Statistics in the High School

by *Dennis G. Haack*

The author is the section leader of the Biostatistics and Epidemiology Section of the Tobacco and Health Research Institute at the University of Kentucky. As a member of the National Council of Teachers of English (NCTE) Committee on Public Doublespeak, Dr. Haack writes and lectures on the misuse of the language of statistics.

The high school mathematics curriculum is continually changing. One of the more recent changes has been the inclusion of a course in statistics (see Pieters, 1976). As to the specific makeup of a high school statistics course, there is not likely to be agreement. As to the primary objective of such a course, there should be agreement. The purpose of this paper will be to look at the objective of a high school statistics course.

The key to the development of any course in statistics is deciding what statistics is. Statistics has, since the publication of R.A. Fisher's *Statistical Methods for the Research Worker* in 1925, been thought of as a set of research tools. In this regard statistics is the investigation of a population. The population of interest may exist or may be created by the researcher.

The study of an existing population is by a sample survey. A part of, or sample from a population is selected and studied. Examples of sample surveys include opinion polls, marketing research surveys, TV-viewing and radio-listening surveys, and pre-election polls. A 100 percent sample is referred to as a census. Of interest in the study of sample survey techniques is how a survey is designed as well as how to analyze and interpret survey data.

On the other hand, a researcher may wish to study a population which he creates. For example, an agricultural researcher might test a fertilizer on a crop which he has planted on a test plot. The researcher is stimulating the use of the fertilizer by farmers, that is, he tries to create a population which would exist if farmers used the fertilizer on their crops.

Another example of the investigation of a created population involves research on the effects of a drug. A population is created in the laboratory which would simulate use of the drug if the drug were put on the market.

As with the study of an existing population, the study of a created population involves a researcher with the design of his experiment as well as with the analysis and interpretation of experimental data. So we see that statistics is the study of a population which exists or is created. Statistics provide a set of tools which are required by an investigator for the design of a population study and the analysis and interpretation of the data generated by the study.

Traditionally, statistics courses at all levels have been an attempt to teach statistics as the study of a population. Distinction between experimental and sample survey investigations may or may not be made.

But statistics has become more than a research tool. Listen to the news this evening, or read a newspaper or a news magazine. Listen to public officials and advertisers. Statistics has become a language in its own right. This language pervades the media making it nearly impossible to understand a newscast without being quite familiar with the language of statistics. What are these words we hear - "estimates," "significance," "projections," "averages," et cetera? We are bombarded by numbers. But what do the numbers mean?

This is what statistics is to most Americans: a language which is very often used and too often misused. Statisticians have, for the most part, not taught about the language of statistics. Even students who have completed a traditional course in statistics cannot usually understand this language.

Statistics can be thought of as the tools required for the study of a population or statistics can be thought of as a language. We must decide which type of statistics we are to teach our students.

A first course in statistics should not try to teach statistics as a research tool. There are two main reasons for this. First, the study of statistics, a research tool, requires students to memorize the use of formulas, if not to memorize the formulas themselves. Students become so involved with learning to calculate statistics that they fail to learn what the statistics mean. Retention of the manipulative skills is minimal, causing students to have little, if any knowledge of statistics after a course of this type is completed.

A second reason why a first course in statistics should not teach statistics as a research tool is that students, after taking a traditional statistics course, are no better able to understand the statistics they'll encounter in the media than they were before the course started. The better students might be able to run a t-test, but they are not likely to have a feeling for what is involved with a determination that, say, significantly more animals in a treatment group developed cancer than did animals in a control group. Some of the students might be able to calculate the probability of selecting a red ball from an urn, but they may not know how to interpret the statement, "The probability of rain is 20 percent today." That is, the most we can hope of a student is that he or she will become a manipulative "whiz." A student might become quite good at "plugging and chugging": plugging numbers into a formula and chugging until a number results. Yet our students are not likely to be able to interpret the statistics they might have learned to calculate.

Statistics should be taught as a language rather than as a research tool. Students should first be taught how to interpret statistics. A student will be much better off being able to understand statistics than only able to calculate them.

Statistics can be taught as a language. It is being done at the University of Kentucky (see Haack, 1976). The idea behind the course is to downplay the calculation of statistics while concentrating on how to interpret statistics. In fact, students do not calculate any statistics in the course. There is, therefore, no need for mathematical formulas. The course is conducted in a strictly verbal, nonsymbolic manner. Examples used in the course come from the media. Ideally, students will be able to apply the principles they learn to statistics they will encounter, or have encountered in other areas.

One of the major drawbacks with a nonsymbolic statistics course has been the lack of a text, requiring a large amount of work by the teacher. Texts are now becoming available (see, for example, Haack, 1979).

