



Truth, Damn Truth, and Statistics

Paul F. Velleman

To cite this article: Paul F. Velleman (2008) Truth, Damn Truth, and Statistics, Journal of Statistics Education, 16:2, , DOI: [10.1080/10691898.2008.11889565](https://doi.org/10.1080/10691898.2008.11889565)

To link to this article: <https://doi.org/10.1080/10691898.2008.11889565>



Copyright 2008 Paul F. Velleman



Published online: 29 Aug 2017.



Submit your article to this journal [↗](#)



Article views: 696

Truth, Damn Truth, and Statistics

Paul F. Velleman Cornell University

Journal of Statistics Education Volume 16, Number 2 (2008),
www.amstat.org/publications/jse/v16n2/velleman.html

Copyright © 2008 by Paul F. Velleman all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

Key Words: Damn lies; Twain; Ethics; Statistics education.

Abstract

Statisticians and Statistics teachers often have to push back against the popular impression that Statistics teaches how to lie with data. Those who believe incorrectly that Statistics is solely a branch of Mathematics (and thus algorithmic), often see the use of judgment in Statistics as evidence that we do indeed manipulate our results.

In the push to teach formulas and definitions, we may fail to emphasize the important role played by judgment. We should teach our students that they are personally responsible for the judgments they make. But we must also offer guidance for their statistical judgments. Such guidance requires that we acknowledge the role of ethics in Statistics. The principle guiding these judgments should be the honest search for truth about the world, and the principle of seeking such truth should have a central place in Statistics courses.

The remark attributed to Disraeli would often apply with justice and force: "There are three kinds of lies: lies, damn lies, and statistics".
—[Mark Twain](#)

This may be my least favorite quotation about Statistics. But I wish to address what underlies both the quotation and the gleeful willingness of many who know nothing at all about Statistics to quote it as if it justified their low opinion of the discipline.

This quotation has infiltrated discussions in many disciplines. Surely you have had it quoted back to you if you were foolish enough to admit in polite company that you teach Statistics. Nigel Rees's *Quote...Unquote*¹ claims that this is the single most quoted remark in the British media.² A *Google books* search of "lies, damn lies, and statistics" turns up 495 books, and a general Google search finds "about 207,000" hits. A small (nonrandom) sample of these references shows that most are meant to suggest dishonest manipulations and interpretations.

1. Sources

The origin of the lies, damn lies... remark is unclear. Online materials on the history of Statistics posted by the University of York³ provide an excellent discussion. Two things are clear: Twain did not originate it (nor, of course, did he claim to), and he was most likely mistaken in attributing it to Disraeli. Twain's source isn't known. The most likely source may be an address given at Saratoga Springs in 1895 (not far from Twain's Elmira NY home) by [Leonard Henry Courtney](#) (1832-1918), the British economist and politician, who said:

After all, facts are facts, and although we may quote one to another with a chuckle the words of the Wise Statesman, "Lies—damn lies—and statistics," still there are some easy figures the simplest must understand, and the astutest cannot wriggle out of.⁴ p. 25).

It is plausible that Twain heard or read the speech and mistakenly assumed that the "Wise Statesman" was Disraeli. Lord Courtney is clearly pointing out the value of statistics in spite of the amusing comment he quotes, which is not surprising since he was President of the Royal Statistical Society from 1897 to 1899. But, as we shall see, this may not be inconsistent with what Twain actually intended.

There are a few references that pre-date both Twain and Courtney. Perhaps the most intriguing is a report in the *Manchester Guardian* of 1892 of a political speech by Arthur James Balfour. Balfour uses the phrase exactly as it would be used today to accuse a political opponent of manipulating statistics by failing to note a changing denominator:

But the improvement in Ireland from 1886 to 1892 was an inconvenient fact, and therefore the Gladstonians set themselves to work to prove that the fact was no fact at all. ... There were certain propositions so obvious to every man who knew the facts that it was in vain to parade, he would not say cooked statistics, for that would be offensive, but manipulated statistics, before the eyes of any audience in the country.—(Hear, hear.) Professor Munro reminded him of an old saying which he rather reluctantly proposed, in that company, to repeat. It was to the effect that there were three gradations of inveracity---there were lies, there were d---d lies, and there were statistics.

