DOCUMENT RESUME

ED 047 007 TM 000 376

AUTHOR Mayo, Samuel T.
TITLE Trends in the Teaching of the First Course in Measurement.
INSTITUTION Loyola Univ., Chicago, Ill.
PUB DATE 70
NOTE 19p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Behavioral Objectives, *Education Courses, Education Majors, Evaluation, Innovation, *Measurement, Statistics, *Teacher Education Curriculum, *Teaching Techniques, *Tests

ABSTRACT

A survey regarding the content of measurement courses showed the most frequent topics to be item writing, item analogies, and choosing standardized tests. A complete list of the unique activities reported is appended. Another study revealed that seniors in teacher training showed but mediocre competency in measurement, not only at graduation but also two years later. This low performance may correlate with a lack of commitment to anything associated with measurement (such as mathematics, statistics, norms, scales, etc.) and could be improved by the introduction of innovative teaching in measurement courses. The innovative approaches and experiences of 3 instructors of an undergraduate measurement course at Loyola University are an example, while some findings of the Loyola Mastery Studies also have relevance. The most important point of these studies is that techniques have been discovered which yield not only mastery learning of content and concepts, but also favorable attitudes toward statistics. Perhaps these techniques are the most important trend in measurement teaching that can be reported. (CK)

Mayo's NCME Symposium Paper - 1970

Trends in the Teaching of the First Course in Measurement

Samuel T. Mayo

Loyola University, Chicago

The word "trends" in the title on this paper can cover a multitude of sins. You know how it is; at some point in time you are on the spot and you have to respond to the call for papers and come up with a proposed title of a paper you are going to write. Then, when you get around to the actual writing you find that the title is restrictive, so you have to take liberties. Then, you feel some obligation to the audience to explain the discrepancies. In the present case, I originally thought of this as a tracing of the history of measurement courses and a survey of the changes in content and methods of instruction. In actually implementing the paper, however, I found that it would not be feasible to do any intensive review of practices in past history. This did leave as an alternative approach a look at present status and perhaps a look a few years back. Again, there were limitations, due to lack of time and a shoestring budget, to obtaining either a large or a representative sample of institutions. However, a very open-ended questionnaire was mailed to some likely persons, about half of whom responded. Some replies were very lengthy and helpful. We also have some intensive data from instructors at Loyola.

What did our survey show? It would appear that most, if not all, courses include practice in item writing. Some include practice in analyzing items. It appears that frequently some attention to and practice in choosing standardized tests are included. It was not clear to what extent statistics is emphasized in measurement courses. It was not mentioned frequently, although one instructor said he de-emphasized it.

The handout lists all the unique activities which were given in the responses. Among these you will find some which are well-known and widely-used. There are others which seem to be genuine innovations. It is gratifying to see the innovations. My CRP study reported that measured competency in measurement was mediocre among not only graduating seniors in teacher training but among the same persons two years after graduation. The research report went on to speculate that a correlate of this low performance was probably a lack of commitment to anything associated with measurement (such as mathematics, statistics, norms,

scales, etc.). The report recommended that we professors should search for teaching innovations in measurement courses.

So much for course content and instructional methodology.

Another debatable issue is that of whether the first course should come at the undergraduate or graduate level, whether it should be required or be an elective, and whether it should be a single, survey course or a sequence of two or more courses.

Still another debatable issue is whether enough measurement competency can be developed in prospective teachers at the time they graduate with the bachelor's degree if their only exposure is as part of another course, such as the educational psychology or methods course. What it boils down to is this: Shall we send a beginning teacher out without knowledges and skills in evaluation and hope that he or she won't flounder too badly until he can come back to graduate school for proper training?

At this point I wish to share with you results of innovations which some of my colleagues at Loyola have tried in our first course. It is an undergraduate course, it is required for all elementary education majors and is an elective for majors in secondary fields, many of whom take it. I shall cite three instructors' experiences.

Instructor One

At the end of each chapter in Thorndike and Hagen is a list of Suggested Readings. It is from these lists, for the most part, that the students choose their articles. They are asked to present orally to the class the main points of the article and tie it in with similar material presented in the text - Thorndike and Hagen. The use of the blackboard and various types of audio-visual aids are emphatically encouraged.

The role the students should attempt to assume is that of a classroom teacher teaching information - with clear objectives for her

presentation and specific learning or reinforcing activities as she progresses with the lecture.

The student submits the objectives and several test questions relating to these objectives. These questions along with the objectives are pooled and an examination using these questions is given to the entire class.

Instructor Two

The first day of class the student is asked to make direct personal contact with a teacher in his subject area and ask permission to work with him. The student must listen to what the teacher says he is trying to do in his classroom and then formulate and write objectives for the class as he sees it. This procedure gives him the much needed practice or experience in interpreting and crystilizing vague objectives.

