DOCUMENT RESUME

ED 098 266 95 TM 004 065

AUTHOR Baker, Eva L.
TITLE Practicum for Instructional Developers. Final Report.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
BUREAU NO R020520-FR
PUB DATE Jun 74
GRANT OEG-0-72-4329
NOTE 73p.

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS Checklists; *Formative Evaluation; Glossaries; Instructional Improvement; *Instructional Materials; *Instructional Programs; Post Testing; Pretesting; *Program Administration; Questionnaires

ABSTRACT

The report describes the development of a project which produced an integrated set of materials on formative evaluation that are appropriate for use in graduate training, research and development settings, and public school personnel use. The objectives of the project were: (1) to identify relevant data sources for formative evaluation and to outline a plan that incorporated appropriate data sources; (2) to be able to display data from a variety of sources, to summarize information in nonstatistical ways, and to draw inferences about program components that require revision; and (3) to suggest revisions of instructional materials when provided with a program description and an unanalyzed data set. The materials were developed in the course of eight tryout and revision cycles and now meet the expectations of their designers. (Author)

Final Report

Project No. RO20520
Grant No. OEG-0-72-4329

PRACTICUM FOR INSTRUCTIONAL DEVELOPERS

Eva L. Baker

Department of Education
University of California
405 Hilgard Avenue
Los Angeles, California 90024

June 1974

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education

National Institute for Education

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The set of materials, evolved as Evaluating Instructional Programs, was developed under the broader funding of "Practicum for Instructional Developers" (Project No. RO20520, Grant No. OEG-0-72-4329). The purpose of these materials was to prepare individuals in the competent trial and revision of newly developed instructional products. The materials were planned in a largely self-instructional format to foster their eventual dissemination. They were designed for use by graduate students, teachers and curriculum specialists who were to assume responsibility for formative evaluation of instruction. Ideally, the materials would articulate with Research-Based Techniques for Instructional Design (Baker and Quellmalz, 1972), developed under the auspices of the national Center for Educational Research and Development, U.S. Office of Education.

What follows will be an unelaborated treatment of the development history of Evaluating Instructional Programs, including the decisions made and options foregone as the product came into being. While there is always faint hope that a process in retrospect will serve as a positive example for future developers, this report will attempt to include the significant aspects of the development activity, both good and bad.

The need for a set of materials to teach formative evaluation techniques was apparent at the time of proposal and continues to this time largely unsatisfied by efforts elsewhere. While there have been rules and guidelines produced that suggest data gathering and analysis techniques, they tend to be inadequate for a number of serious reasons:

1. They appear to savor dealing technically with problems that might be treated more simply.
2. They do not highlight data gathering for the purpose of improving instruction but rather focus on the full description of what a program does.
3. They often do not fall within an established frame of reference for development activity.
4. None have been subjected to empirical tryout and revision processes themselves.

The designer of the materials also had a long history of teaching courses in instructional development and, from that experience, could anticipate the class of problems of concern to novices in the area. Thus, the materials in the early planning stage were designed to meet the following sets of needs:

1. To present clearly and simply information that would help individuals plan formative evaluation trials.

2. To emphasize that data gathering should be instrumental to the improvement of instruction and not a rite in and of itself.

3. To build in opportunities for individuals to work with simulated data.

4. To integrate information about revision of programs with the data collection/analysis activity.

5. To present problems that had a strong flavor of verisimilitude in an attempt to reduce the abstract and unreal character that "practice" exercises often share.

6. To design materials that subjects will not mind using, by incorporating the reactions of subjects to the materials in the revision process.

The objectives of the program were directly formulated and remarkably stable throughout the development process. They were as follows:

After completing the materials, the learner would be able to:

- Identify relevant data sources for formative evaluation and to outline a plan that incorporated appropriate data sources.

- Display data from a variety of sources, to summarize information in non-statistical ways and to draw inferences about program components that require revision.