The average receipts, said Mr. Munro, from passengers during the years 1881 to 1885, when Mr. Gladstone was in office, were 1,098 per mile, while the average receipts between the years 1886-90, when Mr. Balfour was in office, were only 1,092 per mile, showing a decrease of 6 per mile. ...[However} There had been a great extension of light railways in Ireland, and the result had... the effect, of course, of gradually increasing...the number of miles over which the average traffic receipts had to be calculated. (1892, pp 5-6)

This, of course, raises the possibility that it was Balfour, a power in the Conservative Party for 50 years, and not Disraeli, who was the "wise statesman" credited by Lord Courtney⁵. If, as Balfour claims, the *bon mot* is an "old saying," we may never know its true origin.

2. Truth

Of course, one *can* wield the tools of Statistics to mislead. But even those who repeat the quotation don't believe that the *purpose* of the discipline of Statistics is to mislead, or that there is something fundamentally dishonest about statisticians. Statistics doesn't lack respect because people think that statisticians are crooks. You'd buy a used car from a statistician. Why, then, is the insult so widespread?

And what can we as teachers of Statistics learn from the ubiquity of the phrase?

Let's approach these questions by first considering truth. Philosophers have debated the nature (and even the existence) of truth for centuries. I want to avoid as much of that debate as I can, so I'll state my axioms and propose not to debate them here:

- There is a world outside of ourselves.
- There are facts about that world that are true in the ordinary sense of the word. That is, they are true regardless of what you, or I, or anyone else, may believe about them.
- To be useful, theories, models, and explanations of the world must account for observed facts.
- Scientific (and thus, Statistical) thinking seeks ways in which existing understanding might be wrong because such exceptions illuminate the path forward to better understanding of the truth about the world.
- One of the major goals of Statistics—in fact, the *principal* goal of Statistics—is facilitating the discovery, understanding, quantification, modeling, and communication of facts about the world.

Other disciplines see Statistics in this way. Some regard statistical analysis as a gatekeeper. Statistical significance is the first requirement for publication in many social sciences. Medical disciplines discuss "evidence-based medicine," which means selecting treatments based on the best scientific (usually statistically-based) evidence to date.

These concerns enter our teaching in the elementary statistics courses. The rubrics for the AP Statistics test insist that the final step of a free-answer question be a sentence or two that relates the result to the real-world question that motivated the exercise. That attitude—that the motivation for a statistical analysis is a question about the world and the conclusion of the analysis is a statement addressing that question—will prove to be central to our investigation.

3. Facts and Process

The term "Statistics" has two meanings that concern us here. It can refer to isolated facts, or it can refer to analyses and processes by which we work with these facts to attain deeper understanding.

Those who select, define, and summarize facts must make judgments. The shifting denominator such as the one noted by Balfour is typical of the judgments that even simple facts can require. The risks of making a wrong judgment or of being deceived by one have long been recognized. According to [William Guy](#), The Transactions of the British Association establishing the Statistical Society of London (precursor to the Royal Statistical Society) in 1834 stated that its purpose would be the procuring, arranging, and publishing of

...facts calculated to illustrate the conditions and prospects of society, ...the first and most essential rule of its conduct [was] to exclude carefully all *opinions* from its transactions and publications—to confine its attention rigorously to facts... (p 482)

Guy notes the impossibility—and the undesirability—of meeting this goal, and the very evident fact that it was almost immediately ignored by leading members of the Society.

For our purposes here, it is sufficient to recognize that even the simplest statistic involves judgments. There is, then, no practical difference between the two meanings in their reliance on judgment.

4. Damn Truth

John Tukey taught that Statistics is more a science than it is a branch of Mathematics. For a mathematics theorem to be elegant, it is sufficient that it be beautiful and true. But Statistics is held to the additional standard imposed by science.⁶ A model for data, no matter how elegant or correctly derived, must be discarded or revised if it doesn't fit the data or when new or better data are found and it fails to fit them. [Thomas Huxley](#) referred to this as

The great tragedy of science - the slaying of a beautiful hypothesis by an ugly fact.
- *Biogenesis and Abiogenesis*

Statistics lives on the empirical, rather than the theoretical side of science. Statisticians deal with data, and build and examine models for those data. When the model and the data diverge, it's often a sign of progress rather than a failure, so that's where we often focus our attention. As Isaac Asimov reminds us

The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'That's funny...' ⁷

This constant reference to understanding the world is one of the things that makes a Statistics course more difficult to teach than a Mathematics course at a comparable level of technical sophistication. Our corresponding insistence that students write a sentence or two tying their calculations to a conclusion about the world sometimes surprises and frustrates students who thought they'd signed up for a math class (and, of course, pleases others who had feared that was the case.)

