7 | Mathematical Deception

Until now, we've mostly been talking about how words can be used to convince or bamboozle. But numbers have their own special power when it comes to explanation and persuasion, a subject explored in detail in Charles Seife's 2011 book *Proofiness: How You Are Being Fooled by the Numbers*, which points out the psychological and cultural factors behind our near-religious faith in quantitative information.

In Western culture, this faith goes back to our old friends the Ancient Greeks. While most students only know the early mathematician Pythagoras through his famous geometry formula of $a^2 + b^2 = c^2$ (with A and B being the length of the legs of a right triangle and C being the length of the hypotenuse), Pythagoras was also a philosopher, some would say a cult leader, whose followers joined him in believing in the near mystical power of numbers, mathematics, and geometry.

Some stories regarding Pythagoras are likely myths, such as the ones saying he had a thigh made of gold and could dematerialize and perform other feats of magic. But the historic Pythagoras, or at least his philosophy, certainly seems to have influenced Plato, whose work can be seen as trying to find equivalents in human life to the perfection of numbers.

For when you think about it, numbers are not only perfect but may represent the only perfect things we encounter in life. Words change their meaning, people their nature, mountains can be leveled with an earthquake, and great lakes can dry up, but two-plus-two will always equal four. That is true wherever you live and whatever culture you belong to, and it will continue to be true even if you choose not to believe it.

I don't think it was an accident when, in the novel *1984*, George Orwell used as an example of total submission to Big Brother getting someone to both swear and truly believe that 2 + 2 equaled five since this represents a denial of one of the few absolute truths we can conceive of as human beings.

A problem arises, however, once you take that number out of its abstract realm of perfection and drag it down to the real world by attaching a unit to it.

As Seife points out, 2 + 2 may equal 4, but what happens when you add the length of two pieces of wood, each of which is two feet? Well, if you've got a good enough ruler, it will turn out that neither of those pieces of wood is exactly two feet long. Both will be slightly longer or shorter, even if only by a fraction of a fraction of an inch. Which means if you add these alleged twos together, you won't get four, you'll get a number that is slightly less or more than four.

At first, this might seem like a ridiculous concern. After all, we have to accept a certain level of inexactitude as irrelevant if we want to measure a piece of wood to build a cabinet, as well as accomplish most of the other things we do in our lives that involve measurement and practical arithmetic. But what happens when the level of uncertainty inherent when numbers get applied to real things become not just significant but critical?

We can start thinking about this not by looking at complex economic formulas or statistics but instead talking about the simplest of all mathematical operations: counting things. Surely something as ridiculously simple as counting comes as close to the Pythagorean/Platonic ideal of perfect numbers as we can in real life, right?

Wrong! For if you hand a complete deck of cards to a dozen different people and ask them to count them up, I guarantee you that at least one person will tell you that you've got one too many or are missing one. This is just an example of human frailty. The mind wanders, especially when doing repetitive tasks. You can easily get distracted, which is why kids have such fun screaming out random numbers while a parent or friend is trying to tally something up. Or some people are just plain better at this skill than others.

This is why companies spend days doing warehouse inventories; they don't trust people to get it right the first time and would prefer to spend more time and money repeating the count over and over until they get what they believe to be the "correct" numbers.

Now scale this counting/inventorying process up further to an exercise where we have to count millions of pieces of paper, in multiple formats, each with different markings, collected from different places at different times, all of which pass through multiple people's hands on their way to being tallied.

Welcome, folks, to the 2000 presidential election.

Actually, welcome to every presidential election, including the next one. In fact, welcome to any election ever. For every one of them ultimately requires a vote to be tallied, and given that you'll get errors when you ask a dozen friends to count fifty-two identically sized and shaped playing cards, what are the chances that you'll get even more errors when you count millions of differently shaped, worded, and managed ballots?

