

SOCIOLOGICAL AND STATISTICAL SIGNIFICANCE IN EDUCATIONAL RESEARCH

By

ISIDORO PANLASIGUI *

I. INTRODUCTION

I wish to start this paper by briefly explaining the meaning of the two terms in the title "Sociological and Statistical Significance in Educational Research".

"Statistically significant" refers to the commonly accepted meaning in statistical language. It refers to the probability that an obtained measure deviates more or less from, or to, the true measure.

"Sociologically significant" refers to the sociological validity of experimental results (or any educational research) as bases for decision or prediction. For example, if an experiment yields statistically significant results, would the application of the conclusions based on these results yield sociologically significant results? Have the predicted and expected results desirable social values? For example, in a given education experiment, method of teaching A and method of teaching B are compared in order to find which method is better. The result of the experiment shows that A is statistically superior to B. Now, if method A is adapted as the method of teaching in a school system will the results be socially desirable in space and in time?

In this paper I also apply the term "sociologically significant" to any theory or belief whose validity is based only on traditions, customs, or authority of some individuals or a body of individuals without scientific basis. For example, during the middle ages the beliefs that the earth was flat and that the sun moved were at that time sociologically significant. But it is not so at present.

* Dean, College of Education, University of the Philippines, Quezon City.

SIGNIFICANCE IN EDUCATIONAL RESEARCH

These two terms — statistically significant and sociologically significant — suggest that a student of education should consider the necessity and the desirability that his decision or prediction are based on both the statistical significance of already obtained results and the sociological significance of the predicted results. The first is of accomplished fact, the second is yet to be accomplished. It is prophetic in its nature.

II. SOCIOLOGICALLY SIGNIFICANT

In the history of civilization we find many different beliefs and assumptions which were accepted as sociologically significant truth. Although statistical method, as we know it now, was not consciously applied to prove whether these beliefs and assumptions were statistically significant or not, leaders in those days consciously or unconsciously accepted them to be so. They were used as bases for decision or prediction.

As a matter of fact, even in this present age of scientific discoveries and technologies there are still a large number of people who, although scientifically oriented, tenaciously accept many beliefs and assumptions as both sociologically and statistically significant. Some college students in the department of Psychology, University of the Philippines, believe in superstitions. More than 60% of them believe in some superstitions like the following:

Primitive people can see and hear better than a civilized group of people.

The first and only child is usually a spoiled child.

Adults sometimes become feeble-minded from over study.

Children of cousins married to each other are practically certain to be of inferior intelligence.

In spite of the "revelations" through the voice of science there are still a large number of people who tacitly but tenaciously adhere to the assumptions, beliefs, etc. which are based on the astronomy of Jashua, and the geography of Dante.

These assumptions, superstitions, beliefs, etc. which have not been scientifically proved, and which have enjoyed sociological validity stemming only from traditions, are found not only in the areas of superstitions, moral and religion, but also in education. However, there were many beliefs in connection

with the educational theories and practices that had been discarded. These beliefs had enjoyed religious devotion of many teachers simply because they were considered sociologically significant.

Results of educational researches continue to reveal new principles and laws that may discard more of the traditional practices and ushered in new ones. I believe that in the near future our teachers may yet see some fallacies in the beliefs that certain constellations of subjects *per se* have the magic power to generate in the heart of man "sweetness and light" which transforms Caliban to Ariel or Circe to Penelope. This persistent belief has its origin in the time of the Greeks and it had received the benediction of Plato and Aristotle. Since then education must be so implemented. But if we look at the silver screen of history we can readily see that, in space and in time, in the whole range and area of education we find today as in the past, men and women whose hearts do not palpitate with "sweetness and light". Consequently, there are people who loudly lament over the fact that the moral and religious atmosphere of the present social universe is not conducive to the maintenance of spiritual life. And a world catastrophe is dangerously eminent.

The above observations shows clearly that there are men and women — educated or not — who readily accept beliefs, theories, etc. on purely sociological significance, that is, based only on the authority of customs and traditions.

