

The Perils of Misinformation

by Tatiana Bailey, Ph.D.

Director, *UCCS Economic Forum*

My first day in graduate school I had plenty of mechanical pencils and a ruler at the ready for all the fun graphs I planned to draw in the economic, business and statistics classes I was so looking forward to. As I walked into my first statistics class, I remembered that everyone in the economics department was warned that the hardest classes were statistics with Dr. Chamberlin and econometrics with I don't remember who. I remember Dr. Chamberlin though. On the first day he handed out photocopies (no laptops yet) of articles from prominent newspapers with "data." He gave us a few minutes to read them and then asked the class what we thought about the information. The content was interesting, but several of us didn't take the bait. Something had to be wrong with the way the statistics were presented. I was bold and raised my hand to point out one mistake. Turns out there were several layers of mistakes and since Dr. Chamberlin's class, I have become ruthless in my examination of data. Not for the sake of being a nitpicker, but because I realized during that year how pervasive misrepresentation of data is. And how little we are taught about how to scrutinize data.

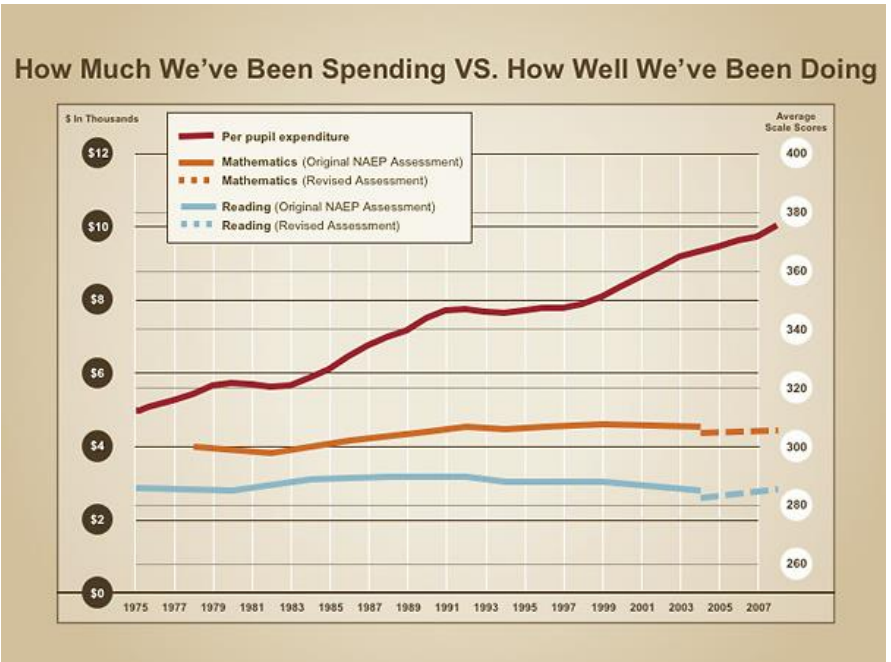
Fast-forward to today where snippets of information are distorted in tables or graphs and then amplified through social media most often on phone screens. This misinformation flies through the virtual world as does the bias that underlies it. We now have access to a sea of information thanks to the digital age, which most would agree is a wondrous thing. It does, however, inundate us and it is difficult to keep up. Enter quick data, especially when it looks official and feeds our biases, and we soak in what we can as fast as we can. It's easier than ever for individuals and organizations to dupe us, and as the resident data nerd, I thought it might help to point out how we are sometimes misled. I will first address this mostly through examples. Then I will talk in a subsequent article about how some academic experts argue that the structure of our two-party political system inherently perpetuates the misrepresentation of information and that this has led to historically high levels of polarization. Sensitive topics, but perhaps particularly important in an election year.

The age-old questions of who, what, where, when and why can help categorize and simplify the essential statistical questions. Let's start with who. Studies do not typically include an entire population, so it is relevant to ask who is missing. For example, a medical trial for a new cold therapy could solicit volunteers at a college campus where students can get \$50 for participation. What if particularly lazy students who live on Chick-fil-A enroll? Results might be biased because on the one hand, students are younger and more resilient, but on the other hand, students often have poor diets and compromise their immune systems with boozy weekends that start on Wednesdays at noon. And if the pharmaceutical company tries to branch out and enrolls more "normal" populations, how do they account for the fact that paid participants can be poorer with suboptimal health, or interested in the study because they are disproportionately sicker with colds?