I'll tell you what it is: it's 100 percent. In every election that's ever been held of any scale (I'm ignoring votes for classroom hall monitor where the votes of thirty elementary school students can be checked and double-checked accurately enough), there has been a margin of error of several hundred or several thousand votes.

Normally we don't take any notice of this problem because the size of the victory overwhelms this error factor to such a degree that it is irrelevant, just as the difference between two feet and 2.001 feet is irrelevant for most carpentry projects. But when the vote is close, that is, when it is within the range of error, the whole system breaks down.

And by breakdown I mean the 2000 election with its hanging chads, delayed decision on who's really the president, courtroom battles, and near constitutional-crisis levels of chaos.

If you think back to that vote, much of this chaos arose from the fact that no one—not the candidates, not the media, not even the public—was willing to accept the fact that there might not be an "accurate" number lurking somewhere. If we can only put enough time and money into counting and recounting the ballots, we told ourselves, we will get to this "true" figure. After all, 2 + 2 must equal 4.

At first, it was assumed that this problem resulted not from the human condition but from unusual conditions related to Florida. The ballots were poorly designed, this politician or that was in someone's pocket, vote counters were given irrational directions, etc., etc. But even when states that learned lessons from Florida and made their balloting processes more consistent and rational had a close vote (as Minnesota did during their 2008 US senate race) the same problem arose when the victory margin was within the margin of error.

Keep in mind that I'm not talking about candidates and their lawyers trying to put their thumbs on the scale to tip the vote count their way by including their supporters and excluding their opponents. That's just one more human factor to add to a mix that already separates a mathematical operation that exists in the highly messy real world (counting ballots) from the metaphysical perfection that we believe numbers must possess.

If we could bring ourselves to accept that finding the true, accurate, final number in any close vote involving complex balloting might not be possible, then we might make changes that could help us deal with the fact that there is mathematical ambiguity in the real world, changes such as requiring a do-over in such close elections that could replace the current method for breaking a tie in many states: the flipping of a coin. But our devotion to the notion of mathematical perfection means we'll continue to spend millions of dollars and live in chaos, all to ensure we never have to admit that in some numerical situations "the answer" might not exist.

So if just counting stuff can get us into this much trouble, what happens when we get into more complex mathematical matters like, say, polling.

Nothing demonstrates our devotion to quantitative information better than our fascination with poll results, even from nonsense polls like ones you can plug onto a Web page these days without having to write a single line of code.

When such polls are used to gauge innocent matters like ice-cream preference, there is little harm done. But whenever they are used to measure important political preferences, such as support for one presidential candidate vs. another, their key flaw of lack of control over who gets to vote becomes apparent. To cite two examples, I routinely get invitations e-mailed to me to participate in ballot stuffing on this or that Internet poll, and I've even heard of a presidential candidate (who shall not be named) who paid supporters to vote for him whenever his name appeared on any poll found anywhere on the Web.

But what about professional polls that people pay thousands of dollars to have experts in survey techniques and statistics put together?

Again, when you're dealing with uncontroversial subjects, such surveys are extraordinarily useful. They underlie a century of successful social science and market research, after all.

But once you start talking about things like who's ahead and who's behind in the presidential race or where the country stands on particular political issues, our lack of understanding of what a poll is really saying, especially with regard to assumptions of accuracy, begins to bite.

For instance, as Seife points out in *Proofiness*, every poll contains something called a *margin of error*, often stated in the form of plus-or-minus some number, which alerts people to a statistical factor that says the “actual” number may be slightly more or slightly less than the number cited as the result of the poll.

For example, if a poll says that candidate Smith is ahead of candidate Jones by 54 percent to 46 percent with a margin of error of plus or minus three, that would indicate that Smith is likely to win since, even if you assume the poll is off by the largest margin of 3 percent, Smith still beats Jones 51 to 49 percent.

This statistical margin of error affects every poll, although it can be brought down by increasing your sample size. A poll of a million voters, for example, will almost certainly have a smaller margin of error than the same poll given to just one thousand people.