One of the more interesting aspects of teaching statistics as a language is that students become genuinely excited about being able to detect misuses of statistics. When I started this experiment in teaching a few years ago, I did not look forward to trying to find examples of the misuse of statistics. Such examples are, of course, very instructive. As I began looking for cases of the misuse of statistics, I became awed by how easy examples were to find. I became more and more convinced that a course of this type was needed. Students also relish catching advertisers and public officials misusing statistics, that is, detecting doublespeak.

Doublespeak is the "involved, inflated, and often deliberately ambiguous use of language" (*Webster's New Collegiate Dictionary*) (see Rank, 1974, and Dieterich, 1976). The misuse of the language of statistics is statistical doublespeak. Statistical doublespeak can be avoided if statistics are properly understood (see Haack, 1977). This is the objective of the course I propose.

It is possible to teach statistics as a language. It is a challenging, yet rewarding undertaking. As you contemplate offering a course of this type, you might want to look at some of the books which can be used as reference material. There are a few good, readable books which may help you teach about statistics, the language.

With emphasis on sample surveys there are:

1. Gallup, G. *The Sophisticated Poll-Watchers Guide*. Princeton Opinion Press, 1972.
2. Roll, C.W., Jr. and A.H. Cantril. *Polls: Their Use and Misuse in Politics*. Basic Books, 1972.
3. Wheeler, M. *Lies, Damn Lies, and Statistics*. Liveright, 1976.

These books lack adequate discussion of the science of studying an existing population but do give a good discussion of the "art" of sample surveying.

On the general topic of statistics and statistical doublespeak consider:

1. Bross, I.D.J. *Scientific Strategies in Human Affairs: To Tell the Truth*. Exposition Press, 1957.
2. Campbell, S. *Flaws and Fallacies in Statistical Thinking*. Prentice Hall, 1974.
3. Federer, W.T. *Statistics and Society*. Dekker, 1973.
4. Hauser, P.M. *Social Statistics in Use*. Russell-Sage Foundation, 1975.
5. Huff, D. *How to Lie with Statistics*. Norton, 1954.
6. Messick, B.M. *Mathematical Thinking in Behavioral Sciences*. Readings from *Scientific American*. Freeman, 1968.
7. Mosteller, F. (editor) *Statistics By Example*. Addison-Wesley, 1973.
8. Reichard, R. *The Figure Finaglers*. McGraw-Hill, 1974.
9. Taner, J. (editor) *Statistics: A Guide to the Unknown*. Holden-Day, 1972.

You will find these books to be very interesting. Taner's collection of essays is an excellent source for the statistics course I propose. The essays are on the application of statistics in just about any area that students might have an interest.

References

Dieterich, D. (editor) *Teaching Public Doublespeak*. NCTE/Citation Press: Urbana, IL. 61801, 1976.

Haack, D.G. "A Nonsymbolic Statistics Course." *Communications in Statistics*, A(5) 10:943-47, 1976.

_____. "Statistical Doublespeak." *Kentucky Council of Teachers of English Bulletin*, 27: 1, Fall 1977.

_____. *Statistical Literacy: A Guide to Interpretation*. Duxbury Press: North Scituate, MA. 02060, 1979.

Pieters, R.S. "Statistics in the High School Curriculum." *American Statistician*, 30: 134-39, 1976.

Rank, H. (editor) *Language and Public Policy*. NCTE/Citation Press: Urbana, IL. 61801, 1974.

CREATIVE MATH EXPERIENCES FOR THE YOUNG CHILD
(Grades K-1)

- Forte, MacKenzie

Contents:

Learning about Shapes; Learning to Read and Write Numbers; Counting by Sets; Using Size Work; Finding Parts of Things; Learning about Measuring; Learning about Money.

The mastery of foundational concepts and understandings is developed in nature and should be presented to the child in proper sequence at the strategic time in keeping with his readiness for them in terms of his past experiences and presently recognized needs. This phase of the young child's intellectual development should never be left to chance, but planned and carefully monitored by an adult. It is important to remember that learning is continuous and developmental, and that each stage of mathematical growth is dependent upon success in the stage that preceded it. In keeping with this belief, the activities in *Creative Math Experiences for the Young Child* are sequentially planned and should be introduced to the child in the order of their presentation in the book. A test is included at the end of each section. Cost is $7.25; over 200 pages.

This book is also developed in three duplicating master books - Learning about Shapes and Sizes; Learning to Read and Write Numbers; Learning about Sets, Fractions, Measuring and Money. Each book costs $7.25; 36 masters in each.

Available at:

Western Educational Activities Ltd.
Box 3806
Edmonton, Alberta T5H 2S7

Publisher: Incentive Publications.