Five years ago, two teenagers from the Midwest decided they wanted to build an airplane from scratch. Not just a glider, but a twin-engine jet that would carry fifty passengers and fly above five thousand feet. The fact that neither of them knew anything about principles of flight, or engines, or had ever built anything didn’t daunt them—they reasoned that if other people could build jets, they could, too. They sought out books on the subject, but they decided early on that they didn’t want to be constrained by what other people had done before—that their own intuitions should carry just as much weight as a textbook. After all, they reasoned, early airplane designers such as the Wright brothers didn’t have textbooks to rely on, did they, and their airplanes turned out just fine.

They set up a work area in an open field in town. A few weeks into the project, the kids invited members of their high school and the larger community to participate. People could come by any time of the day or night and add to the airplane, or, if they were so inclined, they could take away something that didn’t seem right to them, and either replace it or leave it to the next person to do so. A posted sign instructed passersby that the construction project was open to anyone, regardless of background or ability, that it was a true community project with equal access to all. Why should airplane construction be the domain of an elite few? Anyone with a desire to contribute was encouraged to do so.

At one point, an aeronautical engineer visiting relatives in the town passed by the project and was concerned (horrified might be a better word).
He added an emergency shutoff valve to the fuel system, and installed an oil cooler he found at a nearby junkyard. Before leaving town, he left extensive instructions about wing design and jet engine control, along with warnings and cautions that he said must be heeded before attempting to fly the plane. A few days later, an eleven-year-old county champion of paper airplane making, with a do-it-yourself attitude, brought a wrench to the building site and removed the oil cooler the engineer installed, and then threw away all of the engineer's instructions and warnings. Of course, this was all well within the ethos of the project, so no one stopped him.

The airplane was completed after two years, and a test flight was arranged for ten lucky members of the community. They drew straws to see who the pilot would be, reasoning that this was another job that was overrated and should be given to anyone who was interested.

Would you like to be a passenger in this plane? Of course not! But why not exactly?

For one thing, you might find the blatant disregard for expertise to be off-putting. Most of us believe that building (not to mention flying) an airplane requires special training, that it cannot be left to just anyone. In an organized society, we set up specialized schools to provide aeronautical training. These schools need to be accredited and certified by independent agencies in order to guarantee that the training provided is sound. We endorse in general a system in which licenses and certifications of various kinds need to be issued to surgeons, lawyers, electricians, construction companies, and the like. They assure us that high standards of quality and safety have been met. In short, we accept that there are experts in the world who know more than we do, that this expertise is valuable and indeed necessary to undertake important projects.

This story is entirely fictional, but it is a very close analogue to what Wikipedia does. I say this with some trepidation because Wikipedia has done at least two very admirable things: It has made information massively, unprecedentedly, ridiculously accessible, and it has made it free. I wholeheartedly agree that information should be accessible, and moreover, I believe that this is the foundation of a successful society—well-informed citizens are better able to make decisions about our mutual governance, and better able to become happy and productive members of their community.

But there was a trade-off: an antipathy toward expertise. This is according to no less an authority than Lawrence Sanger, the cofounder (with Jimmy Wales) of Wikipedia! The problem, he notes, is that anyone—anyone—can edit a Wikipedia article, regardless of their knowledge or training. There is no central authority of credentialed experts who review the articles to ensure that they are factual or that they are being edited by someone with knowledge on the topic. As a reader of Wikipedia, you have no way to know whether you're reading something accurate or not. And this isn't an unwitting side effect; it was part of Wikipedia's very design. Jimmy Wales has stated that experts should be accorded no more respect than novices, that there should be "no elite, and no hierarchy" of them to get in the way of newcomers who want to participate in Wikipedia.

If you were looking at the rudder of that community jet plane, you'd have no way to know whether it was designed by an expert or a novice, especially if you yourself were a novice. And when a true expert did come by, the visiting aeronautical engineer, his work and contributions were given no more weight ultimately than those of an eleven-year-old. Moreover, if you were coming upon the airplane for the first time and knew nothing of its history, you might very reasonably assume that it was designed by a professional because that is our expectation when we see such a considerable capital project in this country. We expect bridges not to collapse, car gas tanks not to explode, and dams to hold.

Conventional encyclopedias employ editors who are recognized leaders in their respective fields. The editors in turn identify and hire world-renowned experts in various domains to write the topic entries. Those entries are then reviewed for accuracy and bias by still other world experts in the field, who have to buy into the treatment of the subject. The authors sign their work with their academic credentials so that any reader can see who was responsible for the article and what their qualifications are. The system is not foolproof. At least three sources of inaccuracy emerge in articles: intrinsic bias, maintaining the status quo, and a preselection effect in those who agree to write the articles. An expert on Chinese art may devalue Korean art (intrinsic bias); new ideas and scholarship that challenge well-established ones may take some time to become accepted by the entrenched experts who have risen to a sufficient prominence in their field to be considered encyclopedia authors (a drive to maintain the status quo); scientists with active research programs and who are the most knowledgeable about emerging trends may not take
the time to write encyclopedia articles, which are not considered "important" scholarship by their academic peers (preselection effect).

But although the system is not foolproof, its failures occur within a value system that both acknowledges and respects expertise, a system that both implicitly and explicitly establishes a meritocracy in which those who demonstrably know more about a topic are placed in a position to share their knowledge. I can’t put too fine a point on this. Under the Wikipedia model, a neurosurgeon has as much say about an entry on brain aneurisms as a high school dropout. No rational person would choose the high school dropout as their brain surgeon, but if Wikipedia is to be the de facto and default source of information about technical topics such as aneurisms (and nontechnical ones as well), we should all be able to have confidence in its articles’ origins.

Sure, eventually, someone knowledgeable may come along to correct the high school dropout’s inexpert advice, but when? And how can you know that it’s been done? It might be just before or just after you consult the article. Also, without overseers or curators, the entries have little consistency. Details that have grabbed the attention of a single individual can loom large in an entry, while important things can get much less treatment if no one has the knowledge or interest to fill out those sections. What is lacking is an editorial hand to make a decision on questions such as “Is this fact worth knowing about this entry and is it more important than other facts?” In the extreme, an encyclopedia entry could tell you every possible fact about a person or place, leaving nothing out—but such an entry would be too unwieldy to be useful. The usefulness of most professional summaries is that someone with perspective has used their best judgment about what, in the scheme of things, should be included. The person most involved in editing the entry on Charles Dickens may have no connection to the person writing the entry on Anton Chekhov, and so we end up with idiosyncratic articles that don’t give equivalent weight to their lives, their works, their influences, and their historical place.

For scientific, medical, and technical topics, even in peer-reviewed journals, the information sources aren’t always clearly on display. Technical articles can be difficult to understand without specific training, and there are controversies in many fields that require experience to understand and resolve. An expert knows how to weigh different sources of information and to resolve such apparent contradictions.

Some of the most relentless contributors to Wikipedia revisions appear to be people who have simply read a contrary account in a textbook or were taught something different in high school than what current experts believe. (“If it’s in a textbook, it must be right!”) What many novices don’t know is that it can take five years or more for new information to filter down to textbooks, or that their high school teachers were not always right. As Lawrence Sanger puts it, Wikipedia articles can end up being “degraded in quality by the majority of people, whose knowledge of the subject is based on paragraphs in books and mere mentions in college classes.” Articles that may have started out accurate can be hacked into inaccuracy by hordes of nonexperts, many of whom are inclined to believe fervently that their own intuitions, memories, or feelings ought to hold just as much weight as a scientific paper or the opinions of true experts. The root problem, Lawrence Sanger says, is a “lack of respect for expertise.” As one Wikipedia commentator noted, “Why would an expert bother contributing his valuable time to a project that can be ruined by any random idiot on the net?”

Wikipedia does have two clear advantages over conventional encyclopedias. One is that it is nimble. When there is breaking news—an outbreak of violence in a troubled country, an earthquake, the death of a celebrity—Wikipedia is quick to respond and can report on those events within minutes or hours, unlike print encyclopedias, which take so long to compile. A second advantage is that topics that might not rate inclusion in a print encyclopedia can exist in an online format where space and printed pages are not limiting factors. There are thousands of words written about the computer game Dungeons & Dragons and the TV show Buffy the Vampire Slayer, far more than is written about President Millard Fillmore or Dante’s Inferno. For popular TV shows, Wikipedia features plot summaries for each episode, and includes rather extensive information about guest stars and cast. Entries like these are examples of its strength and where the crowdsourcing ethos can shine. Anyone who is moved by the performance of a particular bit actor on CSI can look for her name in the credits at the end of the show and add her to the Wikipedia entry for that episode, all this without having to be an expert of any kind. Other fans of the show who are engaged with it can be trusted to correct faulty information, relying on the broadcast credits for the program.

