DANIEL LEVITIN

From the bestselling author of *The Organized Mind*

WEAPONIZED LIES:

how to think critically in the post-truth era

'Critical thinking for our shrill, data-drenched age . . . every page is enlightening' Charles Duhigg, author of *The Power of Habit* and *Smarter, Faster, Better*

Contents

ABOUT THE AUTHOR

Dr Daniel J. Levitin has a PhD in Psychology, and trained at Stanford University Medical School and the University of California Berkeley. He is the author of the number-one bestseller *This Is Your Brain On Music* (Dutton, 2006), published in nineteen languages, and the bestsellers *The World in Six Songs* (Dutton, 2008) and *The Organized Mind* (Viking, 2014). Currently he is Dean of Social Sciences at the Minerva Schools at KGI in San Francisco, and a faculty member at the Center for Executive Education in the Haas School of Business at UC Berkeley.

PENGUIN BOOKS

WEAPONIZED LIES

‘The world is awash with data, but not always with accurate information. [This book] does a terrific job of illustrating the difference between the two with precision – and delightful good humour’ Charles Wheelan, author of *Naked Economics*

‘Daniel Levitin’s field guide is a critical-thinking primer for our shrill, data-drenched age. From the way averages befuddle to the logical fallacies that sneak by us, every page is enlightening’ Charles Duhigg, author of *The Power of Habit* and *Smarter, Faster, Better*

‘A guide for those who wish to test the authenticity of information that inundates us from every corner, dark and light, of the Web’ *Washington Post*

‘[This book] by the neuroscientist Daniel Levitin lays out the many ways in which each of us can be fooled and misled by numbers and logic, as well as the modes of critical thinking we will need to overcome this’ *Wall Street Journal*

‘Smart, timely and massively useful’ *Globe and Mail*, Toronto

‘Much like Nate Silver’s (*New York Times* bestselling!) *The Signal and the Noise*, Levitin’s is that rare book that makes statistics both understandable and at times even intriguing’ *MacLean’s*

‘A valuable primer on critical thinking that convincingly illustrates the prevalence of misinformation in everyday life’ *Publishers Weekly*

‘The timing could not be better … a survival manual for the post-factual era. Levitin offers a set of intellectual tools to help distinguish the real from the unreal, and often surreal … both engaging and rewarding’ *Literary Review of Canada*

‘Valuable tools for anyone willing to evaluate claims and get to the truth of the matter’ *Kirkus Reviews*

‘Misinformation is a curse of the information age, and Levitin offers blow-by-blow demonstrations of how words, numbers and graphics can be manipulated to distort truth’ *Stanford Magazine*

‘Levitin talks about the crucial role of critical thinking and seeking out the truth in today’s media landscape’ Michael Krasny, NPR Forum

To my sister Shari,
whose inquisitive mind made me a better thinker

took place. It appears that this ignorant citizen does not know what it is to compile and evaluate evidence. In this case, one might look for a link between Hillary Clinton and the restaurant, behaviors of Clinton that would suggest an interest in running a prostitution ring, or even a motive for why she might benefit from such a thing (certainly the motive could not have been financial, given the recent kerfuffle over her speaking fees). He might have observed whether there were child prostitutes and their customers coming in and out of the facility. Or, lacking the mentality and education to conduct one's own investigation, one could rely on professionals by reading what trained investigative journalists have to say about the story. The fact that no dedicated professional journalist gives this any credence should tell you a lot. I understand that there are people who think that journalists are corrupt and co-opted by the government. The U.S. Bureau of Labor Statistics reports that there are 45,790 reporters and correspondents. The American Society of News Editors, an independent trade group, estimates there are 32,900 reporters working for the nearly 1,400 daily newspapers in the United States. Some journalists may well be corrupt, but with this many of them, it's very unlikely that they all are.

Facebook is making an effort to live up to its social responsibilities as a source of information by "making it easier for its 1.8 billion members to report fake news." In other words, to call a lie a lie. Perhaps other social media sites will take an increasingly curatorial role in the future. At the very least, we can hope that their role in weaponizing lies will decrease.