III. STATISTICALLY SIGNIFICANT

Historically, among the first American educators who introduced and popularized the use of statistics in psychology and education was Dr. E. L. Thorndike of Teachers College, Columbia University, New York. His textbook, the "Mental and Social Measurements", was first published in 1904, when a large majority of American educators did not subscribe to his idea along the objective measurement of intelligence and education. However, from that time on the idea gained ascendancy and popularity. Dr. Thorndike gained the distinction of being considered the "Father of Objective Examination". The application of statistics greatly helped and accelerated achievements of scientific theories in psychology and education.

SIGNIFICANCE IN EDUCATIONAL RESEARCH

During the early years of scientific study of education, many students in education, with the new instrument of research (statistics) in their hand, started researches on many different aspects of education — content, methods, skills, theories, beliefs, practices, learnings, motivations, emotions, etc.

Many of these studies yielded conflicting and confusing results. This fact was due primarily to lack of experience on the part of the researchers in the technique of educational research and to the simple and elementary nature of statistics that were used. If one examines the published studies during the years before 1930, one can find that the *median* or the *mean* was the only statistical measure used for the analysis and interpretation of educational data. Occasionally, the correlation and the SD were also used. Today, statistics is more refined. There are now many students of education who have sufficient training and experience in the technique of research and in the use of more refined statistical techniques. Textbooks in statistics published after 1940 contain not only the measure of central tendency, dispersion, correlation and reliability, but also of other more refined and complicated measures or procedures such as experimental designs, the analysis of variance, factorial analysis, the levels of confidence, tests of significance, the method of random sampling, and the different forms of correlations, such as multiple and partial correlations. Of course some of these measures appeared in some old textbooks in educational statistics, but they were not taught in the classroom.

In this age of scientific investigations in psychology and education in which statistics play an important role, the students of education began to discard the socially significant criterion and swing to the extreme — accepting only the statistically significance as the valid one.

In other words, the use of statistical techniques, simple or refined, in education has led many students of education to believe that after the data of a study or research or an experiment are analyzed with statistical techniques and yielded statistically significant results, the conclusions derived from them are accepted accordingly and are used as guiding principles for new educational theory or method without inquiring

whether the predicted results will be sociologically significant or not. Thus we see that statistical techniques have created blind faith in the magic power of statistics to purify all kinds of ores into pure gold.

The point I am trying to drive home is this. Although the results of an educational investigation or experiment are statistically significant they may not be necessarily sociologically significant in space and in time. What is statistically true in America may not be so in the Philippines and vice versa. What is statistically true today may not be true in the future, ten or more years hence.

IV. STATISTICALLY AND SOCIOLOGICALLY SIGNIFICANT

Now, let us raise a question!

When is the result of an experiment or investigation statistically significant and yet not sociologically significant? The answer is briefly stated. The results of an experiment may be highly statistically significant as far as its purpose is concerned, but it may not be sociologically significant if the application of its conclusions does not yield desirable educational sociological values. To illustrate this point let us consider the following observations:

1. Illustrative examples. Mention has been made above that there are now very many different experiments in education and psychology. Among these experiments are found those that inquired into the relative effects of different means of motivation upon learning. These motivations were studied in the learnings of animals and school children. Results show that certain types of motivations are better than others. Among animals the order of effectivity is punishment and reward, punishment, and reward. Among the school children we find many different motives such as dramatization, knowledge of results, rivalry, etc. Although some results are conflicting, many of these studies yield statistically significant results. This means that upon these significant results we can confidently predict that if a teacher uses one of these means of motivation he will more likely get better result than if he does not use any sort of motivation. Because we are interested only in the increased amount of learned information, we can also confidentially predict that the application of any one of these means of motivation in teaching any subject in any school,

SIGNIFICANCE IN EDUCATIONAL RESEARCH

in any place, in any time, by any teacher we will most likely have the same result. This is so because we are concerned only in a simple event — the improvement of learning or accumulating information. In such a case the result of these experiments are statistically as well as sociologically significant.

The application of such results may be well considered sociologically significant for the simple reason that we are interested only in one particular event — namely, the increased amount that the school children can learn. However, when we raise the question as to the effect of this increased amount of learning, or the effect of certain method of instruction on the social life of the people in its relation to other groups of people within the country, its relation to other countries, its relation to international trends, and its effect upon the people in the future, I believe that statistically significant results are not enough decision-maker; are not enough bases for decision. We should weigh their sociological implications and effects before we make final decision.

