DOCUMENT RESUME

ED 047 984 SE 010 859

AUTHOR Finn, Jeremy D.
TITLE Updating Instruction in Statistics.
INSTITUTION Ontario Inst. for Studies in Education, Toronto.
PUB DATE Feb 71
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (Feb. 4-7, 1971, New York City, N.Y.)

EDRS PRICE EDRS Price MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS College Mathematics, *Instruction, *Mathematics Education, *Objectives, *Research Skills, *Statistics

ABSTRACT

This paper discusses objectives associated with the training of education students in statistical techniques. A historical study shows misconceptions in the areas of evaluation and measurement. Further, the typical preparation in quantitative methodology provides insufficient experience with analysis of problems and development of hypotheses. Objectives presented included identification of researchable problems, development of hypotheses, selection and application of quantitative techniques, and interpretation of the results. It is proposed that the instructional program should be viewed not as courses in statistics, but as a sequence of experiences in evaluation. Some suggestions are made regarding sources of textual materials. This paper was presented at the annual meeting of the American Educational Research Association, New York, February, 1971. [Not available in hardcopy due to marginal legibility of original document.] (RS)

Updating instruction in statistics

Jeremy D. Finn

Ontario Institute for Studies in Education

Abstract

Analysis of the current needs of students, contemporary demands upon social scientists, advances in statistics, and in the philosophy and psychology of education, yields two directions in which the teaching of statistics must move. 1) The integration of statistics with other aspects of the evaluation process, from the identification of research problems, through the interpretation of data, and 2) development of the appreciation of statistics as a refined and consistent symbolic system, offering a large variety of analysis options to the researcher. Suggested objectives and research findings concerning instructional techniques are provided, to aid achievement of "mastery" of the two broad goals.

UPDATING INSTRUCTION IN STATISTICS[1]

Jeremy D. Finn

Ontario Institute for Studies in Education[2]

In 1933, under the direction of Helen Walker, Mr. Ralph Brown completed and published his dissertation, Mathematical Difficulties of Students of Educational Statistics. Brown identified and analyzed the problems specific to statistics courses in Education. He concluded his thesis with a series of suggestions for alleviating the situation -- the imposition of mathematical prerequisites, the grouping of students by background experience, the elimination of certain topics, providing students with the option of computational drill and/or workbooks, and the standardization of statistical symbolism. Now that we have allowed 40 years for these suggestions to be successfully implemented, it is undoubtedly safe to assume that the problem is no longer with us, and is not worthy of further discussion.

In fact, the problems analyzed by Brown have probably not been significantly reduced over four decades. There has been a change however, in social scientists' dispositions about the role of quantitative techniques in the research process. I would like to analyze this role as it is viewed currently, and offer suggestions as to the implications for instruction in quantitative methodology.

---

[1]Paper presented at the annual meeting of the American Educational Research Association, New York, February, 1971. I am indebted to Merlin Wahlstrom and Joel Weiss for their reactions to an earlier draft of this paper.

[2]On leave from State University of New York at Buffalo.

What have been our "objectives" in teaching statistical methodology to students of educational processes? On one level, they are the objectives formally, or informally, developed by an instructor for a given class and semester. We assert that the students should "know, comprehend, apply, analyze, synthesize, and evaluate" the "specifics, ways and means of dealing with specifics, and universals and abstractions" of selected material from the statistical domain.

On another level however, we may ask why it is that we impose these "ends" on large numbers of students in Education programs. Here we discover some of our more basic assumptions. First, we have come to value research, and/or evaluation, as a means for gaining knowledge and making decisions essential to progress in education. And second, we have noted that the quantification of observations, and the treatment of quantified responses, is often a means to facilitate valid research or evaluation outcomes.

Let me describe what is meant here by the processes "research" and "evaluation." In the strictest sense, evaluation is the judgment of worth, value, or "goodness" of some object, event, or idea. Thus in Education we "evaluate" when we choose one set of objectives over another, when we choose a set of learning experiences for our pupils, and when we decide to study particular characteristics of pupils, teachers, environments, processes, and outcomes. That is, in so doing, we are delineating that which we think has sufficient value to warrant our extended efforts in thought and assessment. Further, we evaluate when we select a "sample" of observations for study, and when we select measuring instruments. [I would note parenthetically

however, that testing or "measurement" -- the assignment of numbers to individuals according to their responses to a given set of stimuli -- is not in itself evaluation]. Judgments of the relative or absolute adequacy of the results obtained, according to a set of criteria, and decisions based upon those results, are also evaluations.

The term evaluation does not in itself imply the use of quantitative procedures, nor does it even imply the use of accurate and reliable measurements. Stake (1970), offers an excellent introduction to thinking about less precise evaluations in his Review of Educational Research chapter, "Objectives, priorities, and other judgment data." Nevertheless, the evaluation paradigm with which we have greatest familiarity, involves the formal measurement of a sample of "subjects," and the subsequent mathematical summary and analysis of the outcomes. This special case of evaluation, usually termed "research," is more accurately termed "quantitative evaluation." Thus a distinction is also drawn with non-quantitative research, as in the use of resource or historical material to discover problem solutions.

