DOCUMENT RESUME

ED 405 028 JC 970 134

AUTHOR	Fago, George C.
TITLE	Psychometric Scales: Learning by Doing.
PUB DATE	Mar 96
NOTE	14p.; In: Proceedings of the Annual Conference on Undergraduate Teaching of Psychology (10th, Ellenville, NY, March 20-22, 1996); see JC 970 128.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Active Learning; *Computer Uses in Education; Electronic Mail; Higher Education; Instructional Innovation; *Psychological Testing; Psychology; *Psychometrics; Research Methodology; Scaling; Teaching Methods; *Test Construction
IDENTIFIERS	Ursinus College PA

ABSTRACT

In an effort to improve instruction related to psychometrics and psychological testing for beginning psychology majors, an approach was developed at Pennsylvania's Ursinus College that incorporates active, participatory learning and problem solving. The approach is used in a course on research methods and makes use of a networked computer system that includes electronic mail capabilities. The course covers non-experimental research methods in three units: field research, survey research, and correlational research. Instruction in psychometric scaling is provided during the final unit of the course, in which students are asked to design a correlational study that includes several psychometric scales. While students may incorporate existing scales into the study, they are also required to develop their own scales. The use of computer technology and electronic mail helps eliminate much of the time-intensive labor involved in scale development, with students submitting and rating test items electronically. The following steps are used in developing the psychometric scales: (1) the class selects a concept to be scaled; (2) each student submits five potential items for the scale; (3) the item pool is circulated via electronic mail and students rate each item; (4) the instructor prepares a statistical description of the items; and (5) students select the final items. An outline of the steps in preparing a scale is appended. (TGI)

Psychometric Scales: Learning by Doing

George C. Fago

Ursinus College, PA

Proceedings of the Annual Conference on Undergraduate Teaching of Psychology (10th, Ellenville, NY, March 20-22, 1996)

PSYCHOMETRIC SCALES: LEARNING BY DOING

George C. Fago, Ph.D.

Ursinus College

Psychometric scales or more popularly "psychological tests" are typically of intense interest to the beginning psychology major. Regardless of whether the student is interested in the more biological end of the field or, in contrast, the applied aspects of the field, psychological testing remains 'mysterious,' 'arcane,' and for some a topic for skepticism and/or uneasiness. Regardless of how they 'feel' about the tests, most students are fascinated by them. Yet, teaching students about psychometrics and psychological scales can be a daunting problem. I remember, without even a shred of nostalgia, both an undergraduate course and a graduate course which dealt with the topic. My recollection was that a 'good lecture' was one which was merely profoundly boring, whereas a 'bad lecture' was one which seemed pointless if not deliberately obscure. The problem of course is that the instruments typically lack a high degree of face validity to begin with. To compound the problem the topic is often approached without any reference to concrete examples at all. The highly abstract approach is almost certainly doomed to failure, while using examples with poor face validity is not a great improvement. Given a lack of meaningful context, presenting the steps used to generate the items often leaves students with no appreciation of the reasoning lying behind the steps. In an attempt to rectify this problem I developed a method of instruction that makes the process for the students as concrete as possible. What follows is the approach I use to teach novice psychology majors the basics of scaling in a manner which demands active, participatory learning and problem solving on their part. In doing so it appears to foster much greater depth and appreciation in their level of understanding of the concepts, as well one which reinforces other concepts dealt with in the course.

B: SETTING

It should be noted first of all that for the method to work there must be access available to a networked computer system. At Ursinus all computer labs and faculty offices are tied into a common network. Further, students may access this network from their dorms if they have their own personal computers. The instructional program to be described not only assumes a computer network, it is also designed specifically for a particular student body within a particular curriculum. Thus the context of the instruction is important for purposes of evaluation; although I feel that aspects of this approach are of broad applicability, some description of the students for whom it is intended is necessary. Ursinus is a private, residential liberal arts college located in suburban Philadelphia. The majority of the student body is drawn from the Delaware Valley. The average combined SAT score is approximately 1200.

Instruction in scaling takes place within the context of a course called Research Methods. The course is required of all student majors and minors, and is designed to be taken as the second course in the major, immediately following the completion of the general introductory course. Thus the majority of the students are sophomores, although twenty to twenty-percent usually are upper classmen. The course is conceptualized as an introduction to research methodology and assumes no prior familiarity with statistics or computers. Non-experimental research methods are covered in the course, which is divided into three conceptual units: field research, survey research, and correlational research. Experimental methodology is covered in another course.

As the three types of non-experimental research methodology are presented, they are used as starting points for consideration of other aspects of research. Topics considered during the course include the following: report writing, reliability and validity issues, research ethics, nature of the scientific method, measurement scales, operationalism, sampling, and descriptive and elementary inferential statistics. The course meets for four hours per week, three hours of which are lecture with the remaining hour utilized for laboratory work.

