) .83	–	53.88	(a) 53.32	–
	.792	(b) 1.00	.081		(b) 35.90	.016
Italian poverty	.606	.944	.002	77.71	36.91	.003
CO ₂ emissions	.727	.75	–	51.84	39.18	–
Sleeping beauties	.600	.882	.017	45.08	23.38	.002
Obesity network	.447	.742	.026	33.34	40.04	–

of subjects only briefly observing this diagram, and then making a quick decision that they could not understand the data presented within it. The statistically significant difference in proportions of subjects’ initial stated understanding (Complex: 44.7% vs. Simple: 74%), support this interpretation.

An Analysis of the Results of the Numerical Estimations by Subjects

The survey contained four questions that challenged subjects to numerically estimate certain quantities from the visualizations. For example, in the *CO₂ Emissions* example (Figure 1-f/g), subjects were prompted to fill in the blank in the following statement: “According to the data in the diagram, China’s 2012 total emissions were about _____ times the size of those by India.” Similarly, in the *Sleeping Beauties* example (Figure 1-h/i), subjects were asked to fill in the blanks in the following statement: “In 2011, the paper by Hummers & Offeman received approximately _____ times the number of citations as the paper by Pugh.” Each of these estimates involved assessing an aggregate data pattern rather than assessing single-value lookups. In terms of Jacques Bertin’s in his seminal text from 1967: “all estimates were of [the] *intermediary* or *overall information level*.”⁴⁸ Table 4 contains the results for these estimates.

The data in Table 4 depicts clear evidence of problems inherent in asking subjects to interpret data from each of the complex examples. Subjects took longer to respond to given prompts, or they made larger errors, or both.

48

Bertin, J. *Semiology of Graphics: Diagrams, Networks, Maps*. Madison, WI, USA: Univ. of Wisconsin Press, 1983: p. 153.

Table 4: Results for numerical estimation questions

Case	Estimation Variable	Mean error complex	Mean error simple	t-test p-value	mean(time) complex	mean(time) simple	t-test p-value
2016 warm year	_____ percent of days in 2016 in New York City were warmer than normal.	-25.39	1.89 (a)	< .001	91.23	81.55 (a)	
		-25.39	3.91 (b)	< .001	91.23	53.77 (b)	.10
CO ₂ emissions	China's 2012 total emissions were about _____ times the size of those by India.	.01	-.31		43.65	28.83	.038
	China's 2012 total emissions were about _____ times the size of those by the USA.	-.20	.16	.04	23.06	25.44	
Sleeping beauties	In 2011 paper No 2 received approximately _____ times the number of citations of paper No 13.	-1.80	.12	.007	48.24	47.61	

We also note that for one of the *CO₂ Emissions* questions subjects, on average, significantly underestimated quantities they were asked to discern (indicated by negative error values). This is consistent with the well-documented problem of size comparisons based on graduated circles mentioned earlier. We saw this repeated in subjects responses to the *Sleeping Beauties* example, where the mean underestimation in the complex condition is more than 10 times the size of the overestimation in the simple condition (-1.80 vs. .12).

We also observed that in the *CO₂ Emissions* example, subjects challenged to respond to the Complex version of the visualization cut their response time on the second prompt (USA vs. India) almost in half when compared with that of the first (China vs. India). We offer that this reduction in response time occurred due to the challenges that confronted test subjects with regard to their having to learn to use the basic structure of the visualization in order to effectively glean understanding from it. In this manner, we posit that although a complex diagram may take extra time to learn to read and discern understanding from, once learned, subsequent reads are performed faster. This learning effect is particularly pronounced in the *Italian Poverty* example (Table 5). Whereas response times in the early questions in its Complex version are far longer than in its Simple version, we observed that both the Complex response times and answer accuracy improved sharply once subjects learned how to read the diagram. We observed, however, that by this time, they had not only taken significantly more time overall to engage with the data presented, but

Table 5: Results for the *Italian poverty* case

Question (in sequence)	Proportion of correct answers			Response time mean(seconds)		
	COMPLEX	SIMPLE	Z-TEST P-VALUE	COMPLEX	SIMPLE	T-TEST P-VALUE
% poor in Central Italy	.61	.86	.03	61.63	28.57	.002
% poor 3-member hh in North Italy	.76	.94	.06	31.33	17.40	< .001
% poor 5-member hh in Italy	.82	.83		27.47	24.68	
more members in hh increase likelihood of being poor (agreement)	.97	.97		27.37	24.68	

they had also made significantly larger mistakes in responding to the challenges offered in the earlier prompts.