Many news organizations looked into where the story of the sex-slave pizzeria originated. NBC reported on a thriving community of "fake news" fabricators in the town of Veles, Macedonia, who could well have been the source. This region was in communist Yugoslavia until 1991. *BuzzFeed* and the *Guardian* found more than 100 fake news domain names originating there. Young people in Veles, without any political affiliation to US political parties, are pushing stories based on lies so that they can garner significant payments from penny-per-click advertising on platforms such as Facebook. Teenagers can earn tens of thousands of dollars in towns that offer little economic opportunity. Should we blame them for the gunshots in the pizzeria? Social networking platforms? Or a US educational system that has created citizens complacent about thinking through the claims we encounter every day?

You might object and say, “But it’s not my job to evaluate statistics critically. Newspapers, bloggers, the government, Wikipedia, etc., should be doing that for us.” Yes, they should, but they don’t always, and it’s getting harder and harder for them to keep up as the number of lies proliferates faster than they can knock them down. It’s like a game of whack-a-mole. The Pizzagate story received more than one million hits, while its debunking by *Snopes* received fewer than 35,000. We are fortunate to have a free press; historically, most nations have had much worse. We should never take the media’s freedom and integrity for granted. Journalists and the companies that pay them will continue to help us identify lies and defuse them, but they can’t accomplish this on their own—the lies will win if we have a gullible, untrained public consuming them.

Of course most of us would not believe that Hillary Clinton was running a sex-slave ring out of a Washington, DC, pizzeria. But this book isn’t just about such absurdities. Do you really need this new drug or is the billion-dollar marketing campaign behind it swaying you with handpicked, biased pseudo-data? How do we know if a celebrity on trial is really guilty? How do we evaluate this investment or that, or a set of contradictory election polls? What things are beyond our ability to know because we aren’t given enough information?

The best defense against sly prevaricators, the most reliable one, is for every one of us to learn how to become critical thinkers. We have failed to teach our children to fight the evolutionary tendency toward gullibility. We are a social species, and we tend to believe what others tell us. And our brains are great storytelling and confabulation machines: given an outlandish premise, we can generate fanciful explanations for how they might be true. But that’s the difference between creative thinking and critical thinking, between lies and the truth: the truth has factual, objective evidence to support it. Some claims *might* be true, but truthful claims *are* true.

A Stanford University study of civic online reasoning tested more than 7,800 students from intermediate school through college for eighteen months, ending June 2016. The researchers cite a "stunning and dismaying consistency. Overall, young people's ability to reason about the information on the Internet can be summed up in one word: bleak." They were horrible at distinguishing high-quality news from lies. We need to start teaching them to do so now. And while we're at it, the rest of us could use a refresher course. Fortunately, evidence-based thinking is not beyond the grasp of most twelve-year-olds, if only they are shown the way.

Many said that Pizzagate was a direct result of fake news—but let's call it like it is: lies. There is no "news" in fake news. Belief in lies can be harmless, such as belief in Santa Claus or that these new jeans make me look thin. What *weaponizes* the lies is not the media nor Facebook. The danger is in the intensity of that belief—the unquestioning overconfidence that it is true.

Critical thinking trains us to take a step back, to evaluate facts and form evidence-based conclusions. What got Welch into a situation of discharging a firearm in a DC pizza parlor was a complete inability to understand that a view he held might be wrong. The most important component of the best critical thinking that is lacking in our society today is humility. It is a simple yet profound notion: If we realize we don't know everything, we can learn. If we think we know everything, learning is impossible. Somehow, our educational system and our reliance on the Internet has led to a generation of kids who do not know what they don't know. If we can accept that truth, we can educate the American mind, restore civility, and disarm the plethora of weaponized lies threatening our world. It is the only way democracy can prosper.

Three Kinds of Strategic Defense

I started writing this book in 2001, while teaching a college course on critical thinking. I worked on it in earnest during 2014–2016, and published it with a different introduction and the title *A Field Guide to Lies*. Since then, the dangerousness and reach of lies has become overwhelming. They are no longer just things that people can snark at or giggle over—they have become weapons. This danger may get worse, it may lead to troubles that we have not witnessed for generations. Or it may pass without such drastic consequences. In any case the tools offered here are the same as in the first edition; they are necessary tools, irrespective of the political, social, and economic winds.

Part of the problem is one of source. In the old days, factual books and news articles simply looked authentic, compared to a screed that some nut might have printed in their basement on a home printing press. The Internet has changed that, of course. A crank website can look as authentic as an authoritative, fact-checked one—I give examples later in this book. Misinformation is devilishly entwined on the Internet with real information, making the two difficult to separate. And misinformation is promiscuous—it consorts with people of all social and educational classes and turns up in places you don't expect it to. It propagates as one person passes it on to another and another, as Twitter, Facebook, Snapchat, Instagram, Tumblr, and other social media spread it around the world; the misinformation can take hold and become well known, and suddenly a whole lot of people are believing things that aren't so.

