How to Tell the Liars from the Statiscians

- By Robert Hooke
- 1983 - Marcel Dekker: New York, NY.
- Selected insights by Douglas W. Green, EdD
- DGreen@Stny.rr.com

Why This Book?

- One way I keep up to date is to go to the Binghamton University (NY) Library and browse the books in the New Arrivals shelves. This way I get the first shot at anything that captures my interest. On July 22, 2010, I spotted this book. I was surprised to see the 1983 publication date, but as I read it, I realized that it contained a number of insights valuable to educators and just about anyone who had to deal with modern society. There have been no real breakthroughs in the field of statistics since this book was published. The author presents the material is such a manner that you need not be good in math to understand it. Each of the 76 chapters are only a few pages in length. I have summarized many key concepts, but there is much more I left out. If you find my work interesting, then you certainly want to purchase this book.

Data Pushers

- Data pushers are somewhat like dope pushers, who try to gain control over us with their product. They include those who deliberately try to deceive us with true but misleading data, those whose enthusiasm for a cause leads them to do this unconsciously, and those who merely combine their misinformation with persuasiveness. The data pusher's job is made easier by the fact that there are so many people who shy away from numbers as if they were poisonous reptiles. Probabilities are expressed as decimals, fractions, or percentages, and fear of these things makes one an easy mark for the persuaders. The habit of reading or listening critically when people are quoting numbers is one you should adopt if you don't already have it.

Emphasis on the Consequences

- Anytime you start to worry about a particular possibility, you need to focus on how likely it is to happen, and what are the consequences if it does happen? A favorite device of data pushers is to put so much emphasis on the consequences that you may forget that the event in question rarely or virtually never happens. Even if you are a heavy worrier, you have to tell yourself from time to time that what you are worrying about happens too rarely to give you concern.

Its all about uncertainty.

- All of us are constantly affected by the uncertainty of events, and we make decisions based on what we think the chances of various outcomes are. Unfortunately, little probability is taught in our schools and what is taught often involves things like cards and dice. Knowing the odds of something happening in real life can help you sensibly decide whether or not a given risk is worth taking.

The Broad-Base Fallacy

- Rare events can be made to sound commonplace by the simple act of telling how many people it happens to. This is called the broad-base fallacy because data pushers sometimes use it to appeal to our emotions in cases where we should be considering the odds. A favorite media example is to site the number of fatalities on holiday weekends, which makes travel seem dangerous. A look at the facts reveals that the number of fatalities per vehicle or per mile driven, don't change much. While 10,000, sounds like a lot of deaths from a given disease, an American has a 1 out of 30,000 chance due to the size of our population. (Doug: You need two numbers to know the odds of something happening. Watch out when someone only gives you one number.)

The Unmentioned Base

- If a company's profit increases from 1% to 2%, you can say they had a 1% increase. A more positive way would be to say they had a 100% increase since the profit doubled. In the first case the base is the total sales. In the second case, the base is the total profit. Every percentage is a percentage of something, and this something is called the base. Politicians often play with the base to persuade. When anyone quotes a percentage, you need to ask, "percentage of what?" You can also deceive if you change percentages to totals or rates. If a utility raises the cost of a kilowatt hour from 10 cents to 11 cents, you could call it a 1 cent increase or a 10% increase. The same increase for all customers during a year could be expressed in the millions. The number you pick to emphasize depends on where you stand.

The Fiction Problem

- People who write fiction, find that it is more interesting to write about rascals than saints. This provides a way for false statistics to enter our lives so stealthily that we scarcely notice them. Scoundrels are more common in fiction than in real life. People who fail to notice this tend to conclude that preachers are mostly hypocrites (think Elmer Gantry), business people are mostly crooks, scientists are mostly mad, and private eyes are all smarter than policemen. Racial and other stereotypes in fiction occur from the same effect. A classic is that the majority of Italians in fiction belong to the Mafia. A novel can expose the possibility of some kind of behavior, but a group of novels can't tell us how often that behavior occurs in real life.

When the Truth Is not the Whole Truth

- Data pushers like to present numbers out of context. If a car company says that 90% of its cars are still on the road, it could be that they are only four years old or have gotten much bigger recently. A 1947 study showed that college educated lawyers made less than those without degrees. At the time, however, college education was relatively new for lawyers. This meant that the college educated group was younger and less experienced. When these factors were considered, the college group made more. A report from the 1970's showed a higher cancer rate in cities that added fluoride to their water. When age, race, and sex were considered, the cities with fluoride had somewhat smaller cancer mortality rates.

When Should We Take Chances?