A revealing pattern of subject response and understanding was observed when comparing the proportions of stated comprehension and response times between the two Simple versions of the *2016 Warm Year example* (Figure 1-b/c). Subjects exposed to the bar graph (b) took longer to comprehend it and fewer were able to state their clear understanding of it, even though it contains only half the amount of data of the dual line graph (c). Since the question of whether or not 2016 was a warmer than normal year is a matter of differential, displaying this differential rather than the variables from which the differential can be inferred seemed the simplest means to visually articulate this data. Indeed, we expected that the much simpler structure, as well as the red/blue warmer/colder color coding of the bar graph (b), would have made it easier for subjects to interpret its inherent meanings than the more complex dual line graph (c). However, subjects apparently considered the concept of an explicitly graphed temperature differential more difficult to grasp than an implied differential represented by a comparative display of each of the two temperature ranges. This is reflected in the associated response times for questions pertaining to the diagrams. However, as this experiment revealed, looks can be deceiving. Despite subjects' stated understanding and quickness of response times in the dual line graph condition (c), their mean numerical estimation error was almost twice that of the graph of differentials (b). Apparently, the notion of a

differential may be harder to understand at first, but its simpler graphic representation appears significantly easier for subjects to interpret.

Opportunities for Further Analysis

The results from this experiment pointed in the expected direction. Subjects took less time to understand simpler visualizations, indicated a more effective understanding of the data-cum-information depicted within them, and made smaller errors in reading data from them. Moreover, the differences were statistically significant. The case of the *CO2 Emissions* data also demonstrated that displaying the data more than once, at least in this situation, did not aid comprehension and information retrieval. Indeed, given the reputation of utilizing graduated symbols as a poor means to depict size comparisons, mixing these types of symbol sets with the far-easier-to-comprehend and far-more-familiar 2D point diagram introduces a confounding set of data.

It would, of course, be shortsighted to conclude from these results that the observed effects of the differences between the Complex and Simple versions of data visualizations utilized in this experiment were all just a function of degree of complexity. As we articulated earlier in this article, the success of the Simple version of the *Sleeping Beauties* example may well be due to subjects' familiarity with left-to-right timelines as compared with vertical timelines mixed with graduated circles to indicate magnitude. Similarly, looking up values in a simple data table is obviously easier than inferring them from an alluvial diagram (Figure 1-d), or from a complicated network diagram (Figure 1-j). Perhaps the disadvantages of complex visualizations can be overcome by learning how to read them. The *Italian Poverty* example (Figure 1-d) appears to indicate that indeed, once subjects learned how to navigate the visualization, they performed just as well as those who engaged with the much simpler version of it. We could also compare each of the 'Simple' versions of each of the visualizations used in this experiment with its corresponding 'Complex' version along various design dimensions such as use of color, line, area, size, shape, contrast, texture, *etc.* We might then be able to attribute the differences in the responses among given subjects to particular data visualizations to the differences inherent in those visualizations within a single dimension or a combination of dimensions. Still, if simply re-visualizing the data from a series of seemingly overcomplicated visualizations into a series of straightforward ones yields such significant effects, the advice to 'keep things simple' might be worth heeding.