Each of the three approaches to research constitutes about a third of the course work. In the course of instruction in each of these methods, students are introduced to both topics of general importance and topics of particular importance for the method being considered. As part of this instruction for each method students design and execute a complete study; they generate a hypothesis and a fully operationalized research design, collect data, analyze it, and prepare a report in APA format.

The accompanying one hour lab is scheduled in a general purpose computer lab area, with a computer for each student. The computers are part of a larger campus network. This allows access to SPSS Student Ware (the software package used in the course), as well as local e-mail capability. Laboratory instruction focuses on data analysis by means of statistical software, as well as research design and instrument development.

C: TEACHING PSYCHOMETRIC SCALING

Instruction in psychometric scaling is included in the final unit, the correlational method. The students' task in this unit is to design a correlational study in which they administer several psychometric scales in order to examine possible relationships between the factors or variables being measured. For the purpose of this study students have access to existing scales (*e.g.* Locus of Control, Beck Depression Inventory, Bem Androgyny Scale) which they may incorporate into the research. They must, however, develop and include their own psychometric scales as well. In point of fact, I require both as a means of dealing with the concept of convergent validity.

In my initial attempts to teach this material, I did not ask students to develop their own scales. It quickly became apparent to me, however, that for students at this level presentation of material regarding scaling led to rote learning with little or no appreciation for the design issues, much less the conceptual underpinning of the various methods of scale development. Good students were able to tell me the exact steps to be used in generating a Thurstone Equal Interval Appearing Scale, but showed little understanding of why these steps were necessary or how the methodology related to a scale such as the Locus of Control Scale. For this reason it seemed to me to be desirable to involve students in every step of scale development from operationalizing a concept to final item selection.

My reasons for this decision are not complex. The emphasis on implementing instruction that employs active participatory learning in the past decade or so has been pervasive, and you are all familiar with the arguments. None-the-less, teaching scaling to new majors as part of their first course in research methodology offers a challenge for which a particularly compelling argument may be made for active learning. Students with a very limited exposure to psychology are still operating cognitively in a very concrete fashion, often with more than a few misconceptions about what they are doing and why they are doing it. Although the reasoning underlying a Thurstone scale may seem transparent to the more sophisticated and advanced student, when it is presented to the novice in an abstract fashion, particularly if it is divorced from context, it may be largely inaccessible cognitively necessitating rote learning on the part of even very good students.

However, when students move through each step of the process with the concrete goal of operationalizing or scaling a particular concept, the entire task becomes cognitively accessible precisely because it is concrete. Particularly if there has been prior discussion of reliability and validity issues, I find that students are extremely acute in their questioning of the validity of potential items. If they are engaged in dealing with the problem of selecting a group of items with the best possible reliability and validity characteristics and they have this 'set' from the beginning, then the logic behind the development process becomes much more apparent.

X

In my initial attempts to carry this through I had students design a Thurstone type scale. You will remember that a Thurstone type scale consists of a small number of items (typically 10 or less) each item of which corresponds as closely as possible to a scaled integer value of the concept. That is, given a concept scaled from one to ten, there will be ten scale items each of which is as close as possible to an integer value, the best possible '1' item, etc. I reasoned that if students could develop an understanding of this approach they would be able to easily transfer their understanding to other scale types as well. Informal observation at the time indicated that this was the case. Student understanding of the conceptual issues underlying scale development increased dramatically as a result of this exercise, at least as indicated by their test performance and research reports.

Such gains are not without cost of course. Standard text book presentations of scale development uniformly stress how time consuming and difficult the task is, and not without reason. It is arduous for a single investigator; to coordinate the activities of twenty student 'investigators' adds one more level of demand to an already arduous task. To call this instruction labor intensive is a conservative description.

Remember that to construct a Thurstone type scale the following steps are necessary. A large pool of potential items must be generated; the accepted minimum is 100 items. Each of the items must then be rated by each of a minimum of twenty raters. The resulting ratings are then incorporated into a data file, and means and standard deviations are obtained for each item. These statistics can then be used for item selection. What makes this feasible at all as part of the course work is that the students generate the items and serve as raters. For the instructor this is an obvious savings, but in turn the instructor is faced with the considerable problem of collecting and collating materials, distributing them, collecting ratings, analyzing them, and redistributing the rated items. My estimate is that construction of each scale demanded approximately eight to ten extra-instructional hours of time. This time is spent in editing, typing, reproduction, *etc.* At least some of this can easily be handled by student assistants, but because of time demands, not infrequently I found that I had to do it myself.

While I was convinced that the exercise was an extremely valuable one from a pedagogical viewpoint, the time demands raised questions about the cost benefit ratio. For that reason I began to deliberately ask how the task might be accomplished more efficiently with little or no sacrifice in instructional quality. What follows is my solution to the problem.

C. COMPUTER ASSISTED SCALE DEVELOPMENT

Scale development takes place over the course of approximately three weeks, although it does not require more that two hours of instructional time to accomplish. It requires a number of sequential steps as outlined below. The task remains the same, with the time savings being effected by utilization of readily available computer technology.