This fan-based kind of editing is related to fan fiction, the recent phenomenon of fans writing works featuring their favorite characters from
popular TV shows and films, or adding to existing stories to fill in plot holes or story lines that they feel were insufficient in the original work. This all began with a Star Trek fanzine (of course). This fan-based kind of literature may demonstrate a human need for community storytelling. We are, after all, a social species. We become joined by common stories, whether they are origin stories of humanity or of the country in which we live. Wikipedia clearly answers a need for making storytelling a participatory and communal act, and it inspires millions of people to contribute their enthusiasm and interest (and, yes, often expertise) in the service of what may be one of the most ambitious projects of scholarship ever.

One way to significantly improve Wikipedia would be to hire a panel of editors to supervise the entries and editing process. These experts could ensure uniformity and quality, and moderate disputes. Novices could still contribute—which is part of the fun and excitement of Wikipedia—but an expert panel would have the last word. Such a move would be possible only if Wikipedia had a larger source of income available, either through subscriptions or usage fees, or a benefactor. A loose affiliation of millionaires and billionaires, philanthropists, government agencies, book publishers, and universities could perhaps finance the endeavor, but it is difficult to challenge the grassroots ethos that has grown up around Wikipedia, that its content is democratically determined, and that all information should be free, all the time.

The lack of sympathy for a paid model is similar to a situation that arose in the psychedelic 1960s. When the music impresario Bill Graham started organizing some of the first outdoor rock concerts in Golden Gate Park in San Francisco, many of the hippies complained vigorously about his charging admission to the concerts. “Music should be free,” they cried. Some added that music’s ability to soothe the mortal soul, or its status as the “voice of the universe” virtually mandated that it should be free. Graham patiently pointed out the problem. “OK,” he said, “let’s assume for the moment that the musicians are willing to play for free, that they don’t have to worry about how to pay their rent, or pay for their musical instruments. Do you see that stage? We built it here in the park. It took a team of carpenters and the wood and other materials had to be trucked in. Are they all going to work for free, too? And what about the truck drivers and the gas that their trucks use? Then there are the electricians, sound engineers, lighting, the portable toilets... are all of those people going to work for free, too?”

Of course, as detailed here, this free ethos for Wikipedia leads to all kinds of problems. So for now, the situation is at an impasse, with one notable exception—the emergence of organized and curated editing sessions by public institutions. The Smithsonian American Art Museum in Washington, DC, holds daylong “editathons” to improve the quality of entries by inviting Wikipedia editors, authors, and other volunteers to use the institution’s extensive archives and resources alongside Smithsonian employees. Unfortunately, like the rodents in the Olds and Milner experiment, which repeatedly pressed a bar for a reward, a recalcitrant user can undo all of that curated editing with a single mouse click.

Do the advantages of being able to get a lot of information free outweigh the disadvantages? It depends on how important it is to you that information be accurate. By some definitions of the word, something can be considered “information” only if it is accurate. An important part of information literacy, and keeping information organized, is to know what is true and what isn’t, and to know something about the weight of evidence supporting claims. Although it is important to be respectful of other points of view—after all, this is the way we can learn new things—it is also important to acknowledge that not all points of view are necessarily equally valid: Some really do come from true scholarship and expertise. Someone can believe wholeheartedly that Russia is in the middle of South America, but that doesn’t make it true.

The world has changed for school-age children (not to mention university students and everyone else). Just fifteen years ago, if you wanted to learn a new fact, it took some time. Say, for example, you wanted to know the range of your favorite bird, the scarlet tanager, or the value of Planck’s constant. In the old, pre-Internet days, either you had to find someone who knew or you had to find it yourself in a book. To do the latter, you first had to figure out what book might contain the information. You’d march down to a bricks-and-mortar library and spend an hour—ah, days, has become virtually instantaneous. This utterly changes the role of the classroom teacher from K-12 through graduate school. It no
longer makes sense for teachers to consider their primary function to be
the transmission of information. As the New Yorker essayist Adam Gopnik
put it, nowadays, by the time a professor explains the difference between
elegy and eulogy, everyone in the class has already Googled it.

Of course not everything is so easy to find. The immediate access to
information that Wikipedia, Google, Bing, and other Internet tools pro-
vide has created a new problem that few of us are trained to solve, and this
has to be our collective mission in training the next generation of citizens.
This has to be what we teach our children: how to evaluate the hordes of
information that are out there, to discern what is true and what is not, to
identify biases and half-truths, and to know how to be critical, independent
thinkers. In short, the primary mission of teachers must shift from the
dissemination of raw information to training a cluster of mental skills that
revolve around critical thinking. And one of the first and most important
lessons that should accompany this shift is an understanding that there
exist in the world experts in many domains who know more than we do.
They should not be trusted blindly, but their knowledge and opinions, if
they pass certain tests of face validity and bias, should be held in higher
regard than those who lack special training. The need for education and the
development of expertise has never been greater. One of the things that
experts spend a great deal of their time doing is figuring out which sources
of information are credible and which are not, and figuring out what they
know versus what they don’t know. And these two skills are perhaps the
most important things we can teach our children in this post-Wikipedia,
post-Google world. What else? To be conscientious and agreeable. To be
tolerant of others. To help those less fortunate than they. To take naps.

As soon as a child is old enough to understand sorting and organizing,
it will enhance his cognitive skills and his capacity for learning if we teach
him to organize his own world. This can be stuffed animals, clothes, pots
and pans in the kitchen. Make it into a game to sort and re-sort, by color,
by height, by shininess, by name—all as an exercise in seeing the attributes
one by one. Recall that being organized and conscientious are predictive of
a number of positive outcomes, even decades later, such as longevity, over-
all health, and job performance. Being organized is a far more important
trait than ever before.

Procrastination is a pervasive problem and is more widespread among
children than adults. Every parent knows the difficulties encountered in

trying to get a child to do homework when a favorite TV show is on, to
clean her room when friends are outside playing, or even just to go to bed
at the designated bedtime. These difficulties arise for two reasons—children
are more likely to want immediate gratification, and they are less likely to
be able to foresee the future consequences of present inaction; both of these
are tied to their underdeveloped prefrontal cortices, which don’t fully
mature until after the age of twenty (!). This also makes them more vulner-
able to addiction.

To some extent, most children can be taught to do it now and to shun
procrastination. Some parents even make a game out of it. Recall the motto
of Jake Eberts, the film producer who taught his children: “Eat the frog. Do
that unpleasant thing first thing in the morning and you feel free for the
rest of the day.”

There are a number of critical thinking skills that are important, and
teaching them is relatively straightforward. Indeed, most of them are already
taught in law schools and graduate schools, and in previous generations
they were taught in college preparatory–oriented grades 6–12 schools. The
most important of these skills are not beyond the reach of the average
twelve-year-old. If you like watching courtroom dramas (Perry Mason, L.A.
Law, The Practice), many of the skills will be familiar, for they closely resemble
the kinds of evaluations that are made during court cases. There, a judge,
jury, and the attorneys for both sides must decide what to admit into the
court, and this is based on considerations such as the source of the informa-
tion, its credibility, whether or not a witness possesses the necessary
expertise to make certain judgments, and the plausibility of an argument.

My colleague Stephen Kosslyn, a cognitive neuroscientist who was for-
merly chair of Harvard’s Psychology Department and is now dean of the
faculty at the Minerva Schools at KGI, calls these collectively foundational
concepts and habits of mind. They are mental habits and reflexes that should
be taught to all children and reinforced throughout their high school and
college years.

Information Literacy

There is no central authority that controls how websites or blogs are named,
and it is easy to create a fictitious identity or claim phony credentials. The
president of Whole Foods posed as a regular guy customer to praise the
what to teach our children

store’s pricing and policies. There are many stories like this. Just because a website is named U.S. Government Health Service doesn’t mean it is run by the government; a site named Independent Laboratories doesn’t mean that it is independent—it could well be operated by an automobile manufacturer who wants to make its cars look good in not-so-independent tests.

Newspapers and magazines such as The New York Times, The Washington Post, The Wall Street Journal, and Time strive to be neutral in their coverage of news. Their reporters are trained to obtain information with independent verifications—a cornerstone of this kind of journalism. If one government official tells them something, they get corroboration from another source. If a scientist makes a claim, reporters contact other scientists who don’t have any personal or professional relationship with the first scientist, in order to get independent opinions. Few would take at face value a claim about the health benefits of almonds published by the Almond Growers Association of the United States but nowhere else.

It is true that reputable sources are somewhat conservative about wanting to be certain of facts before running with them. Many sources have emerged on the Web that do not hold to the same traditional standards of truth, and in some cases, they can break news stories and do so accurately before the more traditional and conservative media. TMZ ran the story of Michael Jackson’s death before anyone else because they were willing to publish the story based on less evidence than CNN or The New York Times. In that particular case, they turned out to be right, but it doesn’t always work out that way.