In addition to talking on a one-to-one basis with the teacher, the student is required to sit in on several class sessions in order to become

acquainted with the students and classroom procedure. The student has full responsibility for writing and administering all tests given in the classroom.

The student is then asked to evaluate his objectives in light of the examination data, make specific recommendations for changes in objectives as well as specific changes in test items themselves and then rewrite the examination and the objectives. He must also make a critical analysis of "old" exams written by the teacher and his own exams - using the same objectives.

Complete item analysis for every exam is required as well as specific changes in test items themselves and then rewrite the examination and the objectives. He must also make a critical analysis of "old" exams written by the teacher and his own exams - using the same objectives.

Complete item analysis for every exam is required as well as calculation of such statistics as mean, median, standard deviation, etc..

It is hoped the student will acquire skills and knowledges through this workshop or internship procedure that will have a greater "carry-over" value for him when he graduates from college and accepts the responsibilities of his own teaching assignment..

If a student finds it impossible to get a teacher for himself, a cooperative arrangement is made with a student teacher at Loyola.

Instructor Three

One large group project was assigned for each of two classes of 40 students each in our first course designed to give the students experience in teaching, test-making, test-scoring and grading, and test analysis. Since the classes were large and since only 6 periods were available for presentations, each group consisted of 6 or more students. Each group presented to the class the characteristics, evaluation etc. of one general type of test (such as intelligence tests, high school achievement test, personality inventories, etc.) and an example of a standardized test of this type

in current use. Then, following the presentation, the group administered a short (12-15 items) objective achievement test on what had been presented to the class as a whole. At the beginning of the next class period, the group that had presented the period before returned the test papers to the class and went over each item, providing feedback. Then, within two class periods after the feedback had been given the group turned into the teacher, a basic analysis of the score distribution (showing the most appropriate measure of central tendency and most appropriate measure of variability for that distribution as well as a histogram of the distribution) and an item analysis (index of difficulty and index of discrimination for each item) of the test. Then they submitted a grading system and critical review of the test in terms of the results they had achieved with the class.

Many unexpected outcomes were achieved from this project. To begin with, all 12 presentations were far superior to what the teacher had anticipated.

Students had done a great deal of research and had difficulty in limiting themselves to the time available for the presentation. Most groups supplied the class with handouts ranging from one-page outlines of the talks to 8-page, bound notebooks giving an outline of the talk, a paper on professional criticisms of the standardized test under scrutiny, a cartoon to emphasize a point about the test, and a bibliography of reading about the test type and the specific standardized test in question. Eight of the groups also illustrated the talk with excellent visuals. Most of the visuals were in the form of large, expertly drawn examples of test questions from the test under discussion.

Students from two of the groups took the MMPI and Kuder so that they could speak about them from the point of view of the testee as well as from the point of view of the test administrator and test critic. These tests were administered to them by the Psychology Department.

The students who made the actual presentation were well rehearsed. Ten of the twelve groups actually had the presenters practice before the group before making the presentation before the class.

The tests were excellent and imaginative. Not all of the questions managed to "come off", but they were all worth trying. The item analyses were well done as were all other parts of the project.

The teacher graded the presentations and other parts of the project and gave the groups feedback orally instead of in writing with the hope that she might get more feedback about the projects from the students in turn. This system worked well. It was at such times that the teacher discovered much about how the group had worked.

It seems that most of the groups determined what was to be presented only after heated arguments about what was important and or

significant. The items themselves were also a subject of debate (such as whether items were too easy or two difficult, etc.); all of the groups verbalized their surprise at discovering that the items they thought were easy were missed by many of the class and those they thought were difficult were passed by most of the class. It was interesting to hear and see the students verbalize on this point. They explained this to the teacher as though they were the first persons in the world to have discovered that the test writer cannot determine in advance how a class will perform with his items. Of course, this point had been made earlier in the course in the lecture, but not until the students discovered it themselves did it seem to have meaning. The item analysis became very important to the students because of this discovery of theirs. Moreover, it was at this point that many students (at least 8 groups) requested the teacher to recommend yet more books on the subject of writing items. The teacher recommended Robert Ebel's Measuring Educational Achievement for this purpose.

Students did not always agree on the criticism of the test - that is on what the statistical analysis showed about the test. She, therefore, suggested that those who did not agree with the majority opinion could, if they so desired, submit a minority report. Five of the twelve groups did so.

The vast majority of students had had no college mathematics and had been resistent to the statistics of the course. But when they discovered through the project how statistics could help them understand their tests and their classes, they became more interested and developed a need to know which seemed to account for their improved achievement on the final examination as compared with their achievement in statistics on the midterm. The teacher found herself in discussion concerning some simple elements of inferential statistics with such students.

