Turning Numbers into Knowledge

MASTERING THE ART OF PROBLEM SOLVING

Third Edition

JONATHAN G. KOOMEY, PH.D.

Analytics Press
PO Box 4933
El Dorado Hills, CA 95762
http://www.analyticspress.com
http://www.numbersintoknowledge.com

EPILOGUE

SOME PARTING THOUGHTS

A lie can travel halfway around the world while the truth is putting on its shoes.
—**MARK TWAIN**

Soon after I completed the first edition of *Turning Numbers into Knowledge* in 2001, I investigated a high-profile example of bad statistics run amok. In this case, the tools and techniques described in this book, if widely applied, would have allowed businesses and investors to recognize these erroneous numbers for what they were before any harm was done.

DUBIOUS CLAIMS

Electricity used by computers was in the news around the year 2000, and some dubious claims were making the rounds:

- Claim # 1: The Internet accounted for 8% of all electricity used in the United States in 2000.
- Claim # 2: Computers and information technology equipment, including the Internet, used 13% of all U.S. electricity in 2000—and this total was expected to grow to half of all electricity use 10 to 20 years later.
- Claim #3: The "behind-the-wall" networking, server, and switching equipment needed to support a wireless personal digital assistant used as much electricity as a residential refrigerator.

Most major U.S. newspapers and business magazines, many respected institutions, and politicians of both political parties cited these assertions (the first one even came up in a Doonesbury comic strip at about the same time). Unfortunately, these claims are all incorrect.

These misconceptions originated in an article for *Forbes* magazine in May 1999 ("Dig More Coal—The PCs Are Coming"). They were perpetuated in a supporting report for the Greening Earth Society, in a *Wall Street Journal* op-ed piece, in congressional testimony, in articles for popular magazines, and in numerous interviews.

The claim that computers use lots of electricity spread quickly, driven by a superficially plausible story line and an ostensibly related high-profile crisis in the California electricity sector. *Forbes*, a respected magazine, lent credibility to the argument, as did the writer and publisher George Gilder, who published the *Digital Power Report* (which gave stock recommendations for investors assuming these assertions were true) about one year after the *Forbes* article came out.

The trade press, the investment community, and the popular media found the key claims in the *Forbes* article irresistible. They repeated those claims frequently, often without citing the original source—in effect, enshrining them as conventional wisdom. Leaders in business, government, and academia were misled by this barrage of media attention and cited the statistics widely, ensuring their continued proliferation (Koomey et al. 2002).

For example, reports of six major investment banks published from May to September 2000 accepted the assertions in the *Forbes* article as fact (see **Table EP.1**). Each report claimed to present unique insights on this topic, but each relied heavily on the *Forbes* arguments for justification (sometimes attributing their stock picks to that source, sometimes not).

THE REAL STORY

Over several years, my colleagues and I demonstrated in peer-reviewed journal articles, congressional testimony, conference papers, and technical reports that the figures from the *Forbes* article were incorrect (see the citations in the Further Reading section for this Epilogue). We investigated these claims and in virtually every case they represented overestimates of elec-

TABLE EP.1. Investment banking reports relying on the *Forbes* information technology electricity use figures

Publication and date	Quotation
(1) Deutsche Bank May 2000	"Mark Mills estimates that by 1999, the growth in (sic) Internet and related IT equipment now consumes 13% of our electricity supplies".
(2) Banc of America Securities June 2000	"Internet-related demand for power represented 8% to 13% of electricity consumption in 1999. . . . It is estimated that by 2010, one-half of U.S. electric consumption will be related to the Internet in some way".
(3) Stephens, Inc. August 2000	"The percentage of electricity consumed directly by the Internet is currently estimated to be 10% in the U .S., up from roughly zero in 1993, and there are no signs of slowing growth in the pervasiveness of the web. Some estimates project that the Internet and the equipment to support its growth will consume 50% of domestic power within 10 years".
(4) Credit Suisse First Boston Corporation August 14, 2000	"This increase in dependence on electricity is attributable to the ongoing transformation of our economy from an industrial basis to one centered on knowledge, intellectual capital and technology. According to one study, nearly 8% of all the electricity generated in the United States is consumed by PCs, the Internet, and its related infrastructure. By 2020, that same study projects this consumption to grow to nearly 50%".
(5) JP Morgan Sept 14, 2000	". . . information technology (IT) and telecom should account for an increasingly large piece of the total energy pie (up from about 16% today)".
(6) Salomon Smith Barney Sept. 25, 2000	"In 1995, the U.S. Department of Energy estimated that personal computers consumed approximately 3% of U.S. electricity supplies. Mark Mills, a well-known technology consultant, estimates that in 1999 Internet and related IT equipment consumed 13% of our electricity supplies".