Conclusions

We set out to assess some of the consequences of formally complicating visualizations of scientific data. We selected five real-world examples which we believed qualified as such (based on our criteria), and which we then used to develop and design simpler and easier to interpret versions derived from the identical data. We then randomly exposed subjects to either type of visualization, and asked them to respond to a series of prompts that allowed them to indicate how well they were able to understand the data presented in the visualizations. On average and as expected, subjects took longer to comprehend the complex versions and made larger errors interpreting their data, sometimes dramatically. Although our results showed that subjects eventually learned how to effectively interpret some of the complex versions, this is of little value in cases where a given data visualization is meant to convey a single concept. Since we believe that visually communicating the meaning inherent in single concepts constitutes the vast majority of these types of visualizations, we suggest that information designers, when choosing graphical representations of data, carefully consider whether or not deviating from tried-and-true, graphically simple configurations of the formal elements that represent data is warranted. We further suggest that they consider submitting their new creations to at least some testing against one or two, more graphically simplified versions to determine whether “asserting their creativity” pays off in terms of reader comprehension and information retrieval. To be clear, the goal of our experiment was not to verify and promote graphically simplified data tables, line graphs or point diagrams, which offer clarity but which are also visually bland and aesthetically and conceptually unengaging. We did, however, set out to assess some of the potential consequences of deviating from these types of more primitive symbolisms for cases where they would suffice.

What then might a future bring in which ever more data is being quantified, displayed and consumed? Will those who design visualizations of it understand how to clearly communicate its inherent meaning while still finding ways to render it so that it is aesthetically and conceptually rich and engaging? Will trends in education continue to build on the importance of enhancing visual literacy from a young age forward, in order for future citizens to be able to interpret the information that visualizations like these deliver and then take action based on what they’ve learned. Will they use this knowledge to perhaps learn to critically question a specific visualization’s veracity? Continuing research into why visualizations succeed or fail in communicating information

should shed light on developing, shaping and, as necessary, re-shaping best practices for doing this, and grow our knowledge about what approaches and methods are most and least effective. One thing seems certain: the future will increasingly bring large volumes of data into people's lives, and the fact that data visualization software will be ever more accessible will increase these volumes. These factors confirm the need to educate future generations on how to appropriately read and design with data.

References

- Bateman, S., Mandryk, R.L., Gutwin, C., Genest, A., McDine and D. and Brooks, C. "Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2010*. Association of Computing Machinery, 2010: pgs. 2573-2582. Online. Available at <http://dmrussell.net/CHI2010/docs/p2573.pdf> (Accessed June 12, 2018).
- Bertin, J. *Semiology of Graphics*. Madison, WI, USA: Univ. of Wisconsin Press, 1983: p. 153.
- Bigwood, S. & Spore, M. "When to use numeric tables and why." in *Information Design* edited by A. Black, P. Luna, P., O. Lund and S.E. Walker, pgs. 503-507. London, UK & New York, NY, USA: Routledge, 2017.
- Börner, K. *Atlas of Science: Visualizing What We Know*. Cambridge, MA, USA: The MIT Press, 2010.
- Börner, K. & Polley, D.E. *VISUAL INSIGHTS. A Practical Guide to Making Sense of Data*. Cambridge, MA, USA: The MIT Press, 2014.
- Börner, K. *Atlas of Knowledge; Anyone Can Map*. Cambridge, MA, USA: The MIT Press, 2015.
- Chang, K.T. "Visual Estimation of Graduated Circles." *Canadian Cartographer*, 14 (1977): pgs. 130-138.
- Cheng, K. "Designing Data: Issue 31.4 Feature Introduction." *ARCADE* .31.4 (2013). Online. Available at: <http://arcadenw.org/article/designing-data> (Accessed June 12, 2018).
- Christakis, N.A. & Fowler, J.H. "The Spread of Obesity in a Large Social Network over 32 Years." *New England Journal of Medicine*, 357 (2007): pgs. 370-379.