Most of the time, the truth we hope to describe or model is something that we can't know anyway. What is the true fraction of U.S. adults who smoke, who have children, or who smoke in the presence of their children? The best we can hope for are data that represent the population and let us draw reasonable inferences.

When an objective regulates practice by providing a goal toward which the practice strives, Philosophers call it a *regulative ideal*. The fact that we can never know the truth shouldn't keep statisticians from using it as a beacon to guide our decisions as we seek better and better understanding. One needn't be able to reach a goal for it to serve as a regulative ideal. For example, a perfect score in golf is 18.

5. The Honesty of Statisticians

It can be argued that in acknowledging our uncertainty and quantifying it, Statisticians are in some sense more honest in their statements about the world than others who make absolute claims. After all, we Statisticians don't claim to know things we can't know. Instead, for example, we offer an *interval* of plausible values for an unknown parameter. Not satisfied with that, we spend more effort describing exactly how uncertain we are that even that interval covers the true value and just what we must assume about unknown and unknowable features of the world for those estimates to be correct.⁸

What then is the source of the "damn lies" view? Statisticians are evidently taking great care to be honest, and readily admit their uncertainty. Liars usually assert their lies confidently in their striving to be believed. And when a statistician's conclusion turns out to be wrong, the error is not seen as deliberate deception. Another random sample may yield a different answer, but that isn't blamed on the statistician as a failure of ethical data collection or analysis.

I think that the problem lies in another aspect of uncertainty in Statistical analyses; the fact that Statistics, as a science, is not algorithmic or deterministic. The problem isn't that another sample may give a different answer, but that another statistician working with the same sample may give a different answer.⁹

For example, we find the American humorist [Evan Esar](#) (1899-1995) quipping in his *Comic Dictionary* that

Statistics [is] The only science that enables different experts using the same figures to draw different conclusions.

It seems that we have not made it clear to others – and especially not to our students—that good statistical analyses include judgments, and we have not taught our students how to make those judgments.

6. Judgment

The insight that statistical analyses require judgment is not new. Ironically, it can be found in comments by the statisticians who developed the statistical methods that most often encourage the misapprehension of Statistics as an algorithmic application of mathematics. It is in the area of hypothesis testing that we often see people apply statistics methods blindly, hoping for that statistically significant $P < .05$, but neglecting to employ their judgment. (Indeed, in some disciplines judgment in such analyses is discouraged.)

It is easy to find flow charts for the "hypothesis testing process"—sometimes called the "Scientific Method." But, the originators of the standard inference methods were clear advocates of judgment. [Egon Pearson](#), who, working with Jerzy Neyman, developed frequentist statistical inference wrote

We left in our mathematical model a gap for the exercise of a more intuitive process of personal judgment. (p. 395, cited by [Abelson](#)).

Ironically, the inference procedures developed by Neyman and [Pearson](#) are among the statistics methods most likely to be thought of and taught as mechanical data manipulations that lead to an automatic decision to reject or fail to reject the hypothesis.

Sir Ronald Fisher was the founder of experimental design and Analysis of Variance, and the statistician generally credited with proposing .05 as an acceptable P -value for statistical significance—perhaps the ultimate arbitrary decision made by most users of statistics.¹⁰ Nevertheless, he too found judgment important in reaching conclusions about data. Abelson notes that

Sir Ronald Fisher accused Neyman and [Pearson](#) of making overmechanical recommendations, himself emphasizing experimentation as a continuing process requiring a community of free minds making their own decisions on the basis of shared information. ([Abelson, 1995, p. 2](#))

The important feature of judgments required for statistical analysis is that they are anchored in our knowledge of the world. The "community of free minds making...decisions on the basis of shared information" that Fisher refers to is sharing information about their evolving understanding of how the world works. Statistical judgments should be based on such understanding. That is why Statistics is a required course in so many disciplines; the "shared information" of a scientific discipline provides the

basis for the judgments one must make when analyzing data.