(1) Tirello Jr., Edward J., Barbara Coletti, and Christopher R. Ellinghaus. 2000. *Convergence Redefined: The Digital Economy and the Coming Electricity Capacity Emergency.* Deutsche Banc Alex. Brown. May 12.

(2) Anderson, Hugh M. M., Russell L. Leavitt, James P. LoGerfo, and Daniel L. Tulis. 2000. *The Power of Growth.* New York, NY: Banc of America Securities. June.

(3) Stephens Inc. 2000. *Emerging Power Technology: Industry Report.* Little Rock, Arkansas: Stephens, Inc. August 11.

(4) Pencak, Marko, Neil Stein, Cameron Jeffreys, Christopher E. Vroom, Nat Schindler, Andrew Levi, Steven Parla, and Paul Patterson. 2000. *Energy Technology: An Overview.* Ontario, Canada: Credit Suisse First Boston Corporation. August 14.

(5) Feygin, Anatol, Waqar Syed, and Kyle Rudden. 2000. *Industry Analysis: We Need More Juice—IT and Telecom Growth Fuel Energy Demand—E&P, Natural Gas Pipeline, and Generators Should Benefit.* New York, NY: JP Morgan. September 14.

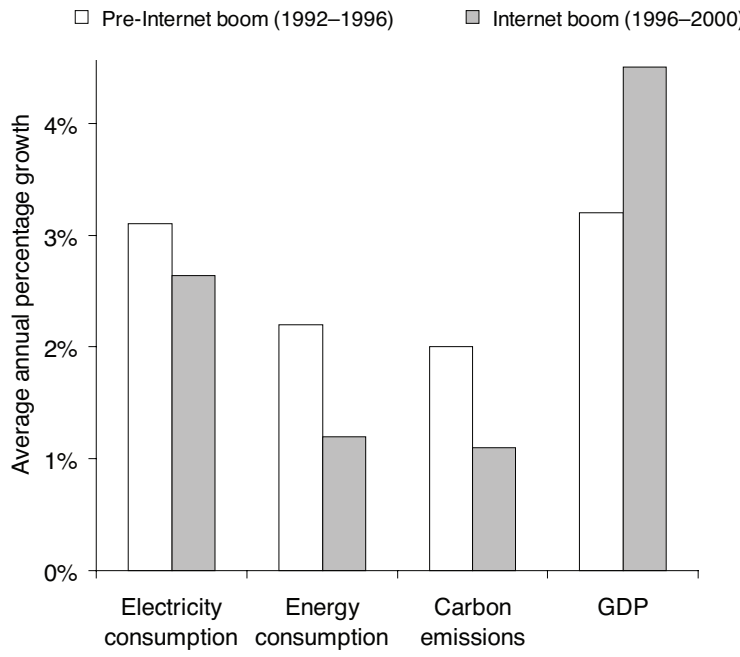
(6) Niles, Raymond C., Joanne M. Fairechio, and Christopher R. Ellinghaus. 2000. *The Power Curve.* New York, NY: Salomon Smith Barney. September 25.

tricity used by office equipment. For example, the Greening Earth Society report (Mills 1999) stated the assumption that a personal computer (PC) used 1,000 watts. Measurements I took at the time indicated that PCs typically used 60–80 watts. Commonly used 17-inch cathode ray tube monitors used about 90 watts in active mode, but the flat panel displays of comparable size (which were becoming more widely used around 2000) used a half or a third that amount. After accounting for other peripherals and "behind the wall" equipment, we concluded that 200 watts (not 1,000 watts) was a reasonable estimate for the active power of PCs, a factor-of-five reduction from the estimates in the *Forbes* article.