- Cox, C.W. Anchor “Effects and the Estimation of Graduated Circles and Squares.” *American Cartographer*, 3 (1976): pgs. 65–74.
- Crawford, P.V. “Perception of Grey-tone Symbols.” *Annals of the Association of American Geographers*, 61 (1971): pgs. 721–735.
- Crawford, P.V. “The Perception of Graduated Squares as Cartographic Symbols.” *Cartographic Journal*, 10 (1973): pgs. 85–88.
- Fabrikant, S.I., Montello, D.P. and Mark, D.M. “The Natural Landscape Metaphor in Information Visualization: The Role of Commonsense Geomorphology.” *Journal of the American Society of Information Science and Technology*, 61 (2009): pgs. 233–270.
- Fischetti, M. China and the U.S. Are Not the Biggest Carbon Emitters on a Per-Person Basis. *Scientific American, Graphic Science*. October, 2015. Online. Available at <https://www.scientificamerican.com/article/china-and-the-u-s-are-not-the-biggest-carbon-emitters-on-a-per-person-basis/> (Accessed June 12, 2018).
- Flannery, J.J. “The Relative Effectiveness of Some Common Graduated Point Symbols in the Presentation of Quantitative Data.” *Canadian Cartographer*, 8 (1971): pgs. 96–109.
- Groop, R.E. & Cole, D. “Overlapping Graduated Circles, Magnitude Estimation and Method of Portrayal.” *Canadian Cartographer*, 15 (1978): pgs. 114–122.
- Haemer, K. “The Pseudo Third Dimension.” *The American Statistician*, 5.28 (1951).
- Hall, P. “Bubbles, Lines and String: How Information Visualization Shapes Society.” In *Graphic Design: Now in Production* edited by A. Blauvelt and A. Lupton, pgs. 170–185. Minneapolis, MN, USA: Walker Art Center, 2011.
- Heer, J. & Bostock, M. “Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design.” in *Proceedings of the SIGCHI Conference in Human Factors in Computing Systems (CHI)*, 2010. Association for Computing Machinery, 2010. pgs. 203–212. Online. Available at <http://vis.stanford.edu/files/2010-MTurk-CHI.pdf> (Accessed June 12, 2018).
- Javed, W., McDonnell, B. and Elmqvist, N. “Graphical Perception of Multiple Time Series.” *IEEE Transactions on Visualization and Computer Graphics*, 16 (2010): pgs. 927–934.
- Ke, Q., Ferrara, E., Radicchi, F. and Flammini, A. “Defining and Identifying Sleeping Beauties in Science.” *Proceedings of the National Academy of Sciences (PNAS)*, 112 (2014): pgs. 7426–7431.

- Kelleher, C. & Wagener, T. “Ten guidelines for effective data visualization in scientific publications.” *Environmental Modelling & Software*, 26 (2011): pgs. 822–827.
- Lai, KK Rebecca. “How Much Warmer Was Your City in 2016” *New York Times* on-line. January 18. 2017. Online. Available at <https://www.nytimes.com/interactive/2017/01/18/world/how-much-warmer-was-your-city-in-2016.html#nyc> (Accessed June 12, 2018).
- Lima, M. *Visualizing Complexity. Mapping Patterns of Information*. New York, NY, USA: Princeton Architectural Press, 2011.
- Istat. “La povertà relativa in Italia nel 2006.” *Statistiche in breve. Istituto Nazionale di Statistica*. Roma. Online. Available at <http://www.istat.it/it/files/2011/03/poverta06.pdf> (Accessed June 12, 2018).
- Meirelles, I. *Design for Information. An introduction to the histories, theories, and best practices behind effective information visualizations*. Beverly, MA, USA: Rockport Publishers, 2013.
- Pettersson, R. *Information Design. An Introduction*. Amsterdam, NL and Philadelphia, PA, USA: John Benjamin’s Publishing Company, 2002.
- Porpora, M. *Italian Poverty Red Thread*, 2009. Online. Available at <https://www.behance.net/gallery/463228/Italian-Poverty-Red-Thread>. (Accessed June 12, 2018).
- Reitsma R., Hsieh, P. - H., Diekema, A., Robson, R. and Zarsky, M. “Map- or List-based Recommender Agents? Does the Map Metaphor Fulfill its Promise?” *Information Visualization*, 16 (2016): pgs. 291–308.
- Reitsma, R. & Trubin, S. “Information Space Partitioning Using Adaptive Voronoi Diagrams.” *Information Visualization*, 6 (2007): pgs. 123–138.
- Schonlau, M. & Peters, E. “Comprehension of Graphs and Tables Depends on the Task: Empirical Evidence from two web-based studies.” *Statistics, Politics and Policy*, 3 (2012). Article 5.
- SciAm *Graphic Science*. Scientific American. Online. Available at <https://www.scientificamerican.com/departments/graphic-science/> (Accessed June 12, 2018).
- Skau, D., Harrison, L. and Kosara, R. “An Evaluation of the Impact of Visual Embellishments in Bar Charts.” in *Proceedings Eurographics Conference on Visualization (EuroVis) 2015*. Online. Available at <https://kosara.net/papers/2015/Skau-EuroVis-2015.pdf> (Accessed June 12, 2018).
- Thudt, A., Walny, J., Perin, C., Rajabiyazdi, F., MacDonald, L., Vardeleon, R., Greenberg, S. and Carpendale, S. “Assessing the Readability of