Fortunately for our teaching, many of these judgments show up early in the first statistics course. For example:

- Whether a histogram should be considered skewed or symmetric,
- Whether a stray point is an outlier,
- Whether that second bump is a real mode (and the data possibly not homogeneous)
- Whether a scatterplot shows a straight enough pattern for correlation and regression to be appropriate,
- Whether the Normal probability plot or histogram of residuals shows a distribution near enough to Normal for t -based inference (or whether the sample size is big enough that this assumption doesn't really matter),
- Whether observational data from two groups can properly be considered paired (or alternatively, whether the groups can properly be considered independent,
- Whether, once we have rejected a null hypothesis, the size of the effect is large enough to be meaningful (and the importance of understanding the difference),
- When to re-express a variable and which re-expression to choose,
- Whether our sample is sufficiently representative (or conversely, whether we can think of important sources of bias),
- Whether the cases in our data are mutually independent (as required by almost all our inference procedures), and
- Whether the regression residuals are really homoskedastic.

As students get more sophisticated, we can discuss bias in data collection, failures of independence, lurking variables, and the need for clarity in defining exactly what has been recorded and about whom. Each of these also usually involves judgments that incorporate knowledge of the real world as well as about the data.

7. BS

Sometimes Statistics is used to lie. The need for judgment opens the door to unethical biasing of results, biased data collection, and partial reporting or manipulation of results with an intent to mislead. The American Statistical Association publishes *Ethical Guidelines for Statistical Practice*¹¹, which calls for (among other things):

The avoidance of any tendency to slant statistical work toward predetermined outcomes.

And

Exposure of dishonest or incompetent uses of statistics.

Some people do use Statistics as part of a deliberate lie. They know they have elected a biased sample or cherry-picked results with low P-values. Since they know what they say is false but nevertheless try to convince others, they are, in [Frankfurt's](#) words,

"...fakers and phonies who are attempting by what they say to manipulate ... opinions. What they care about primarily, therefore, is whether what they say is *effective* in accomplishing this manipulation. Correspondingly, they are more or less indifferent to whether what they say is true or whether it is false." (*On Bullshit*, 2005)

Their use of Statistics is what [Frankfurt](#) has defined as *bullshit*.

However, other sciences and social sciences can and have been distorted to make a false case without their disciplines being tarred as damn lies. We hear the occasional credentialed expert deny Evolution, claim that HIV doesn't cause AIDS, support Regression Therapy, or argue that slavery was good for slaves. These minority views don't lead the mass of mainstream scientists or informed lay-people to label entire disciplines as damn lies. Even when experts in other disciplines who are not out of the mainstream disagree publicly, we don't consider their differing judgments to be lies.

8. Magic

Why then are statistical analyses seen as worse than damn lies? I think the problem is to be found elsewhere. To many people, mathematics, cosmology, and nuclear physics are as mysterious as statistics, being beyond common comprehension. But frankly, most people don't care what they conclude; they are about worlds imagined, or too far away, or too small to seem to matter. However, the conclusions of Statistics are about *our* world. As an editorial in the *Lancet* of Jan 2, 1937 put it

"Statistics...tends to induce a strong emotional reaction in non-mathematical minds. This is because statisticians apply, to problems in which we are interested, techniques which we do not understand."

This places Statistics in the realm of [Arthur Clarke's Third Law](#), which tells us that

Any sufficiently advanced technology is indistinguishable from magic.
- *Profiles of the Future*

Statistics is an advanced technology that is about our everyday world, so it invokes the kind of magic that the wizards of myth and fiction use to change everyday circumstances.

The rules for this kind of magic are well established in tradition and literature:

- You must work the spell in the precise manner specified by the obscure lore.
- You do not need to understand the incantation (which may be in an arcane language). Indeed, there is no encouragement or advantage to doing so—it is *meant* to be obscure.
- Nevertheless, if you get it even slightly wrong then either it just won't work or, like the sorcerer's apprentice, you may cause a result you didn't intend.

Importantly, there is no room for judgment. However, the power of the magic grants a certain degree of authority to those who can (or appear to be able to) control it (and who can then invoke the prohibition against judgment to silence doubters.)