This is a book about how to spot problems with the facts you encounter, problems that may lead you to draw the wrong conclusions. Sometimes the people giving you the facts are hoping you'll draw the wrong conclusion; sometimes they don't know the difference themselves. Today, information is available nearly instantaneously, national leaders show up in your social media accounts, reports of "breaking news" grab your attention daily, even hourly, but when is there time to determine if that new information is packed with pseudo-facts, distortions, and outright lies? We all need efficient strategies for evaluating whether what we are being told is trustworthy.

We've created more human-made information in the past five years than in all of human history before them. Found alongside things that are true is an enormous number of things that are not, in websites, videos, books, and on social media. This is not just a new problem. Misinformation has been a fixture of human life for thousands of years and was documented in biblical times and classical Greece. The unique problem we face today is that misinformation has proliferated and lies can be weaponized to produce social and political ends we would otherwise be safeguarded against.

In the following chapters, I've grouped these strategies into categories. The first part of this book is about numerical misinformation. It shows how mishandled statistics and graphs can give a skewed, grossly distorted perspective and cause us to draw faulty conclusions (and make unsound decisions). The second part of the book investigates faulty arguments, showing how easy it is to be persuasive, to tell stories that drift away from facts in an appealing yet misguided fashion. Included along the way are the steps we can take to better evaluate news, advertisements, and reports. The last part of the book reveals what underlies our ability to determine if something is true or false: the scientific method. It is the best tool ever invented for discovering the most challenging mysteries, and it traces its roots back to some of the greatest thinkers in human history, figures such as Aristotle, Bacon, Galileo, Descartes, Semelweis, and Popper. This last part of the book grapples with the limits of what we can and cannot know, including what we know right now and don't know just yet. I offer a number of case studies in order to demonstrate the applications of logical thinking to quite varied settings, spanning courtroom testimony, medical decision making, magic, modern physics, and conspiracy theories.

Critical thinking doesn't mean we disparage everything; it means that we try to distinguish between claims with evidence and those without.

It is easy for partisans to lie with statistics and graphs because they know that most people think it will take too much time to look under the hood and see how they work. Maybe they think that they aren't smart enough. But anyone can do this, and once you have some basic principles, charts quickly reveal their elegance—or disfigurement.

Take the statistic I quoted earlier, about how the number of books students read declines steadily every single year after second grade. The implication is that our educational system is flawed—children are not developing good learning habits, they're not interested in bettering themselves, and they're not intellectually engaged. Now stop and ask yourself: is *number of books* the right metric for drawing conclusions about this? Second

Plausibility

Statistics, because they are numbers, appear to us to be cold, hard facts. It seems that they represent facts given to us by nature and it's just a matter of finding them. But it's important to remember that *people* gather statistics. People choose what to count, how to go about counting, which of the resulting numbers they will share with us, and which words they will use to describe and interpret those numbers. Statistics are not facts. They are interpretations. And your interpretation may be just as good as, or better than, that of the person reporting them to you.

Sometimes, the numbers are simply wrong, and it's often easiest to start out by conducting some quick plausibility checks. After that, even if the numbers pass plausibility, three kinds of errors can lead you to believe things that aren't so: how the numbers were collected, how they were interpreted, and how they were presented graphically.

In your head or on the back of an envelope you can quickly determine whether a claim is plausible (most of the time). Don't just accept a claim at face value; work through it a bit.

When conducting plausibility checks, we don't care about the exact numbers. That might seem counterintuitive, but precision isn't important here. We can use common sense to reckon a lot of these: If Bert tells you that a crystal wineglass fell off a table and hit a thick carpet without breaking, that seems plausible. If Ernie says it fell off the top of a forty-story building and hit the pavement without breaking, that's not plausible. Your real-world knowledge, observations acquired over a lifetime, tells you so. Similarly, if someone says they are two hundred years old, or that they can consistently beat the roulette wheel in Vegas, or that they can run forty miles an hour, these are not plausible claims.

What would you do with this claim?

> In the thirty-five years since marijuana laws stopped being enforced in California, the number of marijuana smokers has doubled every year.