One would expect to find small differences between any such estimates, but when errors are all large and in the same direction it is likely that there is a bias in the analysis. Our studies found that the *Forbes* article overestimated power used by the Internet by at least eight times (i.e., the Internet used less than 1% of U.S. electricity in 2000). It also overstated the total power used by office equipment by about four times—the actual percentage of U.S. electricity used by all office equipment was 3% in 2000 (for details, see the web site at [http://enduse.lbl.gov/Projects/InfoTech.html]). The claim about wireless Palm Pilot electricity use was too high by a factor of two thousand (Koomey et al. 2004). The consulting firms Arthur D. Little and RAND published other independent analyses, funded by the U.S. Department of Energy, that confirmed our findings (Baer et al. 2002, Roth et al. 2002).

18 We also checked independent data sources, which showed that something was amiss with these claims. Joe Romm of the Center for Energy and Climate Solutions plotted **Figure EP.1** from Energy Information Administration data. The figure shows annual growth rates for U.S. electricity use, primary energy use, GDP, and carbon emissions for the 1992 to 1996 and 1996 to 2000 periods. While GDP grew faster in the second period, electricity use, energy use, and CO_2 emissions all grew more slowly in that period. If the thesis of the *Forbes* article was correct, electricity demand growth rates at the end of the 1990s (the early heyday of the Internet) would have increased, but in fact the opposite occurred. These data contradict the assertion that demand growth was stronger with the widespread use of the Internet.

FIGURE EP.1. Comparison of average annual percentage growth rates in electricity use, energy use, carbon emissions, and GDP for the pre-Internet and Internet boom periods.



SOURCE: Joe Romm of the Center for Energy and Climate Solutions, based on Energy Information Administration data, as cited in Koomey et al. (2002).

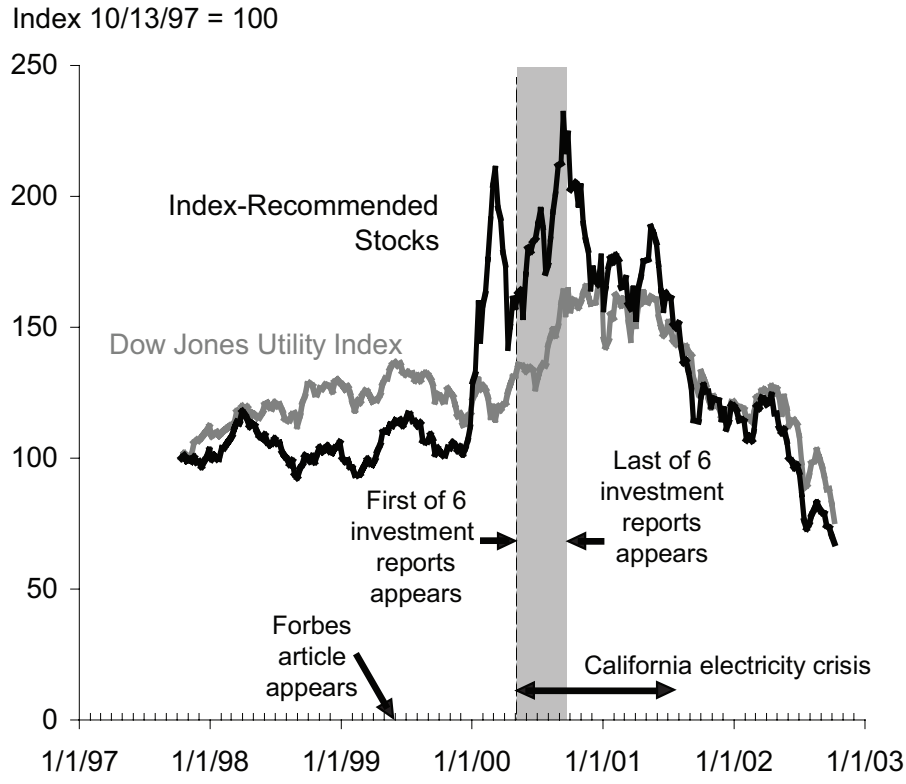
THE COST TO INVESTORS

In the absence of a formal "event study",[111] no one knows for sure if these erroneous data and the associated investment recommendations hurt investors. I present some relevant data below so that readers can make their own judgments.

