Can you lie with statistics?

Some people would have you believe that they have their facts straight by quoting statistics. Statistics is widely used in proving one's point in academic debates or to provide empirical evidences in social and scientific researches. Yet are there ways to fudge supposed statistical truth? Are there studies that use correct statistical methods and yet arrive at unreliable results? What are the guidelines in spotting lies behind the statistics we see everyday and when can we say that the numbers are reliable?

These are the questions that this paper intends to answer. A biased sample equals a biased study There are two factors to consider when talking about bias in statistics. First is the sample bias, second is the selection bias. Sample Bias is created when a person accidentally models a study such that it omits evidence that contradicts one particular outcome (Craig 23). When you conduct a home survey on the preferences of mothers on the brand of milk they buy for their children, your objective is to get the opinion of mothers regarding what brand of milk they choose the most.

However, only making a home survey for the study presupposes that most if not all mothers are usually at home. This is not sensitive to the fact there are a lot of working mothers who also decide on what milk to get their children. Your study then becomes biased to the opinions of housewives which disenfranchise the opinions of working mothers. This is not to say that there will automatically be a difference in the results.

In fact, the biased variable which is in this case type of work (between housewife and all other occupations) may not at all be significant, but the fact that there is a bias creates the risk of unreliability which might not have been there if all mothers were given appropriate representation. One factual

example of sample bias was during the 1920s, United States presidential elections. A telephone survey was conducted to predict who would win the US election; the survey revealed that the Republican candidate would win by a huge majority.

However, he lost and it was found out later that since phones were at that time a lot more expensive relative to the common working man's wages, none of the poorer Democrat voters were able to participate in the telephone survey (Craig 31). Selection bias on the other hand, is created when a person deliberately omits evidence of a study that contradicts one particular outcome or purposely prevents certain groups from participating in the study due to a preconceived notion of how they would react (Craig 35).

Typically, studies that intend to commit selection bias create a screening system for participants in a study which would reject a certain people that probably won't produce the desired result. For example, if you want to make a survey on the favorability of " War on Iraq" and choose only hardliner republicans in your sample, then you are committing selection bias because you know that these people would vote favorably for the war. It doesn't matter how large your sample is, your study will most certainly be unreliable.

Diet programs like " Slim Fast" and " Fat burner" were questioned in 1995 regarding their survey taking methods. The promoters of these programs sent survey questionnaire regarding the program's effectiveness only to those people who were recorded to have actually lost weight in order to make it seem like that program was working for everyone (Craig 37). As a researcher, one must avoid sample bias by considering all possibilities or limiting the study dependent on only which possibilities can be covered.

If you are going to make a survey for high school students, you have to take into consideration that there are different groups in you sample, from the mainstream students in the public school system to the more minor groups who are in private school or are home schooled. If you cannot or do not intend to include these other groups, then you must specifically limit you study in the appropriate chapter. In order to not get fooled by sample or selective biases, a reader must always be mindful about the source that the study had taken respondents from. Spotting a biased sample is not very easy, especially if the bias is accidental.

More reliable surveys are supposed to break down their results from a broad scope to its specific subgroups. A good survey on " War on Iraq" favorability would include results of the whole sample first, and then show individual results depending on political leaning, race, gender and other relevant factors. Pitfalls in significance testing Significance testing is composed of a variety of tools like T-tests, F-tests, or P-tests in order to find out the significance of difference or relationship between two or more variables. Yet researchers are often misled into extensively using these tests and only these to prove their theses.

This dependency of scientific works to significance testing which resulted in statistical lies was exposed in the ' file drawer problem' presented by Rosenthal. The ' file drawer problem' reveals the existence of over-representation in published work of statistically significant results. Because those studies that fail to achieve statistically significant results are less likely to be published, researchers tend to repeat scientific experiments over and

over again until they get one that has a statistically significant result (Rosenthal 639).