During fast-breaking news events, like the Arab Spring, journalists aren’t always on site. Reports from regular citizens hit the Web through Twitter, Facebook, and blogs. These can be reliable sources of information, especially when considered as a collection of observations. Nonprofessional journalists—citizens who are swept up in a crisis—provide timely, firsthand accounts of events. But they don’t always distinguish in their reports what they’ve perceived firsthand from what they’ve simply heard through rumor or innuendo. Our hunger for instant updates on breaking news stories leads to inaccuracies that become settled only later. Early reports contain false or unverified information that doesn’t get sorted out until some hours or days after the event. In the pre-Internet days, journalists had time to gather the necessary information and verify it before going to press. Because newspapers published only once a day, and network TV

news had its primary broadcast only once a day, there was not the rush we see now to run with a story before all the facts are in.

During the August 2013 chemical attacks in Syria, the information flow available via social media was contaminated with misinformation, some of it deliberately planted. With no trained investigative journalists to organize the conflicting and contradictory accounts, it was difficult for anyone to make sense of what was happening. As former New York Times editor Bill Keller noted, “It took an experienced reporter familiar with Syria’s civil war, my colleague C. J. Chivers, to dig into the technical information in the U.N. report and spot the evidence—compass bearings for two chemical rockets—that established the attack was launched from a Damascus redoubt of Assad’s military.” Chivers himself: “Social media isn’t journalism, it’s information. Journalism is what you do with it.”

Two sources of bias can affect articles. One is the bias of the writer or editors. As humans, they have their own political and social opinions, and for serious journalism, these are supposed to be left at the door. This isn’t always easy. One difficulty with preparing a neutral news story is that there can be many subtleties and nuances, many parts of the story that don’t fit neatly into a brief summary. The choice about what parts of an article to leave out—elements that complicate the story—is just as important as deciding what to include, and the conscious or subconscious biases of the writers and editors can come into play in this selection.

Some news sources, such as National Review or Fox (on the right) and MSNBC or The Nation (on the left) appeal to us because they have a particular political leaning. Whether or not this is the result of a conscious filtering of information isn’t obvious. Some of their reporters may feel they are the only neutral and unbiased journalists in the business. Others may feel it is their responsibility to seek out views on their side of the political spectrum to counter what they perceive as a pernicious political bias in the so-called mainstream media.

My former professor Lee Ross of Stanford University conducted a study that revealed an interesting fact about such politically and ideologically based biases in news reporting, dubbed the hostile media effect. Ross and his colleagues, Mark Lepper and Robert Vallone, found that partisans on any side of an issue tend to find reporting to be biased in favor of their opponents. In their experiment, they showed a series of news reports about the 1982 Beirut massacre to Stanford students who had pre-identified
themselves as either pro-Israeli or pro-Palestinian. The pro-Israeli students complained that the reporting was strongly biased toward the Palestinian point of view. They said that the news reports held Israel to stricter standards than other countries, and that the reporters were clearly biased against Israel. Finally, the students counted only a few pro-Israeli references in the reports, but many anti-Israeli references. The pro-Palestinian students, on the other hand, reported exactly the opposite bias from watching the same news reports—they judged the reports to be strongly biased toward Israel, and counted far fewer pro-Palestinian references and far more anti-Palestinian references. They, too, felt that the reporters were biased, but against the Palestinians, not the Israelis. Both groups were worried that the reports were so biased that previously neutral viewers would turn against their side after viewing them. In fact, a neutral group of students watching the same clips held opinions that fell between the opinions of the partisan students, testifying to the neutrality of the clips.

These experiments were conducted with reports that were close to being objectively neutral (as indexed by the neutral students’ responses). It is easy to imagine, then, that a partisan watching a news report that is skewed toward his or her beliefs will find that to be neutral. This is arguably a factor in the new prominence of so-called ideologically driven news commentary such as those by Ann Coulter and Rachel Maddow, a form of journalism that has probably always existed as long as there was news to convey. Herodotus, in ancient Greece, is not only recognized as among the first historians, but the first who allowed partisan bias to enter into his reports, and he was taken to task by Aristotle, Cicero, Josephus, and Plutarch for it. Biases come in many forms, including what is deemed newsworthy, the sources cited, and the use of selective rather than comprehensive information.

We are not always seeking neutrality when finding information on the Web, but it is important to understand who is providing the information, what organizations they are sponsored by or affiliated with (if any), and whether the website content is sanctioned by or provided by officials, experts, partisans, amateurs, or people posing as someone they’re not.

Because the Internet is like the Wild West—largely lawless and self-governed—it is the responsibility of each Internet user to be on guard against being taken in by the digital equivalent of crooks, con artists, and snake-oil salesmen. If this is beginning to sound like still another instance of shadow work, it is. The work of authenticating information used to be done, to varying degrees, by librarians, editors, and publishers. In many universities, a librarian holds an advanced degree and a rank equivalent to that of a professor. A good librarian is a scholar’s scholar, familiar with the difference between a rigorously reviewed journal and a vanity press, and is up to date on controversies in many different fields that arise due to lapses in scholarship, credibility, and where to look for impartial perspectives.

Librarians and other information specialists have developed user’s guides to evaluating websites. These include questions we should ask, such as “Is the page current?” or “What is the domain?” (A guide prepared by NASA is particularly helpful.) Critical thinking requires that we not take at face value the content we find on the Web. The usual cues we evolved to use when interacting with people—their body language, facial expressions, and overall demeanor—are absent. People re-post articles and alter them for their own gain; advertising endorsements are disguised as reviews; and impostors are difficult to detect. Is the page merely an opinion? Is there any reason you should believe its content more than that of any other page? Is the page a rant, an extreme view, possibly distorted or exaggerated?

In evaluating scientific and medical information, the report should include footnotes or other citations to peer-reviewed academic literature. Facts should be documented through citations to respected sources. Ten years ago, it was relatively easy to know whether a journal was reputable, but the lines have become blurred with the proliferation of open-access journals that will print anything for a fee in a parallel world of pseudo-academia. As Steven Goodman, a dean and professor at Stanford School of Medicine, notes, “Most people don’t know the journal universe. They will not know from a journal’s title if it is for real or not.” How do you know if you’re dealing with a reputable journal or not? Journals that appear on indexes such as PubMed (maintained by the U.S. National Library of Medicine) are selected for their quality; articles on Google Scholar are not. Jeffrey Beall, a research librarian at the University of Colorado Denver, has developed a blacklist of what he calls predatory open-access journals. His list has grown from twenty publishers four years ago to more than three hundred today.

Suppose your doctor recommends that you take a new drug and you’re trying to find more information about it. You enter the name of the drug into your favorite search engine, and one of the first sites that comes up is
RxList.com. You haven’t seen this site before and you want to validate it. From the “About RxList” page, you learn, “Founded by pharmacists in 1995, RxList is the premier Internet Drug Index resource.” A link sends you to a list of contributing writers and editors, with further links to brief bios, showing their academic degrees or professional affiliations so that you can decide for yourself if their expertise is suitable. You can also enter RxList.com into Alexa.com, a free data mining and analysis service, where you learn that the site is mostly used by people with only “some college” and, compared to other Internet sites, it is less used by people with college or graduate degrees. This tells you that it is a resource for the typical layperson, which might be just what you’re looking for—a way to avoid the technical jargon in medical descriptions of pharmaceutical products—but for more sophisticated users, it serves as a warning that the information may not be vetted. How reliable is the information? According to Alexa, the top five sites that link to RxList.com are:

- yahoo.com
- wikipedia.org
- blogger.com
- reddit.com
- bbc.co.uk

Only one of these informs us about the site’s validity, the link from the BBC news service. If you follow the link, however, it turns out to be on a message board portion of the site, and is nothing more than a comment by a reader. A Google search of .gov sites that link to RxList.com is more helpful, revealing 3,290 results. Of course the number itself is meaningless—they could be subpoenas or court proceedings of actions against the company, but a random sampling shows they are not. Among the first reported links is from the National Institutes of Health on a page of recommended resources for clinical medicine, as well as links from the State of New York, the State of Alabama, the U.S. Food and Drug Administration, the National Cancer Institute (of the NIH), and other organizations that lend legitimacy and a seal of approval to RxList.com.

Because the Web is unregulated, the burden is on each user to apply critical thinking when using it. You can remember the three aspects of Web literacy by an English word used to express good wishes on meeting or parting: ave. Authenticate, validate, and evaluate.

Much of the information we encounter about health, the economy, our favorite sports, and new product reviews involves statistics, even if it isn’t dressed up that way. One source of misleading data comes from a bias in the way the data were obtained. This most often occurs in statistical summaries that we encounter, but it can also occur in regular news stories. The bias refers to cases when an unrepresentative sample is collected (of people, germs, foods, incomes, or whatever quantity is being measured and reported on). Suppose a reporter wanted to measure the average height of people in the city of Minneapolis, for an investigative story on whether alleged contaminants in the water supply have led to a decrease in height of the population. The reporter decides to stand on a street corner and measure passersby. If the reporter stands in front of a basketball court, the sample is likely to be taller than the average person; if the reporter stands in front of the Minneapolis Society for Short People, the sample is likely to be shorter than the average.

