# STATISTICAL PREVARICATION: TELLING HALF TRUTHS USING STATISTICS

SCHIELD, Milo
W. M. Keck Statistical Literacy Project, Augsburg College
USA

*All too often statistics are characterized as lies. But statistics are more likely to be half truths than lies. This paper studies statistical prevarication – the art of straddling both sides of an issue or idea – involving a statistic. This paper studies statistical prevarications in everyday use and in statistics education. If statistics educators are to avoid a charge of statistical negligence, they should focus more on identifying and eliminating sources of statistical prevarication in their teaching and textbooks. And statistical educators should do more to help students become statistically literate in detecting statistical prevarication.*

## PREVARICATION: HALF TRUTHS

I first leaned about prevarication from my mom. I was a little kid playing with my younger brother. My younger brother began crying. My mom called out, “What happened?” Knowing that my mom disliked lying, I said honestly, “He fell.” My Mom came back and found out my younger brother had fallen because I had pushed him. She looked down at me and said, “You prevaricator!” I had never heard the word before; I knew I hadn’t lied, but I knew that I had told just part of the truth. To this day, I remember my first experience with prevarication.

Prevarication is bending the truth using an ambiguity, an equivocation or an omission of something material For simplicity we use ‘half truths’ to describe both ‘bent truths’ and ‘incomplete truths.’ Statistical prevarication is well described by this analogy. “*Statistics are like a bikini. What they reveal is suggestive, but what they conceal is vital*.” Prevarication is different from falsehoods (lies) and from reader errors.

A most famous statistical falsehood was noted by Joel Best (2001). “During the past 50 years, the rate at which kids died *doubled* each year.” This should have been stated as, “During the past 50 years, the rate at which kids died each year *doubled*.” Other falsehoods include “Three is 3 times more than one,” “Interest rates increased by 2% in going from 1% to 3%” or “Sale price is four times less than retail” (Schield, 1999).

Reader errors can occur even when a statistic is meaningful and accurate. An innumerate reader may think a big difference means a big effect. They hear that “The sun is about 5 million miles closer to the earth at its closest than at its furthest; they mistakenly conclude that difference is so big that it causes the seasons: the difference between summer and winter. In fact, that difference is less than 5% of the average distance and it does not cause the seasons. An innumerate reader may think a big ratio means a big difference. They hear that, “In the US in 1998, the synthetic-drug arrest rate was twice as high in the West (10 per 100,000) as in the Midwest (5 per 100,000). An innumerate reader mistakenly concludes this big ratio (twice as high) means there is a big difference when the actual difference is only 5 such arrests per 100,000 people. An innumerate reader hears that “More doctors like Crest” and presume this means that “Most doctors like Crest.” An innumerate reader hears that interest rates increase from 1% to 3% and mistakenly concludes this is an increase of 2%.

Reader errors may result from their failure to recognize that all statistics are socially constructed (Best, 2001). Readers must learn to “Take CARE” (Schield, 2005).

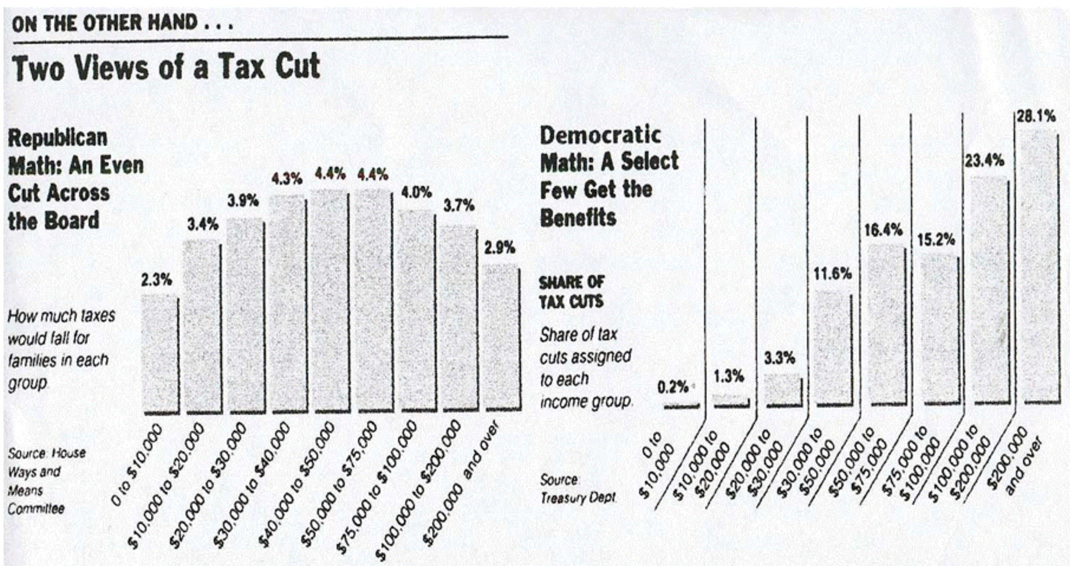
To review, statistical prevarications (half-truths) are different from falsehoods (lies) and from reader errors. Let’s examine some obvious statistical prevarications found in everyday life and then turn to some that are found in statistical education