- Suggest revisions of instructional materials when provided with a program description and an unanalyzed data set.

The major concepts treated in the program were as follows:

1. Collect the least amount of data required for improving instruction limiting the sample size and the investment in slick materials at early stages.

2. Attend to more than prespecified cognitive outcomes.

3. Display data visually rather than only summarize statistically, to foster understanding of what is happening in the program.

4. Revise programs consistent with the extant data base in research on instruction: focusing first on the use of instructional principles, second, on alternative methods, and last, on alternative media.

This recommended sequence also is scheduled in terms of increasing cost.

DEVELOPMENT PLAN

Although this project was planned to be pursued independently of the major "Practicum" effort, the interaction of the two projects produced both positive and negative consequences. One serious difficulty that emerged, in retrospect, was the shared funding between the two projects. Project clerical assistance, for example, was often detained by work on the larger activity (which ultimately had stricter deadlines because full classroom groups were employed for long periods of time). The staff of Evaluating Instructional Programs was very thin in numbers, never involving more than two half-time assistants, for one three month period, and usually consisting of one half-time research assistant, plus the principal investigator. Even so, those individuals were sometimes interrupted in their activity in order to assist the Practicum staff in their development effort. The consequences of these interruptions resulted in an extraordinarily uneven development activity, even more so than normally induced by the "hurry-up and wait" problem endemic to the generation of instructional material.

Schedule

The schedule for development of the project proceeded in the following broad categories: 1) Content development, 2) Product design, 3) Objectives and criterion measure refinement, 4) Exercise development, 5) Text production, and 6) Component integration.

Figure 1

Schedule of Development Events

Fall, 1972-73	Development of objectives.
Winter, 1973	Tryout of content through series of oral presentations (UCLA course and AERA).
	Preliminary exercise and measure development tryout (AERA).
Spring, 1973	Revision of exercises and measures.
Summer, 1973	Development of text, exercises, refinement of measures.
	Addition of section on basic statistics.
	First trial of partially integrated package (Florida).
Fall, 1973	Revision of materials, preparation of additional text material.
	Distribution for tryout at remote Practicum sites (Indiana, Oregon, Arizona).

Winter, 1974	Tryout of materials on curriculum specialists (Lawndale). Revision.
	Tryout on curriculum specialists (Pomona). Revision.
Spring, 1974	Tryout on reading specialist candidates (UCLA). Dissemination at AERA training session on Formative Evaluation.
	Corrections.
	Preparation of Final Report.

Development of Dependent Measures

The dependent measure problem for development activity is particularly vexing. On the one hand, we have been taught, and teach others, that adequate samples of criterion performance are important. On the other hand, we have concerns for the well-being and good humor of those acting as subjects for us. We do not wish to present test batteries that are longer than the instruction itself. Nor does the solution of item sampling meet our needs. While distributing alternative sets of items to different learners is an acceptable solution to test fatigue and idiosyncratic item problems, in the process of formative evaluation, its applicability is limited by the small numbers of individuals who participate in any one trial. Our sample was further limited by the specialized nature of the content. We did not think it fair to use the old standby sorts of subjects (such as students enrolled in Psychology 1) as our population.

Our decision was to use tests that were embedded in the instructional materials as an indication of learner success. These items, designated "tests," were administered in the context of instruction and seemed to be reasonable extensions of the work in which students were engaged. Certainly the threat to validity of sequence prompts can be raised, but our overall concern was reducing the salience of "testing" in the tryout of the materials. The dependent measures, then consisted of criterion situations that required the students to integrate information treated in each component.