Don’t laugh—this type of sampling error is pervasive (though admittedly not always as obvious) even in respected scientific journals! The types of people who volunteer for experimental drug trials are undoubtedly different from those who don’t; they may be from lower on the socioeconomic spectrum and need the money, and socioeconomic status is known to correlate with a range of overall health measures due to differences in childhood nutrition and access to regular health care. When only a certain subset of all possible experimental participants walk through the laboratory door, the sampling bias is called a preselection effect. As another example, if the researchers advertise for participants for a new drug experiment, and a precondition is that they cannot drink any alcohol for eight weeks during the trials, the researchers end up bypassing the average person preselecting people who have a certain lifestyle and all that goes with it (they may be super stressed from not having the relief of an occasional drink; they may be alcoholics in a treatment program; they may be unusually healthy people who are exercise fanatics).

Harvard University routinely publishes data on the salaries earned by recent graduates. The type of mental training that we should be teaching, starting in childhood, would cause anyone to ask the question: Is there a
possible source of bias in Harvard’s data? Could those salary figures some-
how be inaccurate, due to a latent bias in the data-gathering methods? For
example, if Harvard relied on surveys mailed to recent graduates, it might
miss any recent graduates who are homeless, indigent, or in prison. Of
those who did receive the survey, not all may be willing to return it. It
seems plausible that those recent Harvard graduates who are unemployed
or working at menial jobs or simply not making much money might be too
embarrassed to return the survey. This would lead to an overestimate of the
true average salary of recent graduates. And of course there is another
source of error—shocker—people lie (even Harvard students). In a survey
like this, recent grads may overstate their income to impress whoever might
be reading the survey, or out of guilt that they’re not making more.

Imagine that a stockbroker sends an unsolicited letter to your home.

Dear neighbor,

I’ve just moved into the area and I’m an expert at predicting the
stock market. I’ve already made a fortune, and I would like you to be
able to benefit from the system that I’ve developed over many years
of hard work.

I’m not asking for any of your money! I’m just asking that you
give me a chance to prove myself without any commitment from
you at all. Over the next several months, I will make predictions
about the market through the mail, and all you have to do is wait to
see whether they are true or not. At any time, you can ask me to stop
sending letters. But if my predictions are right, you can contact me
at the number below and I’ll be happy to take you on as a client and
help you to make riches beyond your wildest dreams.

To get started, I predict that IBM stock will go up in the next
month. I’ll send you another letter in four weeks with my next pre-
diction.

A month later, you get another letter.

Dear neighbor,

Thanks for opening this letter. As you’ll recall, I predicted last
month that IBM stock would go up—and it did! My next prediction
is that Dow Chemical will go up. Talk to you next month.

A month after that, you get another letter, with the broker pointing out
that he was correct again, and making a new prediction. This goes on for
six months in a row—every time, he called the market exactly as he said he
would. At this point, the average person would be thinking that they should
give him some money. Some might even be thinking of mortgaging their
houses and giving him all their money. Six times in a row! This guy is a
genius! You know from Chapter 6 that the odds of him getting it right by
chance is only 1/2^6 or 1 in 64.

But you’re not the average person. You’ve been trained in habits of
mind and to ask the question, Is there missing information? Is there a
logical, alternative explanation for this broker’s success that doesn’t rely on
his having unheard-of abilities in predicting the market? What informa-
tion might be missing, or obscured from your view?

Consider that in this particular case, you’re seeing only those letters he
chose to send you—you’re not seeing the letters he sent to anyone else. Statis-
ticians call this selective windowing. The case I’m describing actually
occurred, and the broker was imprisoned for fraud. At the beginning of the
whole scheme, he sent out two sets of letters: One thousand people received
a letter predicting that IBM stock would go up, and one thousand received
a letter predicting that it would go down. At the end of the month, he simply
waits to see what happens. If IBM went down, he forgets about the thou-
sand people who received a wrong prediction, and he sends follow-up let-
ters only to those thousand who received the correct prediction. He tells
half of them that Dow Chemical will go up and he tells the other half that
Dow Chemical will go down. After six iterations of this, he’s got a core
group of thirty-one people who have received six correct predictions in a
row, and who are ready to follow him anywhere.

Selective windowing occurs in less nefarious and deliberate ways as
well. A video camera trained on a basketball player making ten free throws
in a row may selectively window the successes, and the one hundred misses
surrounding it aren’t shown. A video of a cat playing a recognizable melody
on the piano may show only ten seconds of random music out of several
hours of nonsense.

We often hear reports about an intervention—a pill someone takes to
improve their health, a government program that defuses tensions in a for-
egnern country, an economic stimulus package that puts many people back to
work. What’s usually missing in these reports is a control condition, that
is, what would have happened without the intervention? This is especially important if we want to draw conclusions about causality, that the one event caused the other. We can't know this without a proper control. "I took Vitamin C and my cold was gone in four days!" But how long would it have taken for your cold to go away if you didn't take Vitamin C? If the peculiar flight and maneuvering pattern a witness attributed to a UFO could be replicated in a conventional aircraft, it takes some of the juice out of the argument that the aircraft could only have been a UFO.

For decades, the professional magician and skeptic James Randi has followed self-proclaimed psychics around the world and has replicated exactly their feats of mind reading. His aim? To counter the argument that the psychics must be using extrasensory perception and mysterious psychic powers because there is no other explanation for their extraordinary feats. By accomplishing the same thing with magic, Randi is offering a more logical, parsimonious explanation for the phenomena. He is not proving that psychic powers don't exist, only that psychics can't do anything that he can't do using ordinary magic. His is a control condition, one in which psychic powers are not used. This leaves the following logical possibilities:

1. Both psychic powers and magic exist, and they are able to produce the same feats.
2. Psychic powers don't exist—psychics use magic and then lie about it.
3. Magic doesn't exist—magicians use psychic powers and then lie about it.

Two of these options require you to toss out everything that is known about science, cause and effect, and the way the world works. One of these asks you only to believe that some people in the world are willing to lie about what they do or how they do it in order to make a living. To make matters more interesting, Randi has offered a $1 million prize to anyone who can do something with psychic powers that he cannot replicate with magic. The only restriction is that the psychics have to perform their demonstrations under controlled conditions—using cards or other objects that are neutral (not ones that they provide or have a chance to doctor)—and with video cameras taping them. More than four hundred people have tried to claim the prize, but their psychic powers mysteriously fail them under these conditions, and the money sits in an escrow account. As Stanford psychologist Lee Ross says, "If psychic powers exist, they are impish, and do not want to be discovered in the presence of a scientist."

When two quantities vary together in some kind of clear relationship, we say they are correlated. Taking multivitamins was correlated, in some older studies, with longer life span. But this doesn't mean that multivitamins cause you to live longer. They could be entirely unrelated, or there could be a third factor X that causes both. It's called X because it is, at least initially, unidentified. There could be a cluster of behaviors called health conscientiousness. People who are health conscientious see their doctors regularly, eat well, and exercise. This third factor X might cause those people to take vitamins and to live longer; the vitamins themselves may be an artifact in the story that does not cause longer life. (As it happens, the evidence that multivitamins are associated with longer life appears to be faulty, as mentioned in Chapter 6.)

The results of Harvard's salary survey are no doubt intended to lead the average person to infer that a Harvard education is responsible for the high salaries of recent graduates. This may be the case, but it's also possible that the kinds of people who go to Harvard in the first place come from wealthy and supportive families and therefore might have been likely to obtain higher-paying jobs regardless of where they went to college. Childhood socioeconomic status has been shown to be a major quantity correlated with adult salaries. Correlation is not causation. Proving causation requires carefully controlled scientific experiments.

Then there are truly spurious correlations—odd pairings of facts that have no relationship to each other and no third factor X linking them. For example, we could plot the relationship between the global average temperature over the past four hundred years and the number of pirates in the world and conclude that the drop in the number of pirates is caused by global warming.

The Gricean maxim of relevance implies that no one would construct such a graph (below) unless they felt these two were related, but this is where critical thinking comes in. The graph shows that they are correlated, but not that one causes the other. You could spin an ad hoc theory—pirates can't stand heat, and so, as the oceans became warmer, they sought other employment. Examples such as this demonstrate the folly of failing to separate correlation from causation.
This type of experiment is performed routinely with experimental medicines, and people are naturally quite willing to volunteer to be in the drug condition if they think it will help cure their illness. But if the experiment carries with it the threat of harm, such as smoking, it is unethical to perform it. The logic behind random assignment is this: There might be some people who are more likely to show an effect in your experiment than others, and random assignment helps to distribute them evenly across the different experimental groups. We know that people who work in coal mines or who live in areas with serious air pollution are more likely to contract lung cancer; it wouldn't be a fair test of smoking if all of them ended up in the smoking group or in the no smoking group—the researcher randomly assigns subjects to even out the potential effects of any preexisting condition, personality type, or other factor that could bias the results.