Plausible? Where do we start? Let's assume there was only one marijuana smoker in California thirty-five years ago, a very conservative estimate (there were half a million marijuana arrests nationwide in 1982). Doubling that number every year for thirty-five years would yield more than 17 billion—larger than the population of the entire world. (Try it yourself and you'll see that doubling every year for twenty-one years gets you to over a million: 1; 2; 4; 8; 16; 32; 64; 128; 256; 512; 1024; 2048; 4096; 8192; 16,384; 32,768; 65,536; 131,072; 262,144; 524,288; 1,048,576.) This claim isn't just implausible, then, it's impossible. Unfortunately, many people have trouble thinking clearly about numbers because they're intimidated by them. But as you see, nothing here requires more than elementary school arithmetic and some reasonable assumptions.

Here's another. You've just taken on a position as a telemarketer, where agents telephone unsuspecting (and no doubt irritated) prospects. Your boss, trying to motivate you, claims:

> Our best salesperson made 1,000 sales a day.

Is this plausible? Try dialing a phone number yourself—the fastest you can probably do it is five seconds. Allow another five seconds for the phone to ring. Now let's assume that every call ends in a sale—clearly this isn't realistic, but let's give every advantage to this claim to see if it works out. Figure a minimum of ten seconds to make a pitch and have it accepted, then forty seconds to get the buyer's credit card number and address. That's one call per minute (5 + 5 + 10 + 40 = 60 seconds), or 60 sales in an hour, or 480 sales in a very hectic eight-hour workday with no breaks. The 1,000 just isn't plausible, allowing even the most optimistic estimates.

Some claims are more difficult to evaluate. Here's a headline from *Time* magazine in 2013:

> More people have cell phones than toilets.

What to do with this? We can consider the number of people in the developing world who lack plumbing and the observation that many people in prosperous countries have more than one cell phone. The claim seems *plausible*—that doesn't mean we should accept it, just that we can't reject it out of hand as being ridiculous; we'll have to use other techniques to evaluate the claim, but it passes the plausibility test.

Sometimes you can't easily evaluate a claim without doing a bit of research on your own. Yes, newspapers and websites really ought to be doing this for you, but they don't always, and that's how runaway statistics take hold. A widely reported statistic some years ago was this:

> In the U.S., 150,000 girls and young women die of anorexia each year.

Okay—let's check its plausibility. We have to do some digging. According to the U.S. Centers for Disease Control, the annual number of deaths *from all causes* for girls and women between the ages of fifteen and twenty-four is about 8,500. Add in women from twenty-five to forty-four and you still only get 55,000. The anorexia deaths in one year cannot be three times the number of *all* deaths.

In an article in *Science*, Louis Pollack and Hans Weiss reported that since the formation of the Communication Satellite Corp.,

> The cost of a telephone call has decreased by 12,000 percent.

If a cost decreases by 100 percent, it drops to zero (no matter what the initial cost was). If a cost decreases by 200 percent, someone is paying *you* the same amount you used to pay *them* for you to take the product. A decrease of 100 percent is very rare; one of 12,000 percent seems wildly unlikely. An article in the peer-reviewed *Journal of Management Development* claimed a 200 percent reduction in customer complaints following a new customer care strategy. Author Dan Keppel even titled his book *Get What You Pay For: Save 200% on Stocks, Mutual Funds, Every Financial Need.* He has an MBA. He should know better.

Of course, you have to apply percentages to the same baseline in order for them to be equivalent. A 50 percent reduction in salary cannot be restored by increasing your new, lower salary by 50 percent, because the baselines have shifted. If you were getting

$1,000/week and took a 50 percent reduction in pay, to $500, a 50 percent increase in that pay only brings you to $750.

Percentages seem so simple and incorruptible, but they are often confusing. If interest rates rise from 3 percent to 4 percent, that is an increase of 1 percentage point, or 33 percent (because the 1 percent rise is taken against the baseline of 3, so 1/3 = .33). If interest rates fall from 4 percent to 3 percent, that is a decrease of 1 percentage point, but not a decrease of 33 percent—it's a decrease of 25 percent (because the 1 percentage point drop is now taken against the baseline of 4). Researchers and journalists are not always scrupulous about making this distinction between percentage point and percentages clear, but you should be.