1. Selection of a concept to be scaled.

In the course of a regular class meeting students chose a concept to be scaled, and are given instruction in preparing potential scale items.

2. Item Pool Generation.

Prior to the weekly lab students are told to individually prepare five items apiece and to bring these potential items to the lab. During this first lab meeting, further material is presented regarding item preparation and selection, and then students in small groups critique each others items checking for clarity, face validity, spelling, etc. Following the critique, students do any necessary revision of their individual items. Once they are satisfied with their items, they then use campus computer network to e-mail the items to me.

Previously students handed in handwritten items which were collated and typed by a student assistant. This typically took approximately two hours. By having the items e-mailed, much of this labor is eliminated. One of my lab assistants simply works from my office. The assistant downloads the individual student item text files, strips the headings, and combines them into a single text file. This process typically takes no more than a half hour.

3. Item ratings.

Within a day after the lab, the items are circulated so that the students can each rate them as they were instructed to do. Circulation is again by e-mail. The item pool also contains specific instructions as to how the resulting ratings are to be returned. Specifically, each student is to e-mail me a message with the body consisting solely of a 10 by 10 data matrix (assuming 100 items to be rated), with each cell representing an individual rating. The lab assistant again downloads the resulting messages, strips the headings, and incorporates the ratings into a single data file. Again this process requires approximately one half hour, as opposed to the two or more hours that it took to create the data file by hand.

4. Statistical Description of the Items.

Currently I not only create the system file, but also actually carry out the data analysis. Conceptually, there is no reason why the students could not do this themselves since at this point they are very familiar with the necessary SPSS procedures. But because they are using the Studentware version of SPSS which is limited to data files with no more than 20 variables this is not currently feasible. I hope to soon be able to use a full scale version of SPSS which will allow this step to be done by students as well.

5. Item selection.

The resulting analysis is circulated to students. The lab assistant simply prepares a text file containing the mean and standard deviation for each item which is then uploaded and e-mailed to the students prior to next lab. During the lab itself, students are given a short lecture on the criteria for item selection. Then each student individually goes through the item pool selecting the items they deem most appropriate for scale inclusion. The results of this are shared, and form a basis for a concluding discussion of the subjectivity inherent in the methodology.

I estimate that the total amount of time spent on the exercise is now no more that two hours to two and a half hours with only about a half hour of my personal time needed. The majority of time saved was formerly spent on collating and typing materials, duplicating, and distributing. By using the e-mail facility all of this is eliminated. Students do much of their own editing, and no typing or other preparation is required. As a result I have expanded the exercise such that students now prepare two scales: both the original Thurstone type scale and now a Likert scale as well. I find that this can typically be accomplished in less than half the amount of time it formerly took for one scale, and with no more that an hour of my personal attention.

D. CONCLUSIONS

There is nothing inherently new in the process I have described above. The process of scale development remains unchanged. And I certainly can take no credit for the 'participatory learning' approach. What is new is using computer technology in a somewhat novel way. What enables the approach described is e-mail. Like my students, I had thought of e-mail primarily in terms of corresponding with individuals at distant sites. Then I became accustomed to communicating with colleagues on campus using e-mail. A year ago, I realized that it had also become entirely feasible to communicate with students using e-mail as well. This realization was forced on my, when I began to wonder why students were hanging around the lab after handing in their assignments. Looking over shoulders established that they were using the e-mail capabilities. I therefore began to issue electronic syllabi (easily updatable), send class announcements and updates, post new assignments and reminders, *etc.* using e-mail. From this it was only a short conceptual step to using e-mail as a communications tool within the laboratory context, thus distributing the work load across the students themselves and freeing me and my assistants to do other tasks. My only claim to originality is a mildly innovative use for an established computer utility.

STEPS IN PREPARING A PSYCHOMETRIC SCALE

STEP 1

1. Choose concept to be scaled
2. Instructions in item preparation

STEP 2

1. Prepare potential items
2. Critique items in lab groups
3. E-mail items to instructor

STEP 3

1. Assistant downloads items, stripping headers
2. Combines items into master file
3. Master file e-mailed to students

STEP 4

1. Students rate each item in master file
2. Students e-mail ratings to instructor
3. Assistant downloads ratings, stripping headers
4. Ratings combined into master data file

STEP 5

1. Item means and standard deviations calculated
2. Item means and standard deviations compiled and e-mailed to students.
3. Students develop individual scale

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Teaching of Psychology: Ideas and Innovations. Proceedings of the Tenth Annual Conference, March 20 - 22, 1996.	
Author(s): Judith R. Levine & Gene Indenbaum, Editors	
Corporate Source:	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

☐ Sample sticker to be affixed to document

Check here

Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document ☐

or here

Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Judith R. Levine	Position: Associate Professor
Printed Name: Judith R. Levine	Organization: SUNY Farmingdale
Address: Psychology Dept, SUNY Farmingdale, Farmingdale NY 11735	Telephone Number: (516) 420-2725
	Date: Jan 26 1997