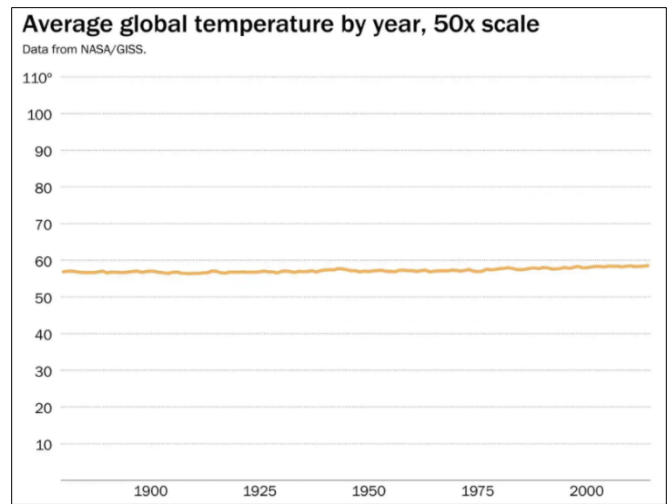
This is called selection bias and it can be one layer in many. Polling results, for elections, controversial issues or even products, are often displayed as truisms. However, in the age of mobile phones and screening of calls, less than 6% of polling calls are answered according to the Pew Research Center. What about the 6% who do participate? Do they have more time on their hands, and if so, aren't they a special population who can perhaps skew results? Many polls still even use landlines. I don't know anyone who still has a landline except my mother, and she is older and extremely opinionated. Could there be a disproportionately high percentage of landline respondents who think Eisenhower is still president?

With respect to polling, one other layer may be what questions are asked, and then how results are presented to the public often excluding the leading question. Consider, "Given that U.S. debt is at \$22 trillion, its highest level ever and more than total GDP in 2019, do you support income tax cuts?" Compare this question to, "Given the reduction in the middle class in the past four decades, do you support reducing the income tax burden for qualifying families?" Either survey question could result in statements such as, "72% of Americans do not support income tax cuts," or "72% of American support income tax cuts." It will depend on how the question was asked and what news station you are watching.

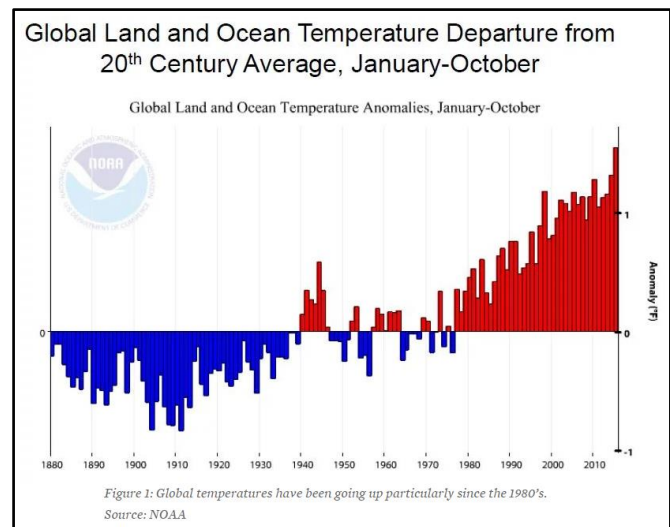
Spuriousness is the root cause of other commonly seen distortions. A 2013 *Slate* article entitled, “Will Teachers Unions Kill Virtual Learning?” makes a statement that per-pupil spending on public education has more than doubled over the past three decades, while student performance has flatlined. A graph is included as evidence with significant increases in funding since 1975 with no increases in test scores. What about other factors that can be impacting test scores? Some might argue it has to do with more Americans below the poverty line, the increasing divorce rate, or an antiquated educational system. In other words, the correlation does not imply causation. Also, where is the data source? Has the way we count per pupil expenditures changed at all since 1975? Have NAEP tests changed at all? Are those per pupil expenditures incorporating inflation increases or are they in nominal dollars? The “when” matters here too. Also, “Teachers” in the title should technically be possessive: “Teachers’ Unions...” The grammatical error speaks more to me about the quality of the education of the author than does the lousy graph. Yes, I can be an annoying nitpicker.



Another aspect of “what” relates to scale and context. Non-believers of climate change show graphs with what is usually a correct scale that starts at zero (first graph). What’s the big fuss in going from 58 degrees to 59 degrees? Believers of climate change show temperatures that deviate from a baseline with a specific time point (e.g. 1900 shown in the middle graph) or that only focus on the past 20 years with a scale that starts at a higher temperature (e.g. the third graph that starts at 55 degrees). Couldn’t you build an argument with any of these graphs? In this case, context matters so it may help to defer to content experts who study the ramifications of a one-degree change in average temperatures.



Data dredging and manipulation can involve looking at large data sets and then cherry-picking components of the data to make an argument. The final graph illustrates the number of people on welfare in the U.S. versus the number of people with a full-time job. The source seems legitimate: the U.S. Census Bureau. The problem is that the Census Bureau has huge data sets on many different topics. The first bar shows a count of *all people within a household (including children)* where at least



one recipient received any kind of government benefit for any period of time. The second bar shows *individuals* with a full-time job; it is not household data.

There are more kinds of data distortion, but I think you get the idea. A disclaimer here that even as a data nerd who only uses mechanical pencils, I am sometimes the sucker myself. The point is that if you know what questions to ask, you are more than halfway there and not easily duped. This may be particularly useful as big data and technology continues to increase logarithmically, and as we make critical decisions with the sea of information we now possess.

Published by the Colorado Springs Business Journal on February 14, 2020; www.csbj.com.

The next article will focus on the two-party political system and how it can intensify the misrepresentation of data and polarization.

For information on receiving the monthly economic dashboard via sponsorship, please contact Tatiana Bailey, Director, UCCS Economic Forum (tbailey6@uccs.edu)