The *New York Times* reported on the closing of a Connecticut textile mill and its move to Virginia due to high employment costs. The *Times* reported that employment costs, "wages, worker's compensation and unemployment insurance—are 20 times higher in Connecticut than in Virginia." Is this plausible? If it were true, you'd think that there would be a mass migration of companies out of Connecticut and into Virginia—not just this one mill—and that you would have heard of it by now. In fact, this was not true and the *Times* had to issue a correction. How did this happen? The reporter simply misread a company report. One cost, unemployment insurance, was in fact twenty times higher in Connecticut than in Virginia, but when factored in with other costs, total employment costs were really only 1.3 times higher in Connecticut, not 20 times higher. The reporter did not have training in business administration and we shouldn't expect her to.

To catch these kinds of errors requires taking a step back and thinking for ourselves—which anyone can do (and she and her editors should have done).

New Jersey adopted legislation that denied additional benefits to mothers who have children while already on welfare. Some legislators believed that women were having babies in New Jersey simply to increase the amount of their monthly welfare checks. Within two months, legislators were declaring the “family cap” law a great success because births had already fallen by 16 percent. According to the *New York Times:*

> After only two months, the state released numbers suggesting that births to welfare mothers had already fallen by 16 percent, and officials began congratulating themselves on their overnight success.

Note that they’re not counting pregnancies, but births. What’s wrong here? Because it takes nine months for a pregnancy to come to term, any effect in the first two months cannot be attributed to the law itself but is probably due to normal fluctuations in the birth rate (birth rates are known to be seasonal).

Even so, there were other problems with this report that can’t be caught with plausibility checks:

> … over time, that 16 percent drop dwindled to about 10 percent as the state belatedly became aware of births that had not been reported earlier. It appeared that many mothers saw no reason to report the new births since their welfare benefits were not being increased.

This is an example of a problem in the way statistics were collected—we’re not actually surveying all the people that we think we are. Some errors in reasoning are sometimes harder to see coming than others, but we get better with practice. To start, let’s look at a basic, often misused tool.

The pie chart is an easy way to visualize percentages—how the different parts of a whole are allocated. You might want to know what percentage of a school district’s budget is spent on things like salaries, instructional materials, and maintenance. Or you might want to know what percentage of the money spent on instructional materials goes toward math, science, language arts, athletics, music, and so on. The cardinal rule of a pie chart is that to be able to represent a whole lot of data with a single number. The median does a better job of this when some of your observations are very, very different from the majority of them, what statisticians call *outliers.*

If we visit a room with nine people, suppose eight of them have a net worth of near $100,000 and one person is on the verge of bankruptcy with a net worth of negative $500,000, owing to his debts. Here's the makeup of the room:

Person 1:	−$500,000
Person 2:	$96,000
Person 3:	$97,000
Person 4:	$99,000
Person 5:	$100,000
Person 6:	$101,000
Person 7:	$101,000
Person 8:	$101,000
Person 9:	$104,000

Now we take the sum and obtain a total of $299,000. Divide by the total number of observations, nine, and the mean is $33,222 per person. But the mean doesn't seem to do a very good job of characterizing the room. It suggests that your fund-raiser might not want to visit these people, when it's really only one odd person, one outlier, bringing down the average. This is the problem with the mean: It is sensitive to outliers.

The median here would be $100,000: Four people make less than that amount, and four people make more. The mode is $101,000, the number that appears more often than the others. Both the median and the mode are more helpful in this particular example.

There are many ways that averages can be used to manipulate what you want others to see in your data.

Let's suppose that you and two friends founded a small start-up company with five employees. It's the end of the year and you want to report your finances to your employees, so that they can feel good about all the long hours and cold pizzas they've eaten, and so that you can attract investors. Let's say that four employees—programmers—each earned $70,000 per year, and one employee—a receptionist/office manager—earned $50,000 per year. That's an average (mean) employee salary of $66,000 per year (4 × $70,000) + (1 × $50,000), divided by 5. You and your two friends each took home $100,000 per year in salary. Your payroll costs were therefore (4 × $70,000) + (1 × $50,000) + (3 × $100,000) = $630,000. Now, let's say your company brought in $210,000 in profits and you divided it equally among you and your co-founders as bonuses, giving you $100,000 + $70,000 each. How are you going to report this?