- Stacked Graphs.” in *Proceedings of the 2016 Graphics Interface Conference*. Victoria, BC, Canada. 2016. Online. Available at <http://innovis.cpsc.ucalgary.ca/innovis/uploads/Publications/Publications/Thudt-Streamgraphs2016.pdf> (Accessed June 12, 2018).
- Tufte, E.R. *The Visual Display of Quantitative Information*. Cheshire, CT, USA: Graphics Press, 1983.
- Tversky, B. “Diagrams. Cognitive Foundation for Design.” in *Information Design* edited by A. Black, P. Luna, P., O. Lund and S.E. Walker, pgs. 349-360. London, UK & New York, NY, USA: Routledge, 2017.
- Wilkinson, L. “Less is More: Two-and Three-Dimensional Graphics for Data Display.” *Behavior Research Methods, Instruments and Computers*, 26 (1994): pgs. 172–176.
- Williams, A. “Some of the Best Science Can Slumber for Years.” *Scientific American, Graphic Science*. January, 2016. Online. Available at <https://www.scientificamerican.com/article/graphic-science-some-of-the-best-science-can-slumber-for-years/> (Accessed June 12, 2018).
- Williams, R.L. *Statistical Symbols for Maps; Their Design and Relative Values*. New Haven, CT, USA: Yale University, 1956.
- World Resources Institute – CAIT Climate Data Explorer *Total CO₂ Emissions Excluding Land-UseChange and Forestry - 2012*. Online. Available at <https://bit.ly/35TMJJI> (Accessed June 12, 2018).

Biography

René Reitsma is a Professor of Business Information Systems at Oregon State University’s College of Business in Corvallis, Oregon, USA. He grew up and was educated in the Netherlands and worked at the International Institute for Applied Systems Analysis (IIASA) in Laxenburg, Austria, the University of Colorado, Boulder, Colorado, USA, and St. Francis Xavier University in Nova Scotia, Canada before joining the Oregon State faculty in 2002. Professor Reitsma teaches courses in Information Systems Analysis, Design and Development and Business Process Management. His research interests are in information visualization and digital libraries.

reistmar@bus.oregonstate.edu

Andrea Marks is a Professor of Design at Oregon State University's College of Business in Corvallis, Oregon, USA. She is a design researcher, author/producer, and has worked as a design educator at Oregon State University since 1992. She received a BFA in Graphic Design from University of the Arts in Philadelphia, Pennsylvania, USA, and was awarded a Fulbright International Education Fellowship for her post-graduate work at the Basel School of Design in Basel, Switzerland. Her research includes producing and co-directing *Freedom on the Fence*, a documentary film chronicling the history of the Polish poster, and she is the author of the book *Writing for Visual Thinkers* (New Riders Press, 2011). Currently, she teaches courses in Design Thinking, Innovation and Entrepreneurship.

andrea.marks@oregonstate.edu