## STATISTICAL PREVARICATION IN EVERYDAY LIFE

Consider some common statistical prevarications involving omissions or equivocations.

- "Sale: 50% off." The half truth omits the basis: 50% off of what starting price? Is this the dealers cost, yesterday's price or the manufacturer's suggested retail price?
- "More doctors prefer Crest." The half truth omits the rest of the comparison. More doctors prefer Crest more than any other toothpaste? Than cigars? Than dentists like Crest?
- "Are you more likely to worry about your health than most people?" Answer: "Yes, most people don't worry about my health." The half truth used an incomplete comparison.
- "Jimmy John's Sandwich franchise is the fastest growing franchise in the US." The half truth omits the base. It is easier for a small operation to grow fast than for a large one.
- "Black sheep eat less grass than white sheep." This half truth equivocates on whether 'sheep' is considered collectively (as a group) or individually (per sheep).
- "The typical salary is $80,000." The half-truth equivocates on whether 'typical' means average (the mean) or most common salary (the mode)?

Sometime the statistics are presented accurately, clearly and unambiguously. But there can still be prevarication in their use. Consider this graph provided by Bernie Madison.




How can the rich (incomes of $200K and up) get only a 2.9% tax cut and yet get 28.1% of the tax cut? It seems they can get either 2.9% or 28.1% but not both. One key word is "of" which introduces the whole or the pie – the entire tax cut, so the 28.1% is their share of the total. The other key word is 'cut.' This 2.9% tax cut is a percentage change – a totally different kind of percentage than the first percentage (part-whole). Both graphs are clearly marked. Both sets of statistics can be true for the same group of people at the same time. As such there is no prevarication. Suppose we asked, "What fraction of the tax cut goes to the rich?" If a politician replied, "The rich only get a tax cut of 2.9%" then they would be prevaricating. Suppose we asked, "What is the reduction in taxes for the rich?" If a politician replied, "The rich will get 28.1% of the tax cut", then they would be prevaricating. In each case, the respondent told a truth, but not the whole truth relevant to the question being asked.

Comparing statistics is a fertile source of prevarication. Consider a comparison where the relevant base is not stated and there are two different – but relevant – candidates. Suppose we have a fixed pie as in the vote for the candidates and we note that one candidate's share increased from 10% to 15%. As a percentage change, it increased by 50%. As a share of the total vote, it increased by 5%. So is the increase 50% or 5%? Unless the basis of the comparison is clearly stated – as part of the sentence – as being the 100% size of the fixed pie, the default is to presume the comparison is a percentage change. Journalists who use this '% increase of 100% pie' without explicitly mentioning the fixed pie participate in statistical prevarication. Rates are a fertile field for statistical prevarication when a relevant whole is omitted.

- Sometimes an omitted term does not involve a half-truth. "In 2001 the death rate was 67% higher in the US (8.7/1,000) than in Mexico (5.0/1,000)." While Mexico and the US delimit the whole (the denominator) they don't identify the people involved. But for death rates, it is well understood that the entire population is normally included.
- Between 1980 and 1999 the birthrate dropped by more than 8% in the US (from 15.9 to 14.5 per 1,000 population), but it dropped by less than 3% (from 68.4 to 65.9 per 1,000 women age 15-44). So did the birthrate decrease by 8% or 4%? Since there is no natural whole for birthrate, the rates and their comparisons are both ambiguous.
- In 1995, the auto accident rate (per 1,000 miles of road) was 5 times as high in Hawaii (35) as in Arkansas (7), but the auto accident rate (per 100,000 vehicles) was 2 times as high in Arkansas (36) as in Hawaii (18). So is the auto accident rate higher in Hawaii than in Arkansas or lower? With these half-truths, we can't say.
- The 4-year graduation rate of students was higher at private colleges than at public colleges. The numerator is slightly ambiguous (graduate in exactly four years or graduate within four years or graduate in about 4 years).The denominator (students) is quite ambiguous: (a) all students who enter as Freshman, (b) all students who enter as Freshman and take a full load of classes every term, or (c) all students who enter as Freshman, take a full-load of classes every term and get a grade of C or better in each class.