But we often take this statistical margin of error to be the only possible variation in poll results. In other words, we take it to be telling us the full range of possible inaccuracy in the poll, even if there might be all kinds of other reasons why those results could be wrong or meaningless.

For example, the sample selected to stand in for the whole may have been chosen poorly or the overall sample might be too small. Survey questions may have been confusing or intentionally designed to push people one way or another. People may lie to the pollsters or people with one political opinion might tend to hang up on telephone pollsters more often than others. The poll may have been designed before a major news story changed the dynamic of a political race. Or maybe someone being paid to collect information screwed up the data entry (remember the human factor) or a PhD in statistics forgot to carry the two.

Such systematic errors may only affect poll results slightly, but if you look at that Smith and Jones example I just used, a one to two percent systematic error can mean the difference between who is winning and losing.

Even if a scientific poll is done perfectly (whatever that means), remember that a poll is simply a snapshot of opinions at a particular point in time. But given how bad polls are at predicting the future (compare what the polls were telling you six months ago vs. today if you happen to be reading this during campaign season to see what I mean), we need to look at polls for what they are: just one more type of information that needs to be evaluated for quality, especially timeliness.

Just as the chaos of the 2000 Florida election was the result of our unwillingness to shake our faith in numbers, our belief in polling data also arises from the extraordinary belief we attach to quantitative information. Like any belief, our readiness to treat numerical data more respectfully than we might other forms of information can be turned against us by those trying to get us to do what they want us to do.

In fact, there are a number of ways numbers are used fallaciously in political argumentation. For instance, there is that old political and business standby: the *unit fallacy*.

This is one you should always look out for when people talk about percentages or rates. In *Thank You for Arguing*, Jay Heinrichs illustrates this technique with the story of a business presentation where someone boasts that company profits are up 20 percent. “Wow!” you say. So since we made 10 percent profit last year, that means we’re up to 30 percent! “Actually,” you’re sheepishly told, we’re only up to 12 percent.

Where did that 12 percent come from? Well if you apply the 20-percent growth rate to last year’s 10-percent profit percentage, 20 percent of 10 percent is two percent. So going from 10 to 12 percent represents a 20-percent increase in the *profit percentage*, not 20-percent growth in actual profit. But saying “profits are up 20 percent” sounds so much better than saying the profit percentage grew by two percent, demonstrating how the unit fallacy can leverage mathematical ambiguity to make meek improvement sound like fantastic success.

You see this all the time in budget politics when someone shows off his or her fiscal responsibility by telling you, for example, that "the rate of growth in government spending has been lower during my administration than in any other time in history." So the budget is smaller than when you came into office? No, it's bigger than ever. But it's not getting any bigger, right? Actually it's still growing so by definition it's getting bigger. But it's getting bigger more slowly than it has in the past.

Feeling like a sucker yet?

Seife, in *Proofiness*, highlights a number of numeric fallacies that travel under fruity names such as *comparing apples to oranges*. This is where you compare two things that sound similar but aren't. Like when President Obama boasted in 2012 that he had created more jobs during the last twenty-seven months of his administration than his predecessor had during both his terms combined.

Pretty impressive, unless you stop to ask why he chose to count just the last twenty-seven months of his administration, ignoring most of his first two years in office. Might that be because these most recent months were the best ones for job creation during his term, meaning including previous months would drive his numbers down? And why compare that to Bush's entire term of office unless comparing Obama's twenty-seven-month apple to Bush's eight-year orange gave him the numbers he needed to make the political point he wanted to make and, just as importantly, his supporters wanted to hear.

That last example is actually a combination of two fruity fallacies. In addition to comparing apples to oranges, Obama was also guilty of *cherry picking*, choosing just the data that served his needs and ignoring numbers that didn't do the job quite as well.