You could say:

Average salary of employees: $66,000
Average salary + profits of owners: $170,000

This is true but probably doesn't look good to anyone except you and your mom. If your employees get wind of this, they may feel undercompensated. Potential investors may feel that the founders are overcompensated. So instead, you could report this:

Average salary of employees: $66,000
Average salary of owners: $100,000
Profits: $210,000

That looks better to potential investors. And you can just leave out the fact that you divided the profits among the owners, and leave out that last line—that part about the profits—when reporting things to your employees. The four programmers are each going to think they're very highly valued, because they're making more than the average. Your poor receptionist won't be so happy, but she no doubt knew already that the programmers make more than she does.

Now suppose you are feeling overworked and want to persuade your two partners, who don't know much about critical thinking, that you need to hire more employees. You could do what many companies do, and report the "profits per employee" by dividing the $210,000 profit among the five employees:

Average salary of employees: $66,000
Average salary of owners: $100,000
Annual profits per employee: $42,000

Now you can claim that 64 percent of the salaries you pay to employees (42,000/66,000) comes back to you in profits, meaning you end up only having to pay 36 percent of their salaries after all those profits roll in. Of course, there is nothing in these figures to suggest that adding an employee will increase the profits—your profits may not be at all a function of how many employees there are—but for someone who is not thinking critically, this sounds like a compelling reason to hire more employees.

Finally, what if you want to claim that you are an unusually just and fair employer and that the difference between what you take in profits and what your employees earn is actually quite reasonable? Take the $210,000 in profits and distribute $150,000 of it as salary bonuses to you and your partners, saving the other $60,000 to report as "profits." This time, compute the average salary but include you and your partners in it with the salary bonuses.

> Average salary: $97,500
> Average profit of owners: $20,000

Now for some real fun:

> Total salary costs plus bonuses: $840,000
> Salaries: $780,000
> Profits: $60,000

That looks quite reasonable now, doesn't it? Of the $840,000 available for salaries and profits, only $60,000 or 7 percent went into owners' profits. Your employees will think you above reproach —who would begrudge a company owner from taking 7 percent? And it's actually not even that high—the 7 percent is divided among the three company owners to 2.3 percent each. Hardly worth complaining about!

You can do even better than this. Suppose in your first year of operation, you had only part-time employees, earning $40,000 per year. By year two, you had only full-time employees, earning the $66,000 mentioned above. You can honestly claim that average employee earnings went up 65 percent. What a great employer you are! But here you are glossing over the fact that you are comparing part-time with full-time. You would not be the first: U.S. Steel did it back in the 1940s.

In criminal trials, the way the information is presented—the framing—profoundly affects jurors' conclusions about guilt. Although they are mathematically equivalent, testifying that "the probability the suspect would match the blood drops if he were not their source is only 0.1 percent" (one in a thousand) turns out to be far more persuasive than saying "one in a thousand people in Houston would also match the blood drops."

Averages are often used to express outcomes, such as "one in *X* marriages ends in divorce." But that doesn't mean that statistic will apply on your street, in your bridge club, or to anyone you know. It might or might not—it's a nationwide average, and there might be certain *vulnerability factors* that help to predict who will and who will not divorce.

Similarly, you may read that one out of every five children born is Chinese. You note that the Swedish family down the street already has four children and the mother is expecting another child. This does not mean she's about to give birth to a Chinese baby—the one out of five children is on average, across all births in the world, not the births restricted to a particular house or particular neighborhood or even particular country.

Be careful of averages and how they're applied. One way that they can fool you is if the average combines samples from disparate populations. This can lead to absurd observations such as:

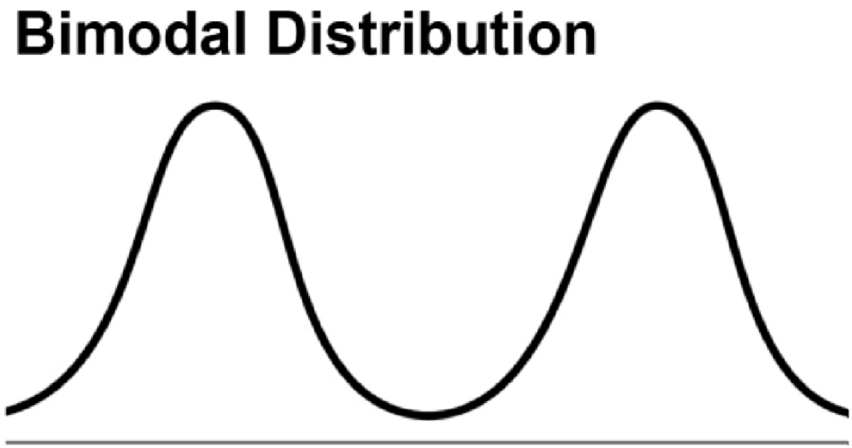
> On average, humans have one testicle.