In general, it seems that the students (at least 90% of them) devoted much effort and time to the project and thus, became more interested in it

than had been anticipated. Because of the intense effort of one group, the presentation ran over into another session for which most of the students stayed. Another asked and received permission from the teacher to give a pre-test so that they would better know what material they could dispense with in the presentation and what they should concentrate on. These students know little if anything about the function of a pretest but reasoned that such a test would be helpful to them and to the students.

One group rearranged the chairs of the class in a semicircle of two rows so that the presenter could be "closer" to the students and be sure that no one missed anything. This group explained to the teacher that intelligence tests were very important and that the class needed to know what would be presented, again as if the teacher had not known this.

The MMPI, which was chosen by groups from both classes, provided another type of learning experience. Both groups criticized this test extensively not only

from the point of view of test theory, but also from the point of view of ethics. It was here that the real realization of the importance of the evaluator of tests results was seen. The students made a more eloquent plea for professionalism in evaluating test results than the teacher had previously and the students in both classes seemed to grasp the significance of what was being said.

All in all, this experience - of making the presentation and doing the subsequent work - seemed to be the best teacher. What had gone before in the way of lectures, class discussions, and reading of the assignments, seemed to form a base not of learning but of reference. Only after the students had the experience afforded them by the project did they really seem to grasp the full implication of what they had been exposed to.

For the most part, things went smoothly for Instructor Three. However, I can't resist telling two anecdotes.

Anedcotes from Instructor Three

Anecdote No. 1

For reasons which we have sti-l not explained one student responed to a completion item on a quiz, in which the keyed answer was "the bell-shaped curve", with her own unique response "the well-shaped curve." (Pause for laughter) Now, you have got to admit that one could psychologize by the hour about what this situation means and implies.

Anecdote No. 2

Here Instructor Three made an innovation which occurred spontaneously. One obviously very bright Political Science major goof e d off throught the entire course and bombed (I think that's the word. Like, man, you know, like he really
the final.) Then he came and asked the Instructor to increase his letter grade from an F to a D. The instructor was wise enough to agree to this only if he retook an exam over the course material but this time an oral exam. He was given two weeks to prepare. As it turned out he performed fantastically well on the oral. As a matter of fact he elaborated at great length on some of the questions about applications of tests or problem

situations. He even showed enthusiasm for the material. He admitted he had crammed hard, had studied the text intensively, and had used fellow students to tutor him. Incidentally this is a beautiful example of mastery learning and with our N = 1 shows how the amount of study time can be treated as a variable. It also illustrates an idea of Edward Thorndike's that the best time for learning something is when you need to learn it.

Loyola Mastery Studies

Some studies of teaching and learning in the first course in statistics at Loyola have relevance here.

Instructor Three and the present author reported a preliminary study at last year's NCME meeting. An abstract of this study appears in Ben Bloom's annotated bibliography of mastery learning studies.

Briefly, our preliminary study demonstrated that application of the Bloom-Carroll Mastery Model to our first statistics course with streamlined objectives did in fact yield not only mastery learning of content and concepts but favorable attitudes toward statistics on the part of all the students, the like of which we had never experienced previously.

We believe that we have discovered some techniques to achieve some of the challenges laid down in the CRP report toward getting students

strongly committed toward recognizing, attending to and using measurement in relation to the other aspects of their life and work. I believe that this is the most important trend I have to report.

Handout for Mayo's Symposium Paper (NCME 1970)
Activities Reported by Instructors

Item writing

Item analysis

Constructing tests - essay, multiple choice, etc.

Administering tests

Interpretation of a given test

Computer generated "scores" for each student to practice on

Reviewing and criticizing a particular standardized test with the accompanying manual

Bibliography prepared, by instructor, on measurement in each area of the curriculum - student reads articles and books in his own specialty or subject area (1) writes reports containing statement of objectives (2) specific techniques that could be used to measure the objectives (3) lists standardized tests available for different purposes and writes a critique

Lab manuals following lectures with exercises and answer key so student can check himself out

Meeting with students to discuss social criticisms of testing

Student writes critiques of specific tests and test users

Designing a testing program for a particular purpose or set of purposes

Promoting favorable attitudes by stressing the value of tests in teaching and counseling (focus on uses)

Statistical concepts only as needed (at a minimum)

Use of colored chalk to differentiate concepts in statistics

Use of materials and handouts from test publishers

Test source and evaluation - use of Buros, judgments of usability of tests based on validity and reliability

Test blueprints - writing objectives

Administering multiple choice tests to students throughout course and providing diagnostic feedback

Examinations are open book and open notes, and sometimes done at home

Visual aids - film strips, transparancies

Oral reports

trips and Internships