You saw the same thing in a 2012 TV ad that attacked Mitt Romney's time as governor of Massachusetts, a state which, the ad claimed, fell to forty-seventh in the ranking of US states with regard to job creation. Pretty damning unless you realize that when Romney took office, Massachusetts was totally in the tank and thus was the worst state (fiftieth) in terms of job creation. Ignoring the question of how you can “fall” from fiftieth to forty-seventh place, forty-seventh place is still pretty lousy. So didn’t that still demonstrate the governor’s stinky record as a turnaround artist?

Perhaps. But by the time Romney left office, Massachusetts had risen to twenty-eighth in this particular ranking. So where did that ranking of forty-seven come from? It came from calculating job creation over the course of Romney’s entire four years in office, which masked the more relevant fact (an improving trend) by using an average instead of looking at more meaningful year-over-year improvement.

These last examples also demonstrate the problem that numbers rarely tell you everything you need to know to understand a complex story. Job creation statistics, for instance, may not provide the economic context in which these statistics played out.

Massachusetts, for example, is affected by the ups and downs of the high-tech industry more than other states. So it can drop to last in the nation or jump from fiftieth to forty-seventh or twenty-eighth based primarily on how the global tech economy is doing, something that may or may not have much to do with who’s sitting in the governor’s office. Similarly, any statistics regarding the economic performance of a president or his predecessor needs to take into account (and include information on) when an economic recession may have begun and ended during either administration or before an administration took office.

Generally, whenever you hear a complex situation explained with a simple and stunning number, watch out, even if (actually especially if) you’re inclined to like the candidate or believe in the issue that such a stunning number supports. This is because “Proofiness” is generally used to convince allies by giving them misleading numbers that support what they already believe. In other words, it's a particularly powerful and pernicious way of playing to someone's confirmation bias.

For instance, in recent years we have been treated to such a stunning number in reports claiming that, despite all of their advances in their workforce over the last several decades, women were still earning just seventy-five cents for every dollar earned by a man, a "huge discrepancy" that demonstrates the continuation of gender discrimination in the workplace.

Or does it? In this case, there are a number of confounding factors that provided a reasonable explanation for *some* of this 25-percent factor, such as how professions are grouped.

For example, highly paid surgeons, where men still predominate, and lower-paid pediatricians, which divides equally between men and women, were all grouped together under the category of "Physicians" to demonstrate gender-based pay disparity within a field. And that 75% statistic does not tell us whether job trends might be in the process of rectifying such imbalances. Then you have failures to include factors such as how choices and timing regarding balancing work and family can play out when calculating averages that span the long period between graduation and retirement.

Keep in mind that this criticism of numbers and how they are presented does not mean women are not discriminated against in the workplace in any number of ways or that the fight against such discrimination should not continue. But it does mean that anyone who feels this way should take extra care to avoid building their case on a powerful-sounding statistic that could be easily discredited because it's a form of "Proofiness." For, as with most confirmation-bias problems, this could end up undermining the very cause you think you are championing.

I realized that I was a little harder on Democrats than Republicans in examples used in this chapter. But the mathematical fallacies you've just read about are so common that it should be a cinch to find Republicans using the same tricks during the next (or any) election cycle.

In fact, the way candidates usually deal with getting caught engaging in Proofiness is to use another fallacy: the *tu quoque fallacy* (which translates to "you, also" or "you, another"). Yup, this is the "so's your old man" fallacy that tries to get out of being caught pulling a fast one by pointing out that your opponent does the same thing all the time.

Like most ad hominem fallacies, tu quoque can be both handy and fun since it lets you turn from defender to attacker without having to admit you got caught trying to mislead, either willingly or inadvertently. But it's not that useful if you're trying to actually get to the truth, the ultimate goal of critical thought.

Given the power numbers will continue to have over us, both to inform and deceive, we must resist the urge to cast away our critical eye when numerical information is invoked, just as we must resist the temptation to be blown away by a particularly persuasive and compelling speaker. For such a speaker might impress you because they really know what they're talking about. But they might also just be skilled at sounding like they do.

And how to tell one from the other is the subject of the next chapter.