One of the most concise and to-the-point treatments of the design and conduct of quantitative evaluation, is provided by W. T. Federer in the introductory chapter of Experimental Design (1955). The process may be conceived as having eight facets (adapted from Federer).

1. Formulation of questions to be answered.
2. A critical and logical analysis of the problems raised.
3. Formulation of expectations, or hypotheses.
4. Selection of a procedure for research.
5. Execution of the research procedures.
6. A complete analysis of the outcomes.

7. Interpretation of results in light of the hypotheses tested and the original questions.

8. Preparation of a complete, correct, and readable report of the evaluation (dissemination of findings).

Such an outline allows us to examine several common misconceptions. The first of these is that quantitative procedures are the only valid procedures for conducting "evaluation." The second, and one of perhaps greater concern, is the equating of measurement and statistics, with the entire evaluation process. A recent unpublished paper entitled "Training an educational researcher " (Timm), provides a program outline and reading list for the research trainee. The outline and list contain 4 books on "basic statistics," 4 on "correlation and regression analysis," 5 on "analysis of variance," 4 on "non-parametric statistics," 4 on "design of experiments," 1 on "survey sampling," 7 on "multivariate analysis," 4 on "matrix algebra," 4 on "measurement-scaling," and 4 on "factor analysis," suggests seminars on Bayesian statistics, sequential analyses, Monte Carlo methods, stochastic models, time series analysis, and power analysis, and concludes with the comment "in addition to the above, the student might also find it advantageous to pursue a substantive area of interest."

This outline surely is extreme in its neglect of six of the eight facets of evaluation listed by Federer. It is not inconsistent with the typical offering of two to three semesters of statistics to graduate Education students, and no formal instruction in the other, perhaps more crucial steps, of the evaluation process. As we would expect, the emphases are further reflected in attitudes carried beyond the graduate school years. To the extent that research in Education is not helpful in solving immediate and obviously-pressing

problems in today's schools, it is not our use or over-use of statistical methods which is at fault, but instead insufficient thought given to a) the problems worthy of consideration, and b) multiple hypotheses and educational outcomes.

What then should our objectives for courses in evaluation methodology include?

1. The student should develop an awarenoss of problems in Education, and be able to apply to them, criteria of problem significance.

2. The student should be able to identify the aspects of educational problems which are amenable to solution through quantitative evaluation. He should be able to state or restate such problems in basic, definable terms. He should be able to construct a "logical chain" between educational practice, theory about human behavior, and specific evaluation needs.

3. The student should be able to derive testable hypotheses from theory about human behavior and prior research.

4. The student should value quantitative techniques as means for obtaining valid and objective tests of his research hypotheses.

5. The student should be able to select appropriate hypothesis-testing devices, and to apply them to the particular situation. He should be able to select and utilize those descriptive devices which will clearly, and most simply, provide elucidation of the reasons why the major hypotheses are, or are not, supported.

6. The student should be able to incorporate both the quantitative outcomes, as well as the theoretical starting points, in deriving interpretations and generalizations from the analysis, for the original problem situation.

These are complex objectives, especially so when we realize that the "students" are not only our students, but ourselves as well. It is thus necessary to ask about structuring a program to promote attainment of these objectives. This is neither a small nor unworthy

task.

The instructional sequence should be viewed not as a course or two in statistics, but as a sequence of experiences in evaluation. Probably the first and most important unit involves the consideration of problems in Education, as drawn from the personal experiences of the participants, from the media, and from readings. Together with the instructor, the students should have direct individual and group observations of pupils in class situations, and of teachers and administrators. In addition, reference to reports of recent observational studies (Jackson, 1968; Smith and Geoffrey, 1968), can play a major role in the identification and formulation of significant researchable issues.

The question of problem significance is a most difficult one, and one which is often avoided. Class discussion of the relative importances of Education problems and outcomes, and of criteria for making such judgments, will at least sensitize each student to the need for consideration of the significance issue. While scientists often disclaim competence or interest in making such value judgments, the recent arguments of Scriven (1966, 1970), Pruzek, and others that we cannot afford not to consider the issue of significance, merit consideration.

The student should repeatedly be required to select aspects of the problem situation, to formalize them in researchable terms, and finally to state several hypotheses about the relationships among the major constructs. We often profess no prior knowledge from which to anticipate the outcomes of our studies. Frequently, probing of the investigator's mind and reference to the findings and theory of others, yield surprisingly strong expectations.