2. The Iloilo Experiment. To illustrate this point let us consider the Iloilo experiment on the use of the vernacular, the Hiligaynon, and the English language. The purpose of the experiment is given below:

- a. As a medium of instruction, which of the two languages, English or Hiligaynon, was more effective in teaching Grade II children to read?
- b. Which of the two languages contributed more to the development of the ability of Grade II children to compute and solve arithmetic problems?
- c. What was the relative achievement in the social studies of Grade II children taught through English and of Grade II children taught through Hiligaynon?
- d. Which group of Grade II children learned language skills faster, the group taught through English or the group taught through Hiligaynon?
- e. What was the relative carry-over of the use of Hiligaynon to English and of the use of English to Hiligaynon?

There is no need here to describe in detail the experimental design. It suffices to say that the experiment had carefully followed the necessary precautions or control as much as possible of relevant and irrelevant elements in order to insure comparable results obtained from the control and experimental groups. The results or data yielded by the experiment were statistically analyzed. Comparing the means of the experimental and control groups in subjects taught gives very high statistically significant differences in favor of the vernacular. The critical ratio for each subjects is Arithmetic 6.23, Reading 8.20, Language 9.05 and Social Studies 10.9. The results of the experiment give the following conclusions:

1. That the local vernacular, Hiligaynon, is much more effective than English as medium of instruction in the first two primary grades. This particularly marked in the social studies, language, reading and arithmetic.

2. That there is a transfer of training in Reading from Hiligaynon to English and from English to Hiligaynon, the amount of carry-over in both cases being approximately the same.

Let us accept the results of the experiment as statistically significant. But just in passing significant observations on the experiment are in order for our consideration.

If we carefully analyze the bases for equating the two groups, we can readily see that they are not properly equated for the simple and obvious reason that the experimental group used the vernacular which is already the familiar language of the children and the control group used the English language which the children in the control group have to learn at the same time. The merit of experiment which compared the effectiveness of two languages is on the use of two languages which are foreign to both groups. As a matter of fact everybody knows that one can learn anything more effectively through his own language than through a foreign language.

While the critical ratios are all large and statistically significant, the differences themselves are very small: Reading 21.3%; Arithmetic 15.2%; Language 19.7% and Social Studies 23.6%. On the average, only around 17.56 of the total possible

SIGNIFICANCE IN EDUCATIONAL RESEARCH

scores in all the tests. This superiority of the experimental group is very small and is certainly due to the familiarity with the vernacular, not its use. One should wonder why the difference is so small!

If the purpose of the experiment is only to find which of the two languages, the Hiligaynon and the English, is a better medium of instruction, the results of the experiment are statistically significant. This statement is based on the assumption that the experimental design is as perfect as could possibly be done. If the purpose is only to teach the Filipino school children to build a good size of vocabulary in arithmetic, reading, social studies, etc., in their own language and make them fully conscious of their own particular dialect and province, and to prepare them for adulthood of limited provincial horizon then the use of the vernacular would be sociologically significant. But if the purpose is to teach Filipino school children to develop their minds for better and more effective thinking power, to teach them to be more conscious of the unity of the whole Filipino people and of their country as a whole and to prepare them to adulthood ready to help realize the objectives of all the people of the earth, the United Nations, the application of the vernacular would not be sociologically significant.

Mention should be made that if we, Filipinos, had only one language, and not many ("ten major native languages spoken by 85 percent of the Filipinos") and if we were concerned only with our people and country then the results of the experiment are socially valid. But even at that, if we are concerned with the trend of internationalism and the increasing interdependence of the different peoples of the world, and the fact that we are now learning and speaking the English language the results are not socially valid.

V. CONCLUDING REMARKS

Before closing this paper I wish to say that a statistically significant result in educational research may not be real but spurious because of many factors that cannot be controlled and whose effects upon the results are difficult to isolate. These factors may be grouped into (1) the mathematical theories, and (2) the methods of measurement.