It can be all too tempting to infer causation from correlational data, especially when controlled experiments can't be done. When we can imagine a plausible, underlying mechanism, the temptation is even greater. The data linking smoking and lung cancer in humans are correlational. The data were obtained by looking retrospectively at people who die from lung cancer and tracking back to see whether or not they were smokers and how much they smoked. The correlation is no perfect: Not every smoker dies of lung cancer and not everyone who dies of lung cancer was a smoker. Some smokers live long lives and die of other things—there are many people who continue to smoke into their eighties and nineties. Some lung cancers appear in nonsmokers, and could be based on genetic or epigenetic factors, exposure to radiation, or other factors. But the correlation between smoking and lung cancer is strong—90% of lung cancers occur among smokers—and scientists have identified a plausible underlying mechanism: toxic chemicals within the smoke-damaged lung tissue. No one has proven that smoking causes lung cancer with a controlled experiment, but we infer causation. It's important to know the difference.

Consider an alternative account, championed by the scientist (and smoker) Hans Eysenck. He has proposed that there is a certain personality type that is prone to smoking. This seems reasonable. Now suppose that there is a gene associated with this personality type and also associated with a propensity to get lung cancer. The gene becomes a third factor—it increases the likelihood that people will take up smoking and it also increases
the likelihood that they'll get lung cancer. Note that if this is true, those people would have gotten lung cancer whether or not they smoked—but because the gene causes them to smoke, we'll never know for sure if they would have contracted lung cancer without smoking. Few scientists take this view seriously, but it is possible that Eysenck is right.

An example of a third-factor-x explanation that turned out to be true is the story of high-tension power lines and childhood leukemia rates in suburban Denver. In the 1980s, the Centers for Disease Control and Prevention became alarmed that the incidence of childhood leukemia was several times greater in certain Denver suburbs than in other parts of the country. An investigation was launched. Researchers discovered that the neighborhoods that had the highest leukemia rates were those with high-tension power lines. And the closer a house was to power lines with a transformer, the greater the risk of leukemia. It was suggested that the strong electromagnetic fields of the power lines disrupted the cellular membrane in children, making the cells more vulnerable to mutation and hence cancer. Here was a correlation and a plausible underlying mechanism. The mechanism was that the power lines were causing the leukemia. A several-years-long epidemiological investigation, however, concluded there was a third factor x that accounted for most of the increase in leukemia rates—socioeconomic status. Because the lines are unsightly, and because much of suburban Denver has power lines buried underground, the houses near visible power lines were less expensive. People who lived near these power lines were thus more likely to be lower on the socioeconomic scale; they had poorer diets, poorer access to health care, and, on average, unhealthy lifestyles. The correlation between living near power lines and developing leukemia was real, but the initial explanation of the cause was not accurate—socioeconomic status was driving both.

Fish oil, rich in omega-3 fatty acids, has been found to be protective against cardiovascular disease, and the American Heart Association has been recommending the consumption of fish twice a week and the supplementation of fish oil capsules, for over ten years. Long-chain omega-3 fatty acids, found in oily fish such as herring, sardines, salmon, and mackerel, are considered essential for human health. They decrease inflammation and have been associated with improved mood, cognitive ability, and energy, and they strengthen heart function. Although some recent studies have cast doubt on the efficacy of fish oil, there is still a body of evidence pointing to its benefits, so many doctors continue to recommend it to their patients.

In the summer of 2013, a study came out that found a strong link between omega-3s and increased risk of prostate cancer in men. Men who were already diagnosed with prostate cancer were found to have higher levels of chemicals found in fatty fish than did cancer-free men. Those blood levels were associated with a 43% increase in the risk of developing prostate cancer. Of course, in a correlational study, there could be some third factor x that is leading to both, and this has not been identified (nor much mentioned in the articles reporting on this finding). Doctors are split as to whether to continue recommending fish oil for their male patients.

The situation is confusing, to say the least. One of the most vocal critics of the study, Dr. Mark Hyman, has a potential conflict of interest: He runs a laboratory that makes its money testing people's blood for omega-3 fatty acids, and operates a website that sells omega-3 capsules. But that doesn't mean he's wrong. He notes that the data are correlational, not the result of a controlled study. He voices concern about the way the blood samples were analyzed. There are many unknowns here, and the various risks and benefits are not well quantified, making the kinds of fourfold table analysis of Chapter 6 unreliable, although not undoable. Thus, a fairly solid body of evidence does suggest that fish oil protects against heart disease, and this is now pitted against a single new study suggesting that fish oil may promote prostate cancer.

To find out how physicians were handling the news, I talked to cardiologists, urological oncologists, and internists. The cardiologists and oncologists were generally split along party lines, with the cardiologists recommending fish oil supplements for the protective effects, and the oncologists recommending against them for the increased cancer risk. An extreme reading of the state of the decision would be: “Die of heart disease or prostate cancer? You get to choose!” Dr. Katsuto Shinohara, a urological oncologist at UC San Francisco, broke the tie among the various physicians, noting that “it is prudent not to put too much stock in a single study.” That single study will surely be followed up on and replicated in the coming years. The evidence for the protective effects of fish oil across dozens of studies, he feels, outweighs the risks of fish oil shown in a single study.

Still, men with previously diagnosed prostate cancer may wish to be especially careful. For them (and perhaps men over fifty who have not
received such a diagnosis), the situation presents no clear solution. If one waits until the new fish oil studies come in, there are risks in the meantime of taking omega-3s and of no: taking them. Incidentally, the American Heart Association also recommends eating tofu and soybeans for their beneficial cardiological effects, and some studies show that soy is preventive against prostate cancer. Other studies have shown that soy doesn't reduce prostate cancer recurrence, and that it can be associated with loss of mental acuity in older men.

The fish oil question is perhaps one for the decision-making equivalent of the junk drawer, a decision that cannot easily be categorized with what we know at present. Sometimes, critical thinking leads to a conclusion that there is no certain answer. And yet we then must make a certain choice.

Whenever we encounter information in the form of numbers, it is important to do a quick mental check to see if the numbers reported are even plausible. To do this, you need to be familiar with certain bits of world knowledge. Every one of us has a mental file drawer full of trivia, such as the population of the United States, the normal operating speed of a car, how long it takes to lose weight, or the amount of time a normal human pregnancy lasts. And any fact of this sort that you haven't committed to memory is just a few hundred milliseconds from finding through a Web search. Performing a quick check of the plausibility of numerical information is one of the easiest and most important parts of critical thinking.

If someone says that 400 million people voted in the last U.S. federal election, that a new economy car has a top speed of four hundred miles per hour, or that so-and-so lost fifty pounds in two days with a juice fast, your general knowledge of the world and your inherent numeracy should raise a red flag about these numerical values.

One of the most important skills, then, that we can teach our children is to think about numbers logically and critically, and enable this sort of querying and verification. The goal of these skills is not to see if the number you've encountered is exactly correct, but only to see if it is approximately correct—that is, close enough to be plausible.

There's a quick trick for evaluating numerical information that is rarely written about in the critical thinking literature: Set boundary conditions. A boundary condition describes what the lowest and highest answers could possibly be. Suppose I ask you how tall Shaquille O'Neal is, and you don't know the answer. Is he taller than four feet? Well, certainly he must be, you might reason; he is a famous member of the NBA, and basketball players tend to be tall. Is he taller than five feet? Again, this is almost certainly true. Is he shorter than ten feet? You might catch yourself searching your memory; you've probably never heard of anyone who is ten feet tall, so yes, you would say he is shorter than ten feet. A quick and not-terribly-precise pair of boundary conditions, then, are that Shaq is somewhere between five feet and ten feet tall. If you know a bit about the NBA and the heights of players, and you know something about human physiological limitations, you might be able to refine your boundary conditions to say that he is probably between five feet six inches and seven feet eight inches. The art in setting boundary conditions is to get them as close as possible to each other while maintaining confidence in your answer. According to the NBA, Shaq is seven feet one inch.

Setting boundary conditions is an essential part of scientific and everyday critical thinking, and is crucial to decision-making. We do this all the time without even knowing it. If you go to the grocery store and buy a bag of groceries, and the checkout clerk tells you that the bill is five cents, you instantly know that something is wrong without having to sum up the cost of every item in your bag. Similarly, if she tells you the price is $500, you also know that something is wrong. Effective approximating is being able to set boundary conditions that aren't so ludicrously far apart. Based on your shopping habits, you might know that 90% of the time a typical bag of groceries at your regular market costs between $35 and $45; you'd be surprised if the total was $15 or if it was $75. So we would say that the boundary conditions for your bag of groceries are $35–$45. Scientists would describe this as your 90% confidence interval—that is, you are 90% sure that the register total should fall within this interval. The closer together your boundary conditions are, the more helpful your approximation is of course.

Part of setting boundary conditions is using your knowledge of the world or acquiring a few landmarks to help in your estimation. If you had to estimate a friend's height, you could use the fact that the average door frame in North America is about eighty inches; how tall is the person compared to the door? Alternatively, if you were speaking to her, would you look her in the eye or have to look up or down? If you need to estimate the width of a car or bus, or the size of a room, imagine yourself lying down in
it—would you be able to lie down without curling up? How many of you
would fit in the space?