When the student encounters his courses in quantitative methods, and in particular, statistics, he will thus have a starting point from which to assess the utility of the techniques presented. Statistical techniques in common use are rapidly increasing in number and complexity. Further, the consideration of "real-life" and thus very complex problems, are likely to require complex models for analysis. To avoid occupying the student's entire graduate career with study of these models, we must increase our efficiency in their presentation. As an alternative to allocating one course unit each, to the topics "one-way fixed effects analysis of variance, two-way fixed analysis of variance, three-way fixed effects analysis of variance, one-way random effects analysis of variance, simple regression analysis, two-predictor regression analysis," and so on, it behooves us to consider such general models as the multivariate general linear (and nonlinear) model, general covariance structure models (Jöreskog, 1970), randomization models, general path analysis models, and others. From these, the students, in roughly the same period of time, can learn to construct all the simpler analyses as special cases, plus an entire set of more complex models, even beyond those presented by Hays (1963). Professors of such courses may, in turn, review the newer material at AERA presessions, and special workshops such as those offered at OISE and elsewhere. While this may seem a great deal of effort, we must ask whether the ultimate "pay-off" will make it worthwhile.

There is currently a materials problem as well. While there are few formal texts to assist us with the newer topics (Draper and Smith, 1967; Hendenhall, 1968), more will be soon forthcoming. In the interim, we may utilize less-formal publications usually available

from authors-to-be (e.g. Jöreskog, McDonald, Pruzek, Bock, Woodson, Timm, Finn). Semi-formal distribution agencies, such as The University of Chicago Statistical Laboratory, The Thurstone Psychometric Laboratory, and others, are likely to be of much assistance.

What about particular concepts within the statistics courses? First, elementary concepts of random variables, distributions of random variables, and descriptive functions, may be achieved through a "mastery learning" approach, as outlined by Bloom (1968). Constant feedback, student work groups, and a plethora of remedial material, have effectively promoted mastery in elementary quantitative methods courses.

In the more general discussions of model fitting, these themes should be repeatedly stressed:

1. The distinctions between "estimation" and "hypothesis testing." Both are essential.
2. The value of multivariate approaches, both in thinking about research problems, and in analysis.
3. The distinctions between experiments and "comparative studies," the disadvantages as well as the strengths of each.
4. The utility and strengths of planned hypotheses and related devices, such as the estimation of planned contrasts among model parameters.

Finally, provide the students with numerous opportunities to analyze real data, to probe the data for hidden meanings and findings, and especially to make recommendations for educational practice from the outcomes.

What I am suggesting then, is not an extended sequence of statistics courses for students of educational problems. It is instead an integration of quantitative analysis techniques into the

research and evaluation processes, at an early point in student's career. Research in Education can be "relevant," and quantitative techniques can play a part in making it so. In order to achieve this, we must expect much of ourselves, in terms of spending a significant amount of time in the systematic formulation of important Education problems. If we, and our students, view the ultimate outcome as worthy, they may expect and possibly attain a higher level of statistical competence than at present. In addition, we may be fortunate enough to achieve parallel attitude changes concerning the utility of evaluation. To maximize the likelihood of such changes, we may employ the achievement-supporting nature of 1) teacher interest and involvement, and 2) pupils' and teachers' expectations. We must also expect much of our colleagues. The students' attitudes reinforced from multiple sources are those most likely to develop.

By reconsidering the role of statistics in social sciences, I hope that we can eliminate discussions of the sort I overheard several months ago. A student in Education at an institution in some distant country, asked a professor to sit on his dissertaion committee. When queried about the topic or problem area; the student replied, "I haven't decided on a topic yet, but I'll probably do a two-way analysis of variance."

References

Bloom, B. S.. Learning for mastery. UCLA CSE Evaluation Comment, May, 1968, 1 (2), 1-12.

Brown, R. Mathematical difficulties of students of educational statistics. New York: Bureau of Publications, Teachers College, Columbia University, 1933.

Draper, N. R., & Smith, H. Applied regression analysis. New York: John Wiley & Sons, 1967.

Federer, W. T. Experimental design. New York: Macmillan, 1955.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1963.

Jackson, P. W. Life in classrooms. New York: Holt, Rinehart and Winston, 1968.

Jöreskog, K. G. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 1969, 34, 183-202.

Mendenhall, W. Introduction to linear models and the design and analysis of experiments. Belmont, California: Wadsworth, 1968.

Pruzek, R. M. The issue of objectives in research on compensatory education. State University of New York at Albany, mimeograph.

Scriven, M. Student values as educational objectives. Proceedings of the 1965 Invitational Conference on Testing Problems. Princeton: Educational Testing Service, 1966.

Scriven, M. The values of the academy. Review of Educational Research, October, 1970, 40 (4), 541-550.

Smith, L. M., & Geoffrey, W. The complexities of an urban classroom. New York: Holt, Rinehart and Winston, 1968.

Stake, R. E. Objectives, priorities, and other judgment data. Review of Educational Research, April, 1970, 40 (2), 181-212.

Timm, N. H. Training an educational researcher. University of Pittsburgh, mimeograph.