This example illustrates the difference between mean, median, and mode. Because there are slightly more women than men in the world, the median and mode are both zero, while the mean is close to one (perhaps 0.98 or so).

Also be careful to remember that the average doesn't tell you anything about the range. The average annual temperature in Death Valley, California, is a comfortable 77 degrees F (25 degrees C). But the range can kill you, with temperatures ranging from 15 degrees to 134 degrees on record.

Or … I could tell you that the *average* wealth of a hundred people in a room is a whopping $350 million. You might think this is the place to unleash a hundred of your best salespeople. But the room could have Mark Zuckerberg (net worth $35 billion) and ninety-nine people who are indigent. The average can smear across differences that are important.

Another thing to watch out for in averages is the *bimodal distribution*. Remember, the *mode* is the value that occurs most often. In many biological, physical, and social datasets, the distribution has two or more peaks—that is, two or more values that appear more than the others.



For example, a graph like this might show the amount of money spent on lunches in a week (x-axis) and how many people spent that amount (y-axis). Imagine that you've got two different groups of people in your survey, children (left hump—they're buying school lunches) and business executives (right hump—they're going to fancy restaurants). The mean and median here could be a number somewhere right between the two, and would not tell us very much about what's really going on—in fact, the mean and median in many cases are amounts that nobody spends. A graph like this is often a clue that there is heterogeneity in your sample, or that you are comparing apples and oranges. Better here is to report that it's a bimodal distribution and report the two modes. Better yet, subdivide the group into two groups and provide statistics for each.

But be careful drawing conclusions about individuals and groups based on averages. The pitfalls here are so common that they have names: the ecological fallacy and the exception fallacy. The ecological fallacy occurs when we make inferences about an individual based on aggregate data (such as a group mean), and the exception fallacy occurs when we make inferences about a group based on knowledge of a few exceptional individuals.

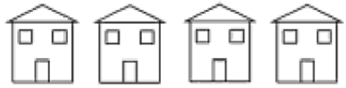
calculate the average (mean) number of siblings, we're sampling children. Each child in the large family gets counted once, so that the number of siblings each of them has weighs heavily on the average for sibling number. In other words, a family with ten children counts only one time in the average *family* statistic, but counts ten times in the average *number of siblings* statistic.

Suppose in one neighborhood of this hypothetical community there are thirty families. Four families have no children, six families have one child, nine families have two children, and eleven families have six children. The average number of children per family is three, because ninety (the total number of children) gets divided by thirty (the total number of families).

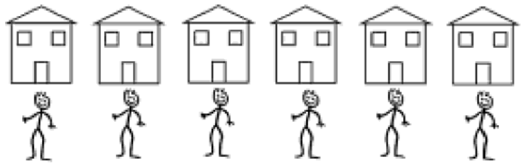
But let's look at the average number of siblings. The mistake people make is thinking that if the average family has three children, then each child must have two siblings on average. But in the one-child families, each of the six children has zero siblings. In the two-child families, each of the eighteen children has one sibling. In the six-child families each of the sixty-six children has five siblings. Among the 90 children, there are 348 siblings. So although the average *child* comes from a family with three children, there are 348 siblings divided among 90 children, or an average of nearly four siblings per child.

	Families	**# Children/ Family**	**Total # Children**	**Siblings**
	4	0	0	0
	6	1	6	0
	9	2	18	18
	11	6	66	330
Totals	30		90	348