The stocks favored by the investment banking analysts (Table EP.1) were all related to electric power. Some were utilities and large energy trading companies, others were smaller firms focusing on power quality, reliability, distributed generation, and power industry services. The looming electric reliability issues in California seemed to validate the focus on these firms.

Figure EP.2 plots an index of the 34 stocks recommended by the reports

FIGURE EP.2. Price-weighted index of investment bank recommended stocks compared to the Dow Jones Utility Index from 10/13/1997 to 10/6/2002, indexed to 10/13/1997 = 100.



SOURCE: Recommended stocks taken from reports listed in Table 1.[112]
Not all stocks were recommended by all reports. Stock prices downloaded from Yahoo! Finance in October 2002.

produced by the investment banks listed in Table 1. This index is weighted by stock price so that it is comparable to the Dow Jones Utility (DJU) index, also plotted here. The chosen stocks consistently underperformed the utility index until about January 2000, when over a three-month period their prices doubled. The price index of the recommended stocks then dropped by about one-third, but rose again to make up those losses and more by the end of September 2000. It is suggestive (but not conclusive) that a large increase in the price of these stocks occurred over the period in which the investment reports were published. Of course, the biggest troubles in California's restructured electricity markets also began to manifest themselves over the same period, but the hype generated by those events fed into the hype created by the investment reports, so they were self-reinforcing.

Let's assume an investor had followed the recommendations in these six reports and bought 100 shares of each of the 34 touted companies back in September 2000 (when the last of these reports came out). By September 2002 he would have lost two-thirds of his investment. For comparison, the DJU index fell by 44%, and while few stock pickers did well during this period, most would regard losing 22 percentage points more than the DJU index over a two-year period as truly dismal investment performance.

THE ADVENTURE CONTINUES

Unfortunately, being refuted in the peer-reviewed literature didn't stop the authors of the *Forbes* article from continuing their crusade. They repeated their claims in a 2005 book (Huber and Mills 2005), without acknowledging their errors. Mills also repeated the claim about iPhones and refrigerators in a 2013 report, which required yet another debunking.[113]

This is another feature of some peddlers of erroneous statistics: they won't fess up to their mistakes, and they keep right on repeating them. In addition, the quality controls on books and white papers are usually far less stringent than those for peer-reviewed journals, and so most anything can be published, regardless of veracity.

As this example shows, it takes many paragraphs (and a lot of work) to refute even just an errant sentence, so debunking of incorrect assertions always lags such claims. A variant of this latter thought has been dubbed "Brandolini's law", after Alberto Brandolini, a software consultant. The law reads

> The amount of energy necessary to refute bulls*$t is an order of magnitude bigger than to produce it.

Truer words were never spoken.

Another example is that of Bjørn Lomborg's book *The Skeptical Environmentalist*, which has been roundly criticized by virtually everyone who knows anything about the topics it covers, but that hasn't stopped its author from continuing to repeat his erroneous and misleading statements.[114] By not admitting to and correcting their errors, these authors have violated

the accepted code of conduct in science, best summarized by John Holdren, past president of the American Association for the Advancement of Science, in his rebuttal to *The Skeptical Environmentalist*:

> The practice of science, which includes the packaging of findings from science for use in the public-policy arena, is governed by an unwritten code of conduct that includes such elements as mastering the relevant fundamental concepts before venturing into print in the professional or public arena, learning and observing proper practices for presenting ranges of respectable opinion and uncertainty, avoiding the selection of data to fit pre-conceived conclusions, reading the references one cites and representing their content accurately and fairly, and acknowledging and correcting the errors that have crept into one's work (some of which are, of course, inevitable) after they are discovered by oneself or by others.
>
> Most scientists follow this code of conduct as best they can out of self-respect and respect for the integrity of science itself. For those for whom these considerations might not be quite enough, there is little that can enforce the code other than concern with the cumulative harm to one's reputation and standing that comes from one's colleagues' awareness of a pattern of infractions, or fear of the public denunciation by colleagues that may follow in the rarer instances of someone's descending into more massive and willful disregard of accepted standards.