Of course, this is the antithesis of a scientific reasoning that incorporates judgment. Therein lies the conflict. There are, for example, those who want to simply recite the spells, make the appropriate mouse—rather, *wand*—motions, achieve statistical significance, and claim the authority of the magic as proof that their conclusions must be correct. They think—they *hope*—that science advances by running data through statistics programs and sifting for P-values less than .05. By taking this approach, they encourage fear and disrespect of the discipline. This is nothing more than stage magic, and those who can see through their weak reasoning may blame their tools as well. Once you see the wire holding up the floating lady, it is natural to dismiss the entire magic act as flim-flam.

9. ...and Statistics

Some consumers of Statistics want to believe in the magic of Statistics. They view Statistics as a type of Mathematics and are offended to learn that judgment is involved at all. After all, they think, the Pythagorean Theorem is not open to opinion. It is the very fact that judgment is a part of a statistical analysis that opens the question of the honesty of Statistics.

Judgments about how to collect data, about the quality of the data, and about what can honestly be inferred from the data are essential parts of a statistical analysis. Yet it can seem that those who practice statistics are ashamed to admit that they made judgments in the course of their work. Rather, they ought to be ashamed to hide those judgments from open view and consideration.

Many users of statistics prefer to present their conclusions as entirely objective and free of judgment. There's a temptation to adopt the mantle of mathematical objectivity, thinking that the resulting conclusions are somehow less open to argument. When judgments are necessary, there's the tradition in many fields of writing in passive voice. Decisions about the study design "are made;" analyses "were performed," using methods that "it was decided to use." The scientists and statisticians making the judgments are nowhere to be found, lest they be accused of tampering with the spells. But statistics without judgment is magic, not logic.

10. Modeling the World as a Regulative Ideal

What we require in statistical judgments is a dispassionate striving for honest judgments. Our guiding principle should be that we seek truth about the world. Judgments guided by that principle can help us frame the questions asked of a statistical analysis. Until we know why we are performing the analysis, it is difficult to know how to seek relevant answers honestly.

[William Hunter](#), advises that

At the outset the most important questions for the statistician to ask is: What is the objective of this investigation?

- "The Practice of Statistics: The Real World is an Idea Whose Time Has Come"

If we know what the issues are, it helps us to make sound judgments, and we must do so dispassionately.

What we want from judgments is that they follow the guiding ethical principle that we seek truth about the world. Of course, honesty in this context does not mean that our analyses will necessarily tell the truth, nor even that we will believe that they do. Rather, we are making an unbiased *effort* to discover the truth. That effort may even include maintaining alternative and even contradictory models of the world if all of them account well for the data we have and the facts we know. The best analysis often arises from Darwinian competition among alternative models. (John Walker has suggested that this competition be called "survival of the best-fit.") As the analysis proceeds, each model spawns new questions, predictions, and residuals to examine, with the result that some lines of reasoning prosper and others die out. This too, can lead some people who don't understand how Statistics really works, to think that Statisticians are playing at the truth rather than seeking it. Surely, they think, it makes no sense to support two conflicting models when they can't *both* be true. Statisticians, (who understand the value of randomized placebo-controlled experiments), realize that knowing that (at least) one model is false isn't the same as knowing which one.

In particular, Statistical reasoning is skeptical. The structure of inference requires that we reject the null hypothesis if we are to make progress. So we often work with models that may be false, and even with those that we hope are false. We can't just state what we believe to be true and gather evidence to support it. This is just sound scientific reasoning. It can be traced back almost 400 years to the founder of modern scientific reasoning, [Sir Francis Bacon](#):

And therefore it was a good answer that was made by one who when they showed him hanging in a temple a picture of those who had paid their vows as having escaped shipwreck, and would have him say whether he did not now acknowledge the power of the gods,

"Aye," asked he again, "but where are they painted that were drowned after their vows?"

And such is the way of all superstition, whether in astrology, dreams, omens, divine judgments, or the like.

--*Novum Organum*, Urbach translation^{[12](#)}

The antiquity of the struggle doesn't make it any easier to teach. It does allow us to point out, however, that even though our discipline is barely a century old in its main parts, the underlying philosophy of discovering truth about the world through observing it, and the difficulty of the judgments required are much older.