Scientists talk about order of magnitude estimates. An order of magnitude
is a power of ten. In other words, as a first rough estimate, we try to decide
how many zeros there are in the answer. Suppose I ask you how many table-
spoons of water are in a cup of coffee. Here are some “power of ten” possi-
bilities:

a. 2
b. 20
c. 200
d. 2,000
e. 20,000

For the sake of completeness, we might also include fractional powers
of ten:

f. 1/20
g. 1/200
h. 1/2000

Now you can quickly rule out the fractions: 1/20 of a tablespoon is a
very small amount, and 1/200 is even smaller than that. You could probably
easily rule out 2 tablespoons as the answer. What about 20 tablespoons?
You might not be so sure, and you might find yourself trying to mentally
convert 20 tablespoons to some other, more useful measure like cups or
ounces. Let’s shelve that for a minute to work on the basis of intuition first,
and calculation and conversion second. To recap: You’re sure there are
more than 2 tablespoons; you’re not sure if there are more or less than 20
tablespoons. What about 200 tablespoons? This seems like a lot, but again,
you may not be sure. But it must be clear that 2,000 tablespoons is too
many. Of the eight estimates listed, you’ve quickly converged on two of
them as plausible: 20 and 200 tablespoons. This is actually quite remark-
able. It is a question you may never have considered before, and with a little
bit of reasoning and intuition, you were able to narrow the answer down to
these two possibilities.

Now let’s calculate. If you bake, you may know that there are 2 table-
spoons in 1/8 cup, and therefore 2 × 8 = 16 tablespoons in a cup. The true
answer is not any of those listed above, but the correct answer, 16, is closer
to 20 than to any of the others. The idea of these powers of ten, of the order
of magnitude estimates, is that we don’t let ourselves get hung up on
unnecessary precision when we’re approximating. It is useful enough for
the purposes of this thought experiment to know that the answer is closer to
20 than it is to 2 or to 200. That is what an order of magnitude estimate is.

If you didn’t know how many tablespoons are in a cup, you might pic-
ture a tablespoon and a cup and try to imagine how many times you’d have
to fill your tablespoon and empty it into the cup before the cup would be
full. Not everyone has the former fact at their command, and not everyone
is able to visualize these quantities, and so that process may end for a lot
of people right here. You might simply say that the answer could be 20 or 200
and you’re not sure. You’ve narrowed your answer to two orders of magni-
tude, which is not bad at all.

We set boundary conditions unconsciously many times a day. When
you step on the scale, you expect it to read within a few pounds of what it
read yesterday. When you step outside, you expect the temperature to be
within a certain number of degrees of what it was the last time you were
out. When your teenager tells you that it took forty minutes to get home
from school, you know whether this is within the normal range of times or
not. The point is that you don’t have to count every single item in your gro-
cery bag to know whether or not the total is reasonable; you don’t need to
carry a stopwatch to know if your commute time is radically longer or
shorter than usual. We round, we estimate, we fudge the numbers, and this
is a crucial operation for knowing quickly whether what we observe is rea-
sonable.

Approximately OK

One of the most important tools in critical thinking about numbers is to
grant yourself permission to generate wrong answers to mathematical prob-
lems you encounter. Deliberately wrong answers! Engineers and scientists
do it all the time, so there’s no reason we shouldn’t all be let in on their
little secret: the art of approximating, or the “back of the napkin” calcula-
tion. Such deliberately wrong answers can get you close enough to the right
answer to make a decision in just a fraction of the time. As the British writer Saki wrote, "a little bit of inaccuracy saves a great deal of explanation."

For over a decade, when Google conducted job interviews, they'd ask their applicants questions that have no answers. Google is a company whose very existence depends on innovation—on inventing things that are new and didn't exist before, and on refining existing ideas and technologies to allow consumers to do things they couldn't do before. Contrast this with how most companies conduct job interviews: In the skills portion of the interview, the company wants to know if you can actually do the things that they need doing.

In a restaurant, the necessary skill might be chopping vegetables or making a soup stock. In an accounting firm, it might be an awareness of tax codes and an ability to properly fill out tax forms. But Google doesn't even know what skills they need new employees to have. What they need to know is whether an employee can think his way through a problem. Students who graduate from top universities in technical or quantitative fields such as computer science, electrical engineering, economics, or business know how to apply what they've learned, and know how to look for information they need. But relatively few can effectively think and reason for themselves.

Consider the following question that has been asked at actual Google job interviews: How much does the Empire State Building weigh?

Now, there is no correct answer to this question in any practical sense because no one knows the answer. There are too many variables, too many unknowns, and the problem is unwieldy. Google isn't interested in the answer, though; they're interested in the process, in how a prospective employee would go about solving it. They want to see a reasoned, rational way of approaching the problem to give them insight into how an applicant's mind works, how organized a thinker she is.

There are four common responses to the problem. People throw up their hands and say "That's impossible," or they try to look up the answer somewhere. Although an answer to this is out there on the Web by now (it has become a somewhat famous problem within the computer science community), Google wants to hire employees who can answer questions that haven't been answered before—that requires a certain kind of mind.

prone to methodical thinking. Fortunately, this kind of thinking can be taught, and it is not beyond the reach of the average person. George Polya, in his influential book *How to Solve It*, showed how the average person can solve complicated mathematical problems without specific training in math. The same is true for this class of crazy, unknowable problems.

The third response? Asking for more information. By "weight of the Empire State Building," do you mean with or without furniture? With or without fixtures? Do I count the people in it? But questions like this are a distraction. They don't bring you any closer to solving the problem; they only postpone being able to start it, and soon you are right back where you began, wondering how in the world you would figure out something such as this.

The fourth response is the correct one, using approximating, or what some people call guesstimating. *These* types of problems are also called estimation problems or Fermi problems. After the physicist Enrico Fermi, who was famous for being able to make estimates with little or no actual data, for questions that seemed impossible to answer. Examples of Fermi problems include "How many basketballs will fit into a city bus?" "How many Reese's Peanut Butter Cups would it take to encircle the globe at the equator?" and "How many piano tuners are there in Chicago?" Approximating involves making a series of educated guesses systematically by partitioning the problem into manageable chunks, identifying assumptions, and then using your general knowledge of the world to fill in the blanks.

How would you solve the problem of "How many piano tuners are there in Chicago?" Google wants to know how people make sense of the problem—how they divide up the knowas and unknowns systematically. Remember, you can't simply call the Piano Tuners Union of Chicago and ask; you have to work this from facts (or reasonable guesses) that you can pull out of your head. Breaking down the problem into manageable units is the fun part. Where to begin? As with many Fermi problems, it's often helpful to estimate some intermediate quantity, not the one you're being asked to estimate, but something that will help you get where you want to go. In this case, it might be easier to start with the number of pianos that you think are in Chicago and then figure out how many tuners it would take to keep them in tune.

In any Fermi problem, we first lay out what it is we need to know, then
list some assumptions. To solve this problem, you might start by trying to estimate the following numbers:

1. How often pianos are tuned (How many times per year is a given piano tuned?)
2. How long it takes to tune a piano
3. How many hours a year the average piano tuner works
4. The number of pianos in Chicago

Knowing these will help you arrive at an answer. If you know how often pianos are tuned and how long it takes to tune a piano, you know how many hours a year are spent tuning one piano. Then you multiply that by the number of pianos in Chicago to find out how many hours are spent every year tuning Chicago’s pianos. Divide this by the number of hours each tuner works, and you have the number of tuners.

**Assumption 1:** The average piano owner tunes his piano once a year.

Where did this number come from? I made it up! But that’s what you do when you’re approximating. It’s certainly within an order of magnitude. The average piano owner isn’t tuning only once every ten years, nor ten times a year. Some piano owners tune their pianos four times a year, some of them zero, but one time a year seems like a reasonable guesstimate.

**Assumption 2:** It takes 2 hours to tune a piano. A guess. Maybe it’s only 1 hour, but 2 is within an order of magnitude, so it’s good enough.

**Assumption 3:** How many hours a year does the average piano tuner work? Let’s assume 40 hours a week, and that the tuner takes 2 weeks’ vacation every year: 40 hours a week × 50 weeks is a 2,000-hour work year. Piano tuners travel to their jobs—people don’t bring their pianos in—so the piano tuner may spend 10%–20% of his or her time getting from house to house. Keep this in mind and take it off the estimate at the end.

**Assumption 4:** To estimate the number of pianos in Chicago, you might guess that 1 out of 100 people have a piano—again, a wild guess, but probably within an order of magnitude. In addition, there are schools and other institutions with pianos, many of them with multiple pianos. A music school could have 30 pianos, and then there are old-age homes, bars, and so on. This estimate is trickier to base on facts, but assume that when these are factored in, they roughly equal the number of private pianos, for a total of 2 pianos for every 100 people.

Now to estimate the number of people in Chicago. If you don’t know the answer to this, you might know that it is the third-largest city in the United States after New York (8 million) and Los Angeles (4 million). You might guess 2.5 million, meaning that 25,000 people have pianos. We decided to double this number to account for institutional pianos, so the result is 50,000 pianos.