Violations of this code of conduct are anathema to scientists. Science is rare among human endeavors in that real progress over decades is easy to show. Better science is more accurate and powerful in its description of the natural world, as demonstrated by experiment, data collection, and the embodiment of scientific knowledge in technology. When serial obfuscators throw sand in the eyes of the scientific and policy process, it is an affront to all those who value truth over self-interest and muddy thinking.

INTEGRITY AND DECISION MAKING

I saw Scott McNealy, chairman of Sun Microsystems, address the World Business Forum in October 2007. He returned again and again to the importance of integrity to success in business, because it enhances your ability to achieve measurable results, improves your credibility, and builds trust between you, your colleagues, and your competitors.

Webster's Ninth New Collegiate Dictionary defines integrity as "firm adherence to a code of especially moral or artistic values: incorruptibility". Wikipedia (accessed June 20, 2017) states that a person with integrity bases her actions "on an internally consistent set of principles" [http://en.wikipedia.org/wiki/Integrity]. Analysts have integrity when they follow the code of conduct outlined by Holdren and the analytical principles outlined in this book. At the core of these principles is a respect for truth above all else, even when the truth contradicts long-held beliefs. **4**

Robert Ringer, one of my favorite authors, summarized his theory of reality as follows:

> This theory emphasizes, first of all, that reality isn't the way you wish things to be, nor the way they appear to be, but the way they actually are. Secondly, the theory states that you either acknowledge reality and use it to your benefit or it will automatically work against you.

Ringer uses the term "reality", but he could just as well have used the words "the truth" in his theory. He captures an essential lesson for those using analysis to improve decision making, and that is, after all, the name of the game.

Getting the numbers right really does matter. All business and policy decisions today are based at least in part on quantitative data, and no good can come of incorrect information being widely accepted. Annoyance, inconvenience, financial hardship, or even disaster can follow from important decisions based on faulty data.

We can't undo what's been done, but there is plenty that all of us can do from now on to ensure that we aren't being misled by bogus information. The widely cited but misleading statistics I mention here are just the tip of the iceberg. Your best defense against being fooled is to think for yourself **11** and do your homework. Never make important decisions based on "com- **14** mon knowledge" unless you can verify its accuracy independently. Make sure you "follow the money" to determine whether someone has a vested **24** interest in one outcome or another. And be skeptical about recommendations in investment reports, even if you know and trust the analyst's integrity **16** and judgment.

You can improve your own analytical integrity by applying the techniques described in this book. Doing so can be time consuming, but if a set of

facts is central to an important decision—for example, a business plan, a marketing strategy, or a new direction in research—then you need to put those principles to work.

CONCLUSIONS

People with integrity are the ones who succeed in the long run. The question remains, however: how can you best develop analytical and personal integrity? I offer these final thoughts:

- *Say what you mean and mean what you say.* If you promise something, always follow through—the recipient will reward your diligence when *you* need help next time.
- *Stick close to the data.* Don't make claims the data don't support, and don't let others do so. Never exaggerate, distort, palter,[115] or dissemble, or the results will come back to bite you in the butt (which is a more colorful way of rephrasing Ringer's "Theory of Reality").
- *Assume that others will replicate your work.* If it's interesting and important enough they surely will, so document it well and be accurate in the claims you infer from it.
- *Focus on getting the job done, and done well.* We've all encountered people whose personal issues, biases, and interests win out over the larger goals, and those are people to avoid if you can. True professionals can properly handle personal interests in a way that still leads to the best outcome for the team, but such professionals are the exception rather than the rule. You should be such an exception.

Analysis is, at its heart, a tool for building a better world. I wish you great success in putting it to work!

Always do right. This will gratify some people and astonish the rest.

—**MARK TWAIN**

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

This is a human-readable summary of (and not a substitute for) the license. Disclaimer.

You are free to:

Share — copy and redistribute the material in any medium or format

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

NoDerivatives — If you remix, transform, or build upon the material, you may not distribute the modified material.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.