It is impossible to know how many ' non-significant' results turned up before a ' significant' one was obtained and presented. Logically, if a researcher was able to find significant results using the same procedure only after several trials, the sample on that particular trial might have been biased, unless the researcher would perform the experiment a lot more times in order to establish the result better. This statistical lie is even more difficult to spot than sample bias; since the common reader would not be able to tell how many times an experiment was performed before good results were obtained.

One measure to take is by scrutinizing the source of the statistic and the study itself. Some Academic journals usually hold a high level of standard for accepting works. Reliable sources are usually sponsored by upstanding universities around the world. These universities have an interest in producing only the most credible, highly inspected researches as they have their reputation to protect. Selective reporting Choosing what to reveal among the many things that one finds out from a study blurs the truth and hides some consequences that might have made the reported results look inconsequential or even detrimental.

A researcher trying to prove the effectiveness of a particular drug may only publish the results that he found indicating effectiveness of the drug without indicating other results such as potentially harmful side effects. One can argue that ethically, a researcher is supposed to collect, summarize and report all the findings that he has found in his study, but personal bias may

get in the way of this ideal from actually happening. In 1978, Hallendroft, a drug manufacturing and research company in Germany was sued by consumers who claimed that the drug it manufactured for diabetics had adverse side effects.

During the cross-examination, some researchers admitted to having known the side-effect beforehand, but did not reveal it because of pressure from management to finish the work (Craig 112). Selective reporting of results is an unethical research practice, but like the previous two sources of statistical lies, it is pretty hard to determine. This type of statistical lie is more prevalent in the hard sciences such as medicine than in social sciences. In fact, medical journals constantly guard themselves from such kind of studies, often with peer-reviews and empirical studies of their own.

The importance of reliability, validity, and consistency checks In the realm of mass academic testing, several examinations are administered to different populations each year to determine academic, emotional, or psychological states of people world wide. International standards mandate that these tools be extensively checked for reliability, validity and consistency in order to make certain that the statistical analysis drawn from the results of these examinations can be trusted. Reliability testing aims to measure the extent to which a test is repeatable and is yielding of unbiased results.

Reliability tests can take several forms, each of which should be used extensively to determine an appropriate measure of reliability. One test involves giving the tool to a sample and then giving the tool to the same sample once more after a short period of time (usually a few days to a week). This test determines whether the respondents are not forced by the

questionnaire to simply guess answers, leading to random statistical results. A tool which was not tested for reliability is unpredictable. Any statistical analysis made from data coming from such a tool is unacceptable.

Validity testing makes sure that a test measures what it intends to measure. One type of validity testing is comparing the results obtained from a sample using the tool in question with results obtained from the same sample using another similar but already validated tool. For example, to test whether a new standardized achievement test in mathematics is valid; the results of a trial run may be compared with the respondents' academic standing in mathematics. A strong, significant correlation should indicate that the tool is valid.

A tool that reveals that a certain population is very gifted in mathematics may only have had very easy questions in the tool if validity testing was not performed. Consistency testing involves making certain that factors that should not be factors are not factors. It usually involves using a tool on two separate groups and determining the existence of bias in the questionnaire for a particular group. Using the same example as the math achievement test, such a tool that is written in English may be administered to native speakers and ESL (English as a Second Language) speakers.

A measure that indicates there being no significant difference means that the tool is consistent. An inconsistent tool may lead people to believe that ESL speakers are less mathematically competent than native English speakers, when in fact it is the language used in the tool that made it harder for the ESL respondents to answer. In spotting these statistical lies, it is

important that one verifies whether or not the required tests of reliability, validity, and consistency have been performed.

A tool for which these three were not established cannot be trusted to give sound results. Conclusion We therefore conclude that yes, you can lie and you can be lied to with statistics. It takes a critical mind to examine where the data is coming from, how the data was handled, and what the tools used on the data were in order to determine the truthfulness of statistical results, and even then we are not a hundred percent sure that nothing has slipped our notice, yet it is all we can do to stay vigilant.