Percents and percentages are also a fertile field for statistical prevarication. Percent grammar, X% of {whole} are {part}, is quite difference from percentage grammar, the percentage of {whole} who are {part} or the percentage of {part} among {whole}. In percent grammar, the phrase 'X% of" always introduces the whole – the pie. But in percentage grammar, the phrase, 'the percentage of' can introduce either the part or the whole. So when a graph is headed by "Percentage of" the reader cannot tell whether the phrase introduces the whole or the part. The title prevaricates. If the part is intended, a better – but longer – title would be, 'Percentage who/that…' See Schield (2000).

Sometimes the keywords 'percent' and 'percentage' are mistakenly treated as being interchangeable when they are not. Percent is a unit of measure (like kilogram) while percentage is the object of measurement (like weight). Percent can always be preceded by a number whereas percentage can never be preceded by a number. This can create a problem in a title of a table or graph as in "Percent of Unemployment." A reader expects that unemployment is whole, when the writer intended that it be part. See Schield (2000).

Accuracy is a fertile field for statistical prevarication. Expressions like '% accurate' or '% chance' can be ambiguous. Consider a medical test for HIV-1 that claims 99.9% accuracy.


An innumerate reader might conclude this means that 99.9% of those who test positive have AIDS (HIV antibodies) But if the person being tested is a member of a group where the actual incidence of AIDS is very low (say one in 1,000), then 50% of the people in that group who test positive will be false positives. How can a test that is 99.9% accurate have a 50% mistake rate? The 99.9% accuracy is accuracy in confirmation: 90% of those who have HIV antibodies will test positive. The 50% is accuracy in prediction: 50% of those who test positive will have AIDS. The term 'accuracy' is equivocal and so the related statistic is a half truth. This ambiguity in '% accuracy' makes the related claim a statistical prevarication.

Chance is an extremely fertile field for statistical prevarication. Chance grammar is the most common form of ratio grammar. Probability is a member of this family. Probability is often taken as resulting from a Bernoulli process where the probability is constant over time, where the trials are independent and the outcomes are otherwise indistinguishable. A Bernoulli probability may be very different from the probability obtained from historical relative frequencies. Consider the death rate in the US of 87 deaths per 10,000 people. If this were the result of a Bernoulli process, a statistician would be entitled to conclude that the 87 per 10,000 is

A1 The chance of randomly selecting someone who would die in the next year.
A2 The chance that any subject would die in the next year.
A3 Your chance of dying in the next year.

But there is nothing in nature that says that dying is adequately modeled by a Bernoulli process. People are distinguishable in ways that matter. In reality, the first claim (A1) is factual when applied against the past (resampling), but is contentious when applied to the present or future. The second (A2) and third (A3) are also contentious when said of a specific individual. To view a given person as being no different from a randomly selected individual without stating that critical assumption is definitely a prevarication.

Comparisons involving 'likely' are a fertile source for statistical prevarication. Consider the comparison, "Men are more likely to die accidentally than women are [to die accidentally]." This is often restated, "Men are more likely to die accidentally than are women." But when the match to the main verb is dropped, the syntax no longer indicates the parallel structure: "Men are more likely to die accidentally than women." But so long as the reader knows the content, they can decode the comparison. When the reader cannot decode the content, the comparison becomes ambiguous and the claim is a half-truth. "X are more likely to Y than Z." See Schield (2001).

## STATISTICAL PREVARICATION IN STATISTICAL EDUCATION

Despite the many misuses of statistics in the everyday press, statistical prevarication may be more common in statistical education. Consider the following instances.

### STATISTICAL PREVARICATION USING CONFIDENCE INTERVALS

Statistical confidence is a fertile field for statistical prevarication – even though statistical educators are absolutely clear on what they think it means. Suppose we are given a fixed 95% confidence interval involving a sample statistic from a random sample along with the unknown value of the associated population parameter. Consider three claims.