So, here are the various estimates:

1. There are 2.5 million people in Chicago.
2. 1 out of 100 people have a piano.
3. There is 1 institutional piano for every 100 people.
4. Therefore, there are 2 pianos for every 100 people.
5. There are 50,000 pianos in Chicago.
6. Pianos are tuned once a year.
7. It takes 2 hours to tune a piano.
8. Piano tuners work 2,000 hours a year.
9. In one year, a piano tuner can tune 1,000 pianos (2,000 hours per year + 2 hours per piano).
10. It would take 50 tuners to tune 50,000 pianos (50,000 pianos + 1,000 pianos tuned by each piano tuner).
11. Add 15% to that number to account for travel time, meaning that there are approximately 58 piano tuners in Chicago.

What is the real answer? The Yellow Pages for Chicago lists 83. This includes some duplicates (businesses with more than one phone number are listed twice), and the category includes piano and organ technicians who are not tuners. Deduct 25 for these anomalies, and an estimate of 58 appears to be very close. Even without the deduction, the point is that it is within an order of magnitude (because the answer was neither 6 nor 600).
Back to the Google interview and the Empire State Building question. If you were sitting in that interview chair, your interviewer would ask you to think out loud and walk her through your reasoning. There is an infinity of ways one might solve the problem, but to give you a flavor of how a bright, creative, and systematic thinker might do it, here is one possible “answer.” And remember, the final number is not the point—the thought process, the set of assumptions and deliberations, is the answer.

Let’s see. One way to start would be to estimate its size, and then estimate the weight based on that.

I’ll begin with some assumptions. I’m going to calculate the weight of the building empty—with no human occupants, no furnishings, appliances, or fixtures. I’m going to assume that the building has a square base and straight sides with no taper at the top, just to simplify the calculations.

For size I need to know height, length, and width. I don’t know how tall the Empire State Building is, but I know that it is definitely more than 20 stories tall and probably less than 200 stories. I don’t know how tall one story is, but I know from other office buildings I’ve been in that the ceiling is at least 8 feet inside each floor and that there are typically false ceilings to hide electrical wires, conduits, heating ducts, and so on. I’ll guess that these are probably 2 feet. So I’ll approximate 10–15 feet per story. I’m going to refine my height estimate to say that the building is probably more than 50 stories high. I’ve been in lots of buildings that are 30–35 stories high. My boundary conditions are that it is between 50 and 100 stories; 50 stories work out to being 500–750 feet tall (10–15 feet per story), and 100 stories work out to be 1,000–1,500 feet tall. So my height estimate is between 500 and 1,500 feet. To make the calculations easier, I’ll take the average, 1,000 feet.

Now for its footprint. I don’t know how large its base is, but it isn’t larger than a city block, and I remember learning once that there are typically 10 city blocks to a mile. A mile is 5,280 feet, so a city block is 1/10 of that, or 528 feet. I’m going to guess that the Empire State Building is about half of a city block, or about 265 feet on each side. If the building is square, it is 265 × 265 feet in its length × width. I can’t do that in my head, but I know how to calculate 250 × 250 (that is, 25 × 25 = 625, and I add two zeros to get 62,500). I’ll round this total to 60,000, an easier number to work with moving forward.

Now we’ve got the size. There are several ways to go from here. All rely on the fact that most of the building is empty—that is, it is hollow. The weight of the building is mostly in the walls and floors and ceilings. I imagine that the building is made of steel (for the walls) and some combination of steel and concrete for the floors. I’m not really sure. I know that it is probably not made of wood.

The volume of the building is its footprint times its height. My footprint estimate above was 60,000 square feet. My height estimate was 1,000 feet. So 60,000 × 1,000 = 60,000,000 cubic feet. I’m not accounting for the fact that it tapers as it goes up.

I could estimate the thickness of the walls and floors and estimate how much a cubic foot of the materials weighs and come up then with an estimate of the weight per story. Alternatively, I could set boundary conditions for the volume of the building. That is, I can say that it weighs more than an equivalent volume of solid air and less than an equivalent volume of solid steel (because it is mostly empty). The former seems like a lot of work. The latter isn’t satisfying because it generates numbers that are likely to be very far apart. Here’s a hybrid option: I’ll assume that on any given floor, 95% of the volume is air, and 5% is steel. I’m just pulling this estimate out of the air, really, but it seems reasonable. If the width of a floor is about 265 feet, 5% of 265 = 13 feet. That means that the walls on each side, and any interior supporting walls, total 13 feet. As an order of magnitude estimate, that checks out—the total walls can’t be a mere 1.3 feet (one order of magnitude smaller) and they’re not 130 feet (one order of magnitude larger).

I happen to remember from school that a cubic foot of air weighs 0.08 pounds. I’ll round that up to 0.1. Obviously, the building is not all air, but a lot of it is—virtually the entire interior space—and so this sets the minimum boundary for the weight. The volume times the weight of air gives an estimate of 60,000,000 cubic feet × 0.1 pounds = 6,000,000 pounds.

I don’t know what a cubic foot of steel weighs. But I can estimate that, based on some comparisons. It seems to me that 1 cubic foot of steel must certainly weigh more than a cubic foot of wood. I don’t know what a cubic foot of wood weighs either, but when I stack firewood, I know that an armful weighs about as much as a 50-pound bag of dog food. So I’m going to guess that a cubic foot of wood is about 50 pounds and that steel is about 10 times heavier than that. If the entire Empire State Building were steel, it would weigh 60,000,000 cubic feet × 500 pounds = 30,000,000,000 pounds.
This gives me two boundary conditions: 6 million pounds if the building were all air, and 30 billion pounds if it were solid steel. But as I said, I'm going to assume a mix of 5% steel and 95% air.

\[
\begin{align*}
5\% \times 30 \text{ billion} & = 1,500,000,000 \\
+ 95\% \times 6 \text{ million} & = 5,700,000 \\
\hline
1,505,700,000 \text{ pounds}
\end{align*}
\]

or roughly 1.5 billion pounds. Converting to tons, 1 ton = 2,000 pounds, so 1.5 billion pounds/2,000 = 750,000 tons. Nicely done!

Another job interviewee might approach the problem much more parsimoniously. Using the same assumptions about the size of the building, and assumptions about its being empty, a concise protocol might come down to this:

Skyscrapers are constructed from steel. Imagine that the Empire State Building is filled up with cars. Cars also have a lot of air in them, they're also made of steel, so they could be a good proxy. I know that a car weighs about 2 tons and it is about 15 feet long, 5 feet wide, and 5 feet high. The floors, as estimated above, are about 265 \times 265 feet each. If I stacked the cars side by side on the floor, I could get 265/15 = 18 cars in one row, which I'll round to 20 (one of the beauties of guesstimating). How many rows will fit? Cars are about 5 feet wide, and the building is 265 feet wide, so 265/5 = 53, which I'll round to 50. That's 20 cars \times 50 rows = 1,000 cars on each floor. Each floor is 10 feet high and the cars are 5 feet high, so I can fit 2 cars up to the ceiling. \(2 \times 1,000 = 2,000 \text{ cars per floor. And 2,000 cars per floor} \times 100 \text{ floors} = 200,000 \text{ cars. Add in their weight, 200,000 cars} \times 4,000 \text{ pounds} = 800,000,000 \text{ pounds, or in tons, 400,000 tons.}

These two methods produced estimates that are relatively close—one is a bit less than twice the other—so they help us to perform an important sanity check. The first gave us 750,000 tons, the second gave us about a half million tons. Because this has become a somewhat famous problem (and a frequent Google search), the Empire State Building's website has taken to giving their estimate of the weight, and it comes in at 365,000 tons. So we find that both guesstimates brought us within an order of magnitude of the official estimate, which is just what was required.

Neither of these methods comes up with the weight of the building. Remember, the point is not to come up with a number, but to come up with a line of reasoning, an algorithm for figuring it out. Much of what we teach graduate students in computer science is just this—how to create algorithms for solving problems that have never been solved before. How much capacity do we need to allow for on this trunk of telephone lines going into the city? What will the ridership be of a new subway that is being built? If there is a flood, how much water will spill into the community and how long will it take for the ground to reabsorb it? These are problems that have no known answer, but skilled approximating can deliver an answer that is of great practical use.

The president of a well-known Fortune 500 company proposed the following solution. While not conforming strictly to the rules of the problem, it is nevertheless very clever:

I would find the company or companies that financed the construction of the Empire State Building and ask to see the supplies manifests... the list of every material that was delivered to the construction site. Assuming 10%–15% waste, you could then estimate the weight of the building by the material's that went into it. Actually, even more accurate would be this: Remember that every truck that drives along a freeway has to be weighed because the trucking companies pay a usage tax to the roads department based on their weight. You could check the weight of the trucks and have all the information you need right there. The weight of the building is the weight of all the materials trucked in to construct it.