- Knute Rockne was the nation's most successful football coach in his time at Notre Dame. His style was very conservative, which was suited to his teams as they usually had more strength and talent. When others copied his style, college football became a dull, low-scoring game. Some professional coaches realized, however, that for less talented teams to have a chance to win, they needed to take more chances like passing. Risky plays have a bigger payoff when they work and are more costly when they fail. This concept can be applied to other sports or even test taking. (Doug: In any situation, you need to analyze the risk and the reward that goes with it as well as where you will end up if the risk results in a negative outcome.)

How many tellers do you need?

- If on the average, a bank has one customer enter each minute and each transaction takes an average of one minute, how many tellers does the bank need? The answer would be one if the customers arrived one minute apart with minute-long transactions. Due to the variable arrival and service times, the single teller will sometimes have nothing to do, and productivity lost during these idle times can never be regained. So with one teller, the average line will build up indefinitely. This type of analysis falls into the field of queuing theory. (Doug: Queuing is what the English call waiting in line, and based on my recent trip to England, I found them to be much more particular about it than Americans.)

What's luck got to do with it?

- Luck is the effect of circumstances over which we have no control. If two contestants are equal in every aspect of skill, then the outcome will be entirely a matter of luck. If you look at the top and bottom performers in any sport one year, it is likely that many will not be on this list the following year. Golf is a good example as it is rare for the same person to win two weeks in a row. In a given league, the top rookies are on the lucky side more often than not. This results in many performing not as well the next year, which gives us the "sophomore jinx." (Doug: The same logic applies to test taking. This includes lucky guesses on multiple choice tests along with luck in regard to knowing the answers to the specific questions that are on the test.)

Two Types of Error

- All methods of detection are subject to two types of error, from smoke alarms, to burglar alarms, to methods of detecting cancer or tuberculosis. Type I errors occur with you detect something that isn't there, such as the smoke detector going off without any smoke. Type II errors are when you fail to detect something that is there. Our judicial system makes both kinds of errors when innocent people are convicted (I) and guilty people get off (II). Liberals are more likely to take action when it isn't needed (I), while conservatives are more likely to avoid action when it is needed (II). (Doug: You can protect yourself from Type I errors by confirming something with a second type of test. In the case of a Type II error, you will just have to deal with the consequences.)

The Problem of Scale

- Most things, be they organizations or physical objects, don't scale well. You can't simply make an organization bigger without changing some of the ways it operates. If you made people or animals a lot bigger, they couldn't support their own weight. (Doug: Designs that work at one size, often breakdown at a very different size. Keep this in mind when you want to do anything at a larger scale. This applies to everything from class size to making a meat loaf. In a small school, just about everyone can play a varsity sport. In large schools, only the most skilled get to play.) We also suffer from linear thinking. If losing 10 pounds is good, losing 100 pounds is 10 times better. In reality, most people would die if they lost 100 pounds. Also, a little fertilizer helps plant grow. A lot will kill it.

Sampling Error

- When you want to test the quality of a product or predict the outcome of an election, you need to select a sample. If you are checking cookies for the ratio of nuts to chocolate, your test will only work well if the batter is well mixed. Human populations, however, are not well mixed, which makes a truly random sample impossible to obtain. The nature of the questions can also drive the results. (Doug: This allows political data pushers to demonstrate just about anything by rigging the questions. You need to know the bias of any organization that publishes political polls.) Your experience in life is a sample from an immense rage of possible experience. (Doug: This means that we all have different experiences, influences, and perspectives.)

Significance and Correlation

- In order to be statistically significant, a result has to have a high probability of not being the result of chance. (Doug: Typical tests will use tables to determine if a result has a 95% or 99% probability of not being the result of chance.) The larger the number in your sample, the easier it is to test for significance. It is also important to select your sample in a random manner.

- Two things correlate if they tend to rise or fall together. People's height and weight have a positive correlation as taller people tend to weigh more. Weight and speed have a negative correlation as heavier people tend to run slower. Correlations vary from +1 to -1 and the farther you are away from 0, the stronger it is. Just because two things correlate, doesn't mean one causes the other.

The best researchers are blind.

- Hooke describes a number of interesting experiments that have serious design flaws. In addition to the random selection of a test group and a control group, the best experiments also require that the subjects and the people doing the testing don't know who is getting the treatment and who is not. In drug testing, half generally get the drug being tested and half get a placebo (sugar pill). Such tests are called "double blind".

- The book ends with the notion that, even though you may find numbers themselves dull, it is important that you question numerical results that are stated as hard facts by people in the media and people who are selling things.

Cool Quotes

- If you can't distinguish between your chances of being killed by a meteorite and your chance of dying of lung cancer after smoking two packs of cigarettes a day for many years, then you are a prime target for data pushers.
- He who never makes a mistake never makes anything.