11. Teaching

Our best hope for changing the view of Statistics as damn lies is education. If we teach Statistics as a mechanistic muddle of magical methods, our students will conclude for themselves that it is a pack of damn lies. But we can do better:

1. We must tell our students that to use statistics they must make judgments, and that there may be no method guaranteed to arrive at the truth. This will distress those who were hoping to just plug new numbers from the exercises and exam questions into the algorithms and formulas found in the little boxes of the textbook.
2. We should advise students to know the motivating reason for the analysis because this will guide them in making these judgments. They should know who (or what) the cases are in the data, what has been measured or recorded about them (and in what units), and when that was done. Even definitions that sound reasonable should be questioned.
3. We should teach that the guiding principle in making statistics judgments is a search for truth about the world. Faced with judgment calls, we make the choice that best supports our efforts to model or understand the world as it is. Where that choice isn't clear, we make an honest attempt to make the best choice. It is fine to entertain alternative or contradictory models for as long as there are no data that allows us to choose among them.
4. We should teach students to resist jumping to conclusions, extrapolating, and proposing explanations for associations that assume causation. And we should teach them to be skeptical of reports of Statistics that don't meet these high standards.
5. Most important, we must present the entire subject as a search for understanding about the world when we have data so that the other principles have a foundation to stand on.

How well do we do these things now?

Most modern statistics texts define Statistics in ways compatible with principle (5). Here are some

examples from current texts:

[Statistics is] The art and science of designing studies and analyzing the data that those studies produce. Its ultimate goal is translating data into knowledge and understanding of the world around us. In short, statistics is the art and science of learning from data" ([Agresti and Franklin](#))

Statistics is a way of reasoning, along with a collection of tools and methods, designed to help us understand the world. ([Deveaux, Velleman, and Bock](#))

Statistics is the science of data. ([Moore](#))

Statistics is much more about the scientific method than anything else, determining research questions; designing studies; organizing, summarizing, and analyzing the data; interpreting results; and drawing conclusions. ([Rumsey](#))

Statistics is a collection of procedures and principles for gathering data and analyzing information in order to help people make decisions when faced with uncertainty. ([Utts and Heckard](#))

Do these texts—and do we as teachers—keep the goals of learning about the world and of making sound decisions based on data central to their discussions throughout the course? There is an inexorable tendency to focus on the definitions and methods and give shorter shrift to the larger goal. This can be especially difficult and dangerous as we discuss inference. The methods are technical, the reasoning follows many steps – too many for most of our students to remember at first—and the conclusions about the world are almost scripted.

We all remind students to plot the data and the residuals, but when the course gets technical and conceptually difficult, it's hard to keep the global view in mind. If we address judgments, we'll have to discuss how to deal with a failure of independence, or non-Normal residuals, or an occasional outlier. It is easier to shunt aside such considerations and talk instead about the formula for the standard error and how many degrees of freedom we have.

Also—not insignificantly—it is hard to write and grade exercises that are tied to real-world concerns and to require conclusions that address those concerns. How tempting to just check whether the P-value is right or whether the null has been rejected in accordance with the key. Publishers don't want to take the pages required to print long answers in the back of the books, so we may have only short numerical solutions without real-world explanations.

Nevertheless, we *can* offer students such exercises, and we can introduce exercises with the simple judgments I listed earlier. Inference methods require assumptions about the data and about the world that generated the data. Some of these assumptions, such as independence and randomization, we judge by understanding how the data were collected, who and what was measured, and by then using our common sense and knowledge of how the world works. Even when we have no data, just thinking about the circumstances is often enough to decide whether a distribution is skewed or a relationship nonlinear. Other assumptions, such as symmetry, normality, and linearity, we judge by displaying the data and judging what we see.

We must not make this just another step in the script that students learn to follow. We should discuss the ethics of honestly seeking truth about the world as a part of introducing Statistics to students. We should remind our students that because Statistics is about modeling and understanding the world through data, we have an obligation to use our goal of good models as a guiding principle in making judgments.

When we center the introductory statistics course on the search for truth about the world, we answer the question most of our students have at the start of the term: "Why am I taking this course?" Our traditional answer, "Because it's required." isn't a very good one. The answer supported by modern statistics education practice, "To help you learn to understand the world from data." is more convincing and more intriguing. The answer I propose, "Because you must make judgments, and those depend on the other knowledge you are gaining in your studies," integrates the course into students' entire course of study.