Is there ever a case where you'd need to know the weight of the Empire State Building? If you wanted to build a subway line that ran underneath it, you would want to know the weight so you could properly support the ceiling of your subway station. If you wanted to add a heavy new antenna to the top of the building, you would need to know the total weight of the
building to calculate whether or not the foundation can support the additional weight. But practical considerations are not the point. In a world of rapidly increasing knowledge, unimaginable amounts of data, and rapid technological advance, the architects of the new technologies are going to need to know how to solve unsolvable problems, how to break them up into smaller parts. The Empire State Building problem is a window into how the mind of a creative and technically oriented person works, and probably does a better job of predicting success in these kinds of jobs than grades in school or scores on an IQ test.

These so-called back-of-the-envelope problems are just one window into assessing creativity. Another test that gets at both creativity and flexible thinking without relying on quantitative skills is the "name as many uses" test. For example, how many uses can you come up with for a broomstick? A lemon? These are skills that can be nurtured beginning at a young age. Most jobs require some degree of creativity and flexible thinking. As an admissions test for flight school for commercial airline pilots, the name-as-many-uses test was used because pilots need to be able to react quickly in an emergency, to be able to think of alternative approaches when systems fail. How would you put out a fire in the cabin if the fire extinguisher doesn't work? How do you control the elevators if the hydraulic system fails? Exercising this part of your brain involves harnessing the power of free association—the brain's daydreaming mode—in the service of problem solving, and you want pilots who can do this in a pinch.

Novelist Diane Ackerman describes playing one session of this game with her husband, Paul, in her book One Hundred Names for Love:

What can you do with a pencil—other than write?

"Use as a spar in a model airplane," Paul had continued. "Measure distances. Puncture a balloon. Use as a flagpole. Roll a necktie around. Tamp the gunpowder into a pint-size musket. Test bonbons for contents. . . . Crumble, and use the lead as a poison."

This type of thinking can be taught and practiced, and can be nurtured in children as young as five years old. It is an increasingly important skill in a technology-driven world with untold unknowns. There are no right answers, just opportunities to exercise ingenuity, to find new connections, and to allow whimsy and experimentation to become a normal and habitual part of our thinking, which will lead to better problem solving.

It is important to teach our children to become lifelong learners, curious and inquisitive. Equally important is to instill in children a sense of play, that thinking isn't just serious, it can be fun. This entails giving them the freedom to make mistakes, to explore new thoughts and ideas outside of the ordinary—divergent thinking will be increasingly necessary to solve some of the biggest problems facing the world today. Benjamin Zander, conductor of the Boston Philharmonic, teaches young musicians that self-criticism is the enemy of creativity: "When you make a mistake say to yourself 'how interesting!' A mistake is an opportunity to learn!"

Where You Get Your Information

As with many concepts, "information" has a special and specific meaning to mathematicians and scientists: It is anything that reduces uncertainty. Put another way, information exists wherever a pattern exists, whenever a sequence is not random. The more information, the more structured or patterned the sequence appears to be. Information is contained in such diverse sources as newspapers, conversations with friends, tree rings, DNA, maps, the light from distant stars, and the footprints of wild animals in the forest. Possessing information is not enough. As the American Library Association presciently concluded in their 1989 report Presidential Committee on Information Literacy, students must be taught to play an active role in knowing, identifying, finding, evaluating, organizing, and using information. Recall the words of New York Times editor Bill Keller—it's not having the information that's important, it's what you do with it.

To know something entails two things: for there to be no doubt, and for
it to be true. "Religious fanatics ‘know’ no less than we scientists," Daniel Kahneman says. "The question might be ‘How do we know?’ What I believe in science is because people have told me things, people I know and whom I trust. But if I liked and trusted other things, I would believe and ‘know’ other things. ‘Knowing’ is the absence of alternatives of other beliefs.” This is why education and exposure to many different ideas are so important. In the presence of alternatives of other beliefs, we can make an informed and evidence-based choice about what is true.

We should teach our children (and one another) to be more understanding of others and of other points of view. The biggest problems the world faces today, of famine, poverty, and aggression, will require careful cooperation among people who don’t know one another well and, historically, have not trusted one another. Recall the many health benefits of being an agreeable person. This does not mean being agreeable in the face of views that are harmful or patently wrong, but it means keeping an open mind and trying to see things from another’s perspective (as Kennedy did in his conflict with Khrushchev).

The Internet, that great equalizer, may actually be making this more difficult to do than ever. Most everyone knows by now that Google, Bing, Yahoo!, and other search engines track your search history. They use this information for auto-complete, so that you don’t have to type the entire search term in the find window the next time. They use this information in two additional ways—one, to target advertising (that’s why if you searched for new shoes online, a shoe ad shows up the next time you log into Facebook), and the other, to improve the search results for each individual user. That is, after you search for a particular thing, the search engines keep track of which of the results you ended up clicking on so that they can place those higher up in the results list, saving you time the next time you do a similar search. Imagine now that the search engines have not just a few days’ or weeks’ worth of your searches, but twenty years of searches. Your search results have been iteratively refined to become ever more personal. The net result is that you are more likely to be given results that support your worldview and less likely to encounter results that challenge your views. While you may seek to maintain an open mind and consider alternate opinions, the search engines will be narrowing down what you actually see. An unintended consequence, perhaps, but one worth worrying about in a world where global cooperation and understanding are increasingly important.

There are three ways we can learn information—we can absorb it implicitly, we can be told it explicitly, or we can discover it ourselves. Implicit learning, such as when we learn a new language through immersion, is usually the most efficient. In classroom settings and at work, most information is conveyed in one of the two latter ways—being told explicitly or discovering ourselves.

The last two decades of research on the science of learning have shown conclusively that we remember things better, and longer, if we discover them ourselves rather than being told them explicitly. This is the basis for the flipped classroom described by physics professor Eric Mazur in his book *Peer Instruction*. Mazur doesn’t lecture in his classes at Harvard. Instead, he asks students difficult questions, based on their homework reading, that require them to pull together sources of information to solve a problem. Mazur doesn’t give them the answer; instead, he asks the students to break off into small groups and discuss the problem among themselves. Eventually, nearly everyone in the class gets the answer right, and the concepts stick with them because they had to reason their own way to the answer.

Something similar happens in art. When we read well-written literary fiction, for example, our prefrontal cortex begins to fill in aspects of the characters’ personalities, to make predictions about their actions, in short, we become active participants in figuring out the story. Reading gives our brains time to do that because we can proceed at our own pace. We’ve all had the experience of reading a novel and finding that we slow down in places to contemplate what was just written, to let our minds wander, and think about the story. This is the action of the daydreaming mode (in opposition to the central executive mode) and it is healthy to engage it—remember, it is the brain’s “default” mode.

In contrast, sometimes entertainment comes at us so rapidly that our brains don’t have time to engage in thoughtful reflection or prediction. This can be true of certain television programs and video games. Such rapidly presented events capture attention in a bottom-up fashion, involving the sensory rather than the prefrontal cortices. But it would be wrong to focus on the medium and conclude “books good, movies bad.” Many non-literary, or pulp fiction, and nonfiction books, even though they allow us to proceed at our own pace, present information in a direct way that lacks the nuance and complexity of literary fiction, and this was the finding of the
study briefly described in Chapter 4: Reading literary fiction, but not pulp fiction or nonfiction, increased the reader's empathy and emotional understanding of other people.

A striking parallel distinction was found in a study of children's television shows. Angeline Lillard and Jennifer Peterson of the University of Virginia had four-year-olds watch just nine minutes of *SpongeBob SquarePants* cartoons, a fast-paced television program, versus watching the slower-paced public television cartoon *Caillou*, or drawing pictures on their own for nine minutes. They found that the fast-paced cartoon had an immediate negative impact on children's executive function, a collection of prefrontal cortical processes including goal-directed behavior, attentional focus, working memory, problem solving, impulse control, self-regulation, and delay of gratification. The researchers point not just to the fast pace itself, but to the "onslaught of fantastical events" which are, by definition, novel and unfamiliar. Encoding such events is likely to be particularly taxing on cognitive resources, and the fast pace of programs like *SpongeBob* don't give children the time to assimilate the new information. This can reinforce a cognitive style of not thinking things through or following new ideas to their logical conclusion.

As in many psychological studies, there are important caveats. First, the researchers didn't test attentional capacity in the three groups of children prior to testing (although they did use the well-accepted method of random assignment, meaning that any a priori attentional differences should be equally distributed across the experimental groups). Second, *SpongeBob* is designed for six- to eleven-year-olds, and so its effect on four-year-olds may be limited just to that age group; the study didn't look at other age groups. Finally, the participants were an uncommonly homogeneous group of largely white, upper-middle-class kids drawn from a university community, and the findings may not be generalizable. (On the other hand, these kinds of issues arise in nearly all experiments in the psychological literature and they are no different for the Lillard and Peterson study than the limitations on most of what we know about human behavior.)

The tentative and intriguing take-home message is that reading high-quality fiction and literary nonfiction, and perhaps listening to music, looking at art, and watching dance, may lead to two desirable outcomes: increased interpersonal empathy and better executive attentional control.