B1. 95% of the intervals from random samples will contain the population parameter.
B2. There is a 95% chance that this kind of interval contains the population parameter.
B3. There is a 95% chance that this particular interval contains the population parameter.
B4. There is a 95% chance that this interval contains the population parameter

For a Frequentist, B1 is clearly valid while B3 is clearly invalid – the population parameter is either inside or outside this particular interval. So long as B2 is a restatement of B1, it is valid. B4 is really ambiguous – somewhere between B2 and B3.

Saying we are 'statistically confident' is also a fertile field for statistical prevarication. Consider the phrase, "95% confident." What does this really mean for confidence intervals?

C1. Ninety-five percent of these intervals will contain the population parameter.
C2. There is a 95% chance that the fixed population parameter is in this fixed interval.
C3. You should act as if C2 were true.

For a Frequentist, C1 is clearly stated and true while C2 is false. So how can C1 be considered a statistical prevarication? Because C1 doesn't address the question: "what should one do if one is 95% confident?" If one has a choice of betting that the next randomly selected ball will be red (when drawing from an urn of 20 balls where 19 are red) or of betting that a fixed 95% confidence interval contains the fixed population parameter, what should one do? Schield (1997) argued that one should be indifferent provided we have no additional knowledge. Regardless of whether this interpretation is correct, the failure of statistics texts to talk about C3 results in statistical prevarication when students ask, "What should I do if I am 95% confident?"

### STATISTICAL PREVARICATION USING CHANCE

The role of chance provides a fertile source for statistical prevarication. Consider this statement: "In sampling, this sample statistic (or one further from the population parameter):"

D1. is extremely unlikely IF due to chance
D2. is extremely unlikely TO BE due to chance
D3. is extremely unlikely due to chance

From a Frequentist perspective, D1 is a valid deductive inference since chance is the premise; D2 is an invalid inference since chance is the conclusion, and D3 is a half truth – a statistical prevarication – since it is ambiguous as to whether chance is a premise (D1) or conclusion (D2). Yet all too many statistics textbooks use D3 in their presentations.

Perhaps the most famous such prevarication was by Sir Ronald Fisher when he said, "First convince us that a finding is *not due to chance*, and only then, assess how impressive it is" McLean & Ernest (1998, p. 18, italics in original).
http://aaaeonline.ifas.ufl.edu/NAERC/2001/Papers/portillo.pdf#search='due%20to%20chance%20ronald%20fisher'

The variations of Sir Fisher's statement would be:

E1. Convince us that a finding is not likely if due to chance.
E2. Convince us that a finding is not likely to be due to chance.
E3. Convince us that a finding is not likely due to chance.
E4. Convince us that a finding is not due to chance.

From a strict Frequentist perspective, E1 is proper and E2 is improper while E3 and E4 are ambiguous half-truths. See Schield (1996).

Note that in each case, it is the second form that the reader is interested in learning about. And in each case the Frequentist statistician is upholding the validity of the first form. So the subsequent forms – the ambiguous forms – stand as prevarications and allow unwary readers to draw false conclusions.

## STATISTICAL PREVARICATION USING 'STATISTICALLY SIGNIFICANT'

In Frequentist thinking a result may be said to be "significant" if it is "statistically significant": the rejection of the null hypothesis of no difference regardless of how small that difference may be (see above). But to readers unaware of this technical usage, saying that a result is "significant" may indicate that the result is material and important – none of which is implied given that the result is just "statistically-significant." The adjective "a statistically-significant result" should never be shorted to read "a significant result" unless one intends to prevaricate by telling a half-truth.

## STATISTICAL PREVARICATION USING 'HYPOTHESIS TEST'

The phrase 'hypothesis test' is a statistical prevarication. Students may presume this means a sample statistic will be used to test the truth of the null or alternate hypothesis. Wrong. This test assumes the null hypothesis is true and tests how likely this sample statistic (or greater) is. A better title would be "A Sample Statistic Test.'

## STATISTICAL PREVARICATION INVOLVING ALPHA AND TYPE-1 ERROR

The definition of alpha is an extremely fertile field for statistical prevarication. In a fixed level test, alpha is typically defined as "the probability of Type 1 error." Type 1 error is typically illustrated as a single cell in a 2x2 table of counts: a simple intersection of two separate but equal conditions – the union of a rejection and a true null hypothesis. So talking about the chance of Type 1 error in hypothesis tests is as ambiguous as talking about the chance of a false positive in medical tests. Statisticians intend their definition of Type-1 error as a conditional relationship: rejecting $H_O$ when (given that) $H_O$ is true. Yet this conditionality is not meaningful in defining the presence of Type I error; it is only meaningful when talking about alpha as a probability. If the goal were to deceive students, one could not do much better than this.