Moreover, by providing a central theme that operates throughout the introductory course, we can save it from being a picaresque tale that passes from graphics to randomness to sampling to probability to inference without much plot holding the incidents together. In its place, we can present a coherent story in which each new method adds insights and fits with the others.

12. Truth, Twain, and Teaching

Let's put the quotation that opened this discussion in its fuller context. This is what Mark [Twain](#) said:

I was deducing from the above that I have been slowing down steadily in these thirty-six years, but I perceive that my statistics have a defect: 3,000 words in the, spring of 1868, when I was working seven or eight or nine hours at a sitting, has little or no advantage over the sitting of to-day, covering half the time and producing half the output. Figures often beguile me, particularly when I have the arranging of them myself; in which case the remark attributed to Disraeli would often apply with justice and force: "There are three kinds of lies: lies, damn lies, and statistics." (1924 , p. 246)

Twain is engaged in some simple statistical analysis. He notes that although he is producing a smaller volume of writing in a day's work, in fact, he has maintained his production rate of about 375 words per hour. He's just working shorter days.

He notes that his initial statistics have a defect because they didn't allow for how long he now works each day. And that leads him to consider how statistics can mislead.

But what did Twain mean by the lies, damn lies... quip? To understand him, we'll need to apply some judgment. As a writer, he chose his words carefully. That makes his choice of "beguile" intriguing, because the word has several alternative meanings.

If the meaning of "beguile" he had in mind was "to pass time in a pleasant way," he would be inviting us to imagine him fashioning lies from figures for his own amusement. That seems doubtful even for the cynical, older Twain.

Most likely, he meant "to mislead or delude." Surely the rest of the quotation tends in that direction. Had Twain just said that "figures often beguile me; I agree with Disraeli when he said...", that would be the natural interpretation. But what should we make of Twain's qualification "particularly when I have the arranging of them myself" coupled with his demonstration of discerning the truth about his productivity by calculating a mean? He is gently chiding the unexamined statistics for misleading him and admitting that this often happens to him when he has "the arranging of them" himself. There is no intent to deceive, but rather the easy slide into a misinterpretation, caught by a more careful examination of the data. This meaning doesn't carry the accusation of deliberate falsification or deception that subsequent quoters of the phrase seem to think of as Twain's meaning. Rather, it is an admonition to look carefully at your data because (not being simple facts free of judgment) they may mislead you.¹³

I would like to think that Twain also had in mind a third meaning of *beguile*: "to win and hold somebody's attention, interest, or devotion; to charm or divert." In that interpretation, Twain could have meant that, when he is given the chance to work with (to "arrange") the numbers for himself, he finds them fascinating because he often discovers truths that might, at first, have been hidden in the data or in simplistic statistics summarizing them.

Given Twain's wide-ranging intellect, his sharp cynicism about humankind, and his full command of the English language, this might be one of the interpretations he intended. And if so, it is a prescient call for us to teach students to make sound, ethical, judgments when they interpret their data so that they, too, might be beguiled by Statistics.

End Notes

¹ <http://www1c.btwebworld.com/quote-unquote/>

² Number two is "The only thing necessary for the triumph of evil is for good men to do nothing." – Attributed to Edmund Burke, but without a firm citation.

³ <http://www.york.ac.uk/depts/maths/histstat/lies.htm>

⁴ It has been proposed by some that the punctuation suggests that Lord Courtney listed only two alternatives: Lies—damned lies, and statistics, but the rest of the phrase suggests otherwise.

⁵ Lord Courtney, as secretary of the treasury in 1882 had been one of those "Gladstonians," [Balfour](#) was raging against, but by 1892 he was moving away from the Liberals and often reached across party lines.

⁶ Throughout this essay, the term "science" should be taken to include the social and behavioral sciences as well – and indeed, any discipline that advances by the scientific method.

⁷ This quotation is widely, and plausibly, attributed to Asimov but without, so far as I've discovered, a citable source.

⁸ Our obsession with our ignorance doesn't end there. Papers that discuss how well or how badly we can estimate the size of a confidence interval are common.

⁹ One example of how two statisticians, each with sound (but subtly differing) interpretations can reach opposite conclusions from the same data is known as Lord's Paradox ([Lord, F.M., 1967](#)), "A Paradox in the Interpretation of Group Comparisons," *Psych Bull* 68, 304-305. For a good explanation, see [Wainer, H. and Brown, L.M.](#) Two statistical paradoxes in the interpretation of group differences: Illustrated with medical school admission and licensing data. *The American Statistician*, 58, 117-123, 2004.