Development of Content

Our experience with the planned content of the program led us to assume that previous treatments of data collection and revision of instructional materials had to be presented more simply (less technically), and in a manner which linked at the outset the kinds of data collected to the range of decisions and improvements possible in a set of instructional materials. In order to obtain feedback on the "principles" or guidelines that were to be taught by the materials, oral presentations were made by the principal author. The first of these presentations was an extended

lecture given to students in the Learning and Instruction program at UCLA during the Winter quarter of 1973. Based upon the reactions and comments of the students, the presentation was revised, principally to add some discussion of pretesting. The presentation was again presented at a Mini Training Course offered at the annual meeting of the American Educational Research Association in New Orleans. The topic of the training session was "Formative Evaluation of Instruction" and the session was very well attended (N=70+). Further refinement of the conceptual base of the product was made as a consequence of these trials. These presentations also permitted the investigator to ascertain some of the desirable prerequisites of the target population. Our feedback from these sessions was generally positive. We were repeatedly told that concepts treated were valuable and unavailable in other instructional formats. The reports of the AERA Standing Committee on Training, as derived from trainee evaluation of the session, were positive. The Committee approved a second training session on the same topic by essentially the same staff which was offered in 1974.

Product Design Activities

The materials were constructed with attention to research-based instructional techniques (Baker and Quellmalz, 1972; Holland, Doran, and Frezza, 1974). Ample opportunity to practice criterion behavior was included. Clear descriptions of tasks and feedback were provided for the learner. The sequence of activities both within and between sections was characterized by a reduction of prompts.

Beyond the studied use of such research-derived techniques, the flow of materials was to present information, to present opportunities for discrimination, and to provide opportunities to practice. Concurrently, a reverse strategy was being tried out in the major Practicum activity, where students were presented with simulations at the outset, designed to stimulate interest and to provide an organizing structure for subsequent activity. Although the tasks treated by the "Practicum" project and Evaluating Instructional Programs were distinct, and there was a considerable instructor effect in the Practicum program, the general sequence of using simulations at the outset did not prove successful, and revisions were made that incorporated a deductive strategy similar to that employed in the materials described in this report.

Practice exercises were produced for the materials by individuals who had strong backgrounds in research in psychology, teaching experience at either elementary and secondary levels, and experience in developing materials for use by teachers. The latter two attributes made these individuals particularly attentive to developing exercises that were interesting in and of themselves. In addition, a content-age level matrix was generated to assure that topics used in exercises sampled across the following areas:

1. Age of learner

2. Classroom or non-classroom use of product

3. Large or small scale development

4. Science, social science, humanities as content areas

5. Affective as well as cognitive goals

Although there were many prerequisite content areas upon which the product was based, there was an attempt to update learners' knowledge in these areas incidentally rather than directly. For example, the product contains no extended treatment regarding the areas of instructional objectives, task analysis, or domain referenced achievement testing models. However, in the exercises employed, objectives presented are operationally stated, statements of sub-tasks are included in the revision exercises, and examples of domain-statements are provided in the revision exercises as well.

Product Format

The design of the product was predicated on a print format, primarily because of cost and the assumption that present modes of dissemination favor print. During an early tryout, a series of cards was used to permit individuals to sequence exercise items of interest, but such a format proved cumbersome and was discarded in favor of a notebook. While the notebook format was feasible for early development and tryout activity (the notebooks could be refilled), cost would prohibit wider dissemination in that mode. In addition, as a casual affective gesture, photographs were inserted in the text for the last two empirical trials. The positive response generated by these photographs was dramatic and they are maintained in the final copy. Materials have been color coded by section to facilitate use.

Adaptations of Criterion

During the development effort, modifications were made in number and format of criterion items, always in the service of providing better coordination between what we asked students to do and what we hoped the program would achieve. This fact of successive item revision, a rather constant feature of development activity of any sort, reduces the clarity with which performance increments from tryout to tryout can be inferred. The goals of the program remained constant, yet sufficient modifications in the criterion measures occurred to make comparability among trials impressionistic rather than decisive.

In addition, during the tryout we collected information regarding student reaction to the ideas conveyed by the materials, the rule and format of the materials, the students' perception of the usefulness of the materials. In addition, we solicited, through questionnaire and interview, students