Note: there is no prevarication in the full statement: alpha is the probability of error in rejecting $H_O$ when/given that $H_O$ is true. While students recognize 'among' and 'of' as whole indicators in a part-whole ratio, they may need help in seeing conditionality as a whole indicator.

## STATISTICAL PREVARICATION INVOLVING P-VALUE

The interpretation of a p-value is a fertile field for statistical prevarication. If the stated p-value is used as the rejection level, then a p-value of 5% could mean:

F1. 5% of random samples from the null distribution have this sample statistic or greater.
F2. 5% of the samples having this statistic or greater come from the null distribution.

F3. There is only a 5% chance that the null hypothesis is true.
F4. One should act as if there is only a 5% chance that the null hypothesis is true.

Again, for a Frequentist, F1 is true (accuracy in confirmation) while F2 and F3 are not valid (accuracy in prediction). Yet F2 and F3 are what readers are interested in. Students want to know which is the stronger evidence for the truth of the alternate: a p-value of 2% when the alternate is that the heights of men and women are different – or – a p-value of 0.02% when the alternate is the truth of psychic phenomena. See Utts (1996). Schield (1996) argued that if the null and alternate hypothesis were equally likely to be true in the decision-maker's mind, then F4 would be true but as the truth of the alternate becomes less plausible, the p-value needed to give strong evidence for accepting the alternate must decrease. Regardless of whether Schield's thinking is accepted, the failure of statistical educators to answer the strength of evidence questions asked by decision makers constitutes a pernicious form of statistical equivocation.

## STATISTICAL PREVARICATION BY OMITTING THE ROLE OF CONTEXT

Omissions are a fertile field for prevarication. Statistical educators are well aware that a statistically significant difference can become statistically insignificant when another factor is taken into account in a multivariate regression – and vice versa. But this crucial fact is seldom if ever found in an introductory statistics textbook. This omission is statistical prevarication when students ask, "What are the most important aspects of statistical significance?" (Schield, 2004). Since most students studying statistics are in majors that use statistical associations from observational data as evidence for causal connections, this omission borders on statistical negligence.

## CONCLUSION

Statistical prevarication is unfortunately common. But statistical prevarication is all too common in statistical education. By allowing statistical prevarication in their teaching and their textbooks, statistical educators may be responsible for much of their students' misunderstanding about conditional probability, confidence intervals, statistical confidence and statistical significance. Their continuing failure to address this problem may constitute statistical negligence.

Statistical educators need to focus more on eliminating statistical prevarication from their writing and speaking, and focus more on informing students of these half-truths so that students can be properly trained to think critically about statistics as evidence.

## ACKNOWLEDGEMENTS

To the W. M. Keck Foundation for their grant "to develop statistical literacy as an interdisciplinary curriculum in the liberal arts."

## REFERENCES

Best, J. (2001). *Damned Lies and Statistics.* University of California Press.

McLean, J.E., & Ernest, J.M. (1998). The role of statistical significance testing in educational research. *Research in the Schools, 5*(2), 15-22.

Schield, M. (1996). Using Bayesian inference in classical hypothesis testing. *ASA Proceedings of the Section on Statistical Education*. www.StatLit.org/pdf/1995SchieldASA.pdf.

Schield, M. (1997). Interpreting statistical confidence. *ASA Proceedings of the Section on Statistical Education.* Copy at www.StatLit.org/pdf/1997SchieldASA.pdf.

Schield, M. (1999). Common errors in forming arithmetic comparisons. *The Association of Public Data Users, 1*. Copy at www.StatLit.org/pdf/1999SchieldAPDU.pdf.

Schield, M. (2000). Statistical literacy: difficulties in describing and comparing rates and percentages. *ASA Proceedings of the Section on Statistical Education.* Copy at www.StatLit.org/pdf/2000SchieldASA.pdf.

Schield, Milo (2004). *Three Graphs to Promote Statistical Literacy*. Presented by distribution at ICME-10. Copy at www.StatLit.org/pdf/2004SchieldICME.pdf.

Schield, Milo (2005). *Designing a Statistical Literacy Textbook*. ASA Proceedings of the Section on Statistical Education. Copy at www.StatLit.org/pdf/2005SchieldASA.pdf.

Utts, Jessica (1996). Exploring Psychic functioning: Statistics and Other Issues. *STATS* #16 p. 3.