Average children per family: 3.0
Average siblings per child: 3.9



4 Families with 0 children



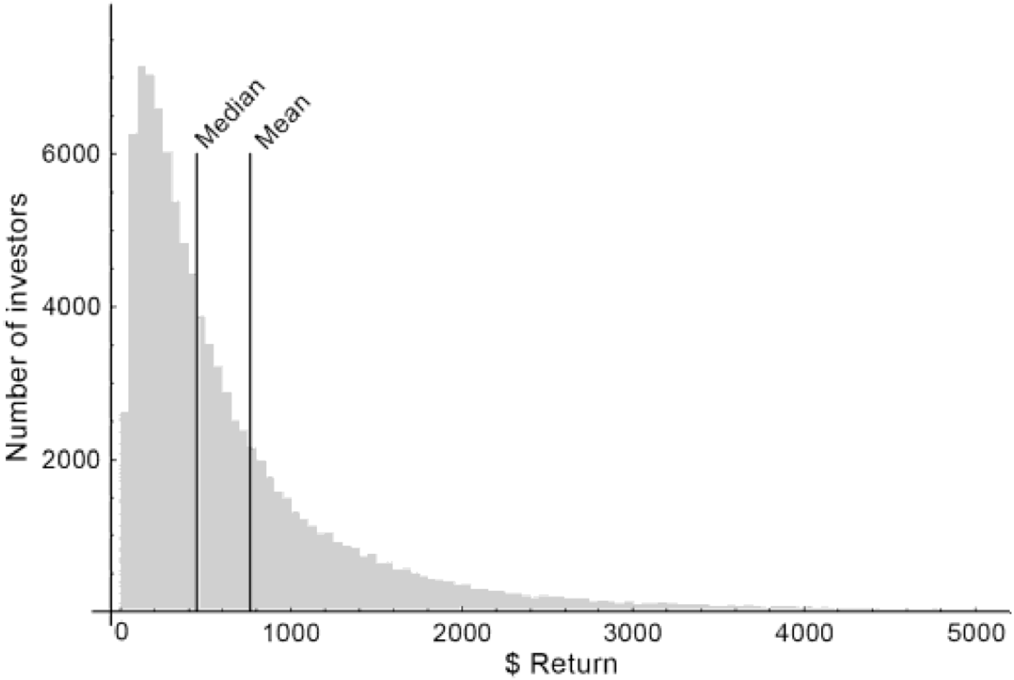
6 Families with 1 child — 6 children with 0 siblings

9 Families with 2 children — 18 children with 1 sibling

11 Families with 6 children — 66 children with 5 siblings

Consider now college size. There are many very large colleges in the United States (such as Ohio State and Arizona State) with student enrollment of more than 50,000. There are also many small colleges, with student enrollment under 3,000 (such as Kenyon College and Williams College). If we count up *schools*, we might find that the average-sized college has 10,000 students. But if we count up students, we'll find that the average student goes to a college with greater than 30,000 students. This is because, when counting students, we'll get many more data points from the large schools. Similarly, the average person doesn't live in the average city, and the average golfer doesn't shoot the average round (the total strokes over eighteen holes).

These examples involve a shift of baseline, or denominator. Consider another involving the kind of skewed distribution we looked at earlier with child mortality: The average investor does not earn the average return. In one study, the average return on a $100 investment held for thirty years was $760, or 7 percent per year. But 9 percent of the investors lost money, and a whopping 69 percent failed to reach the average return. This is because the average was skewed by a few people who made much greater than the average—in the figure below, the *mean* is pulled to the right by those lucky investors who made a fortune.



Payoff outcomes for return on a $100 investment over thirty years. Note that most people make less than the mean return, and a lucky few make more than five times the mean return.

PENGUIN BOOKS

UK | USA | Canada | Ireland | Australia
India | New Zealand | South Africa

Penguin Books is part of the Penguin Random House group of companies whose addresses can be found at global.penguinrandomhouse.com.

Penguin Random House UK

First published as *A Field Guide to Lies* in the United States of America by Dutton 2016
First published in Great Britain by Viking 2016
First published as *Weaponized Lies* in the United States of America by Dutton 2017
Published in Penguin Books 2017

Cover design by Mecob
Photograph: © Shutterstock.com / Mega Pixel

All art courtesy of the author unless otherwise noted.
Images here, here, here, here, and here © 2016 by Dan Piraro, used by permission; image here drawn by Dan Piraro, based on an image by Irving Geis in *How to Lie with Statistics,* © 1954
Image here © 2016 by Alex Tabarrok, used by permission.
Image here was drawn by the author, based on a figure under Creative Commons license appearing on www.betterposters.blogspot.com,
Image here © 2016 by Tyler Vigen, used by permission.
Image here was redrawn by the author with permission, based on a figure found at AutismSpeaks.org.
Image here is public domain and provided courtesy of Harrison Prosper.

ISBN: 978-0-241-31357-2