1. Let us consider first the mathematical theories or facts in their applications to actual and real life activities. Let us consider only the simple arithmetical facts which clearly show that although a theory or a principle is perfectly correct its application may not be necessarily correct in actual practical life activities.

$2 = 2$ is theoretically correct in space and in time. But when we apply this mathematical fact, its validity, statistical or sociological, depends upon the nature of the variables upon which it is applied. For example: 2 inches = 2 inches may still be valid. But when we say 2 apples = 2 apples, we begin to question the validity of the statement from a practical or sociological equality. Its validity is obviously less than that of the first or second situations.

This reduced validity is due to the fact that there are no two apples which are exactly equal to another two apples. Here the equality of the variables lies only in the twoness. There is no equality among the apples. The two apples are not equal to the other two apples in their pairings. Any of the two apples may be two small ones, or one large and the other small. The four apples individually differ from each other in size, in color, in weight, in shape, in contour, in taste, in maturity, etc. There are many elements in an apple which are capable of different combinations and permutations that no one apple equals another one apple, much less two apples equals two apples. This difficulty of equality becomes more evident when we come to equality in human beings and their social relationships.

A hole in the ground can be dug by one man in six days. Theoretically, we can say that a similar hole can be dug in one day by six men. But psychologically it may not be so. The six men working together may be able to dig the hole in less or more than one day depending upon the skills, strength, and motivation of each one of them, the amount of work they put into the digging. It is also possible that the first man who dug the hole, may on second trial, be able to dig it quicker or slower, depending upon the effect of practice or fatigue. We may also consider the nature of the ground where the hole is dug. But of course, we can accept such a conclusion that the six men can dig the hole in one day.

SIGNIFICANCE IN EDUCATIONAL RESEARCH

Another illustration — one auto can run from Manila to Baguio in five hours. Therefore five autos can run from Manila to Baguio in one hour. This conclusion we cannot accept. We can readily see that educational researchers follow the same line of reasoning when they recommend the application of their conclusions. This is so because they forget that the elements which are relevant or irrelevant in the above illustrations are not as many and complicated as those found in educational researches. The sociological significance of results depend upon how well the very many different variables can be controlled.

2. Many students in education have obtained statistically significant results. But we know that some of them are conflicting. These conflicting results may be due to many different contributing factors, such as bad samplings, bad statistics, and bad logic. But assuming that the samplings, the statistical principles, the formula used, and the computations and logic are all correct and appropriate, still the statistically significant results may yet be spurious and invalid. The researchers in education should remember that although there are no errors in statistical principles, there are still many errors that creep in the measurement of learning, motivation, skills, habits, traits, methods of teaching, results of experiments, etc. because of the following reasons:

a. The absence of definite calibration in the instruments of educational measurements — tests or examinations. Two boys who took a spelling test got the same score of twenty do not necessarily have the same ability in spelling. Two boys who made the same score in History do not necessarily have the same ability in history. It is possible that a child who scored 10 in a vocabulary test may have a wider and better vocabulary than anyone who scored 15.

b. The changeable nature of human or social variables to be measured. Such changeability is especially true in intelligence, emotions, efforts, and attitudes which necessarily affect the result of measurement. These factors greatly affect measurement especially in the measurement of personality or character.

c. The complexity of the variable to be measured. The ability of a child to learn foreign vocabulary is not one single ability to measure but it is a hierarchy or combination of many different abilities consisting of sensory and motor abilities, attention, perception, discrimination, intelligence, etc.

d. This fact leads to the difficulty, if at all possible, of isolating one human or social trait, to be measured. There are some statistical techniques for such purpose, as partial correlation, factorial analysis, etc. But these techniques can not isolate an irreducible element or ability as well as chemical methods for isolating an irreducible element from a given substance.

e. The difficulty of controlling human, social and world condition for the purpose of experimental setting in order to obtain results which will be used as bases for decision and prediction.

In view of all these observations the researcher in education should look at his results and the application of his conclusion not only through the mathematical and statistical theories and principles but also through the rapidly changing social order for whose welfare the researcher is working. The student of education should apply his findings not for the sake of statistics but for the sake of the people.