¹⁰ Howard Wainer points out Fisher chose .05 to suit small exploratory studies in the context of ongoing research. His idea was that initial studies tend to be small, so having a too stringent criterion would make it too hard to find things. The cost of correcting a type I error is small (it typically happens with the first replication) whereas the cost of re-finding an effect that was missed is much larger. Hence the

generous criterion of one in twenty. [Wainer, H. & Robinson, D. \(2003\)](#).

¹¹ <http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics>

¹² Twain offers a similar insight in [Mark Twain's Notebook](#):

"Does the human being reason? No; he thinks, muses, reflects, but does not reason...That is, in the two things which are the peculiar domain of the heart, not the mind,--politics and religion. He doesn't want to know the other side. He wants arguments and statistics for his own side, and nothing more."

¹³ Like [Balfour](#), Twain's issue is with a changing denominator.

References

Abelson, R. P. (1995), *Statistics as principled argument*, Hillsdale, N.J., L. Erlbaum Associates.

American Statistical Association, *Ethical Guidelines for Statistical Practice*,
<http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics>

Agresti, A., and Franklin, C. (2007), *Statistics: the Art and Science of Learning from Data*, Upper Saddle River, Pearson Prentice Hall.

Bacon, Sir Francis, (1610) *The Novum Organum; with other parts of, The great instauration* translated and edited by Peter Urbach and John Gibson (1994), Chicago, Open Court.

Balfour, Arthur James, (1892) "Mr. Balfour's Reply to Professor Munro," *Manchester Guardian*, Wednesday, June 29, 1892 Page 5.

Clarke, Arthur, (1973) *Profiles of the Future*, New York, Harper & Row

Courtney, Leonard Henry, (1895) "To my Fellow Disciples at Saratoga Springs," *The National Review* (London) **26**, 21–26.

Deveaux, Velleman, Bock, (2009) *Intro Stats, Third Edition*, Boston, Addison Wesley.

Esar, Evan, (1943) *Esar's Comic Dictionary*, Harvest House.

Frankfurt, Harry G., (2005), *On Bullshit*, Princeton, NJ, Princeton University Press.

Guy, William A, (1865) "On the Original and Acquired Meaning of the term 'Statistics,' and on the Proper Functions of a Statistical Society: also on the Question whether there be a Science of Statistics and, if so, what is its Relation to Political Economy and 'Social Science'," *Journal of the Statistical Society of London*, **28**, No. 4, Dec, pp 478-493.

Hunter, William G. (1981), "The Practice of Statistics: The Real World is an Idea Whose Time Has Come," *The American Statistician*, **35**: 2, 72-76.

- Huxley, Thomas Henry (1959), "Biogenesis and Abiogenesis" in *Collected Essays*, New York, Harper.
- Lord, F.M. (1967), "A Paradox in the Interpretation of Group Comparisons," *Psych Bull* 68, 304-305.
- Moore, D.S. (2007), *The Basic Practice of Statistics, Fourth Edition*, New York, W.H. Freeman.
- Pearson, E.S. (1962). "Some Thoughts on Statistical Inference," *Annals of Mathematical Statistics*, **33**, 394-403.
- Rumsey, D. (2003), *Statistics for Dummies*, Indianapolis, Wiley Publishing.
- Twain, Mark, (1924), *Autobiography*, A. B. Paine (ed), New York and London: Harper Brothers Vol. I.
- ____ *Mark Twain's Notebook*, A. B. Paine (ed), New York, Cooper Square Publishers (1972)
- Utts, J.M. and Heckard, R.F. (2007), *Mind on Statistics*, Belmont, Duxbury, Thomson Brooks/Cole.
- Wainer, H. and Brown, L.M. Two statistical paradoxes in the interpretation of group differences: Illustrated with medical school admission and licensing data. *The American Statistician*, 58, 117-123, 2004
- Wainer, H. Robinson, D. (2003). Shaping Up the Practice of Null Hypothesis Significance Testing. *Educational Researcher*, 32(7), 22-30.

Paul F. Velleman
Departments of Statistical Science and of Social Statistics
Cornell University
Ithaca, NY 14853
U.S.A.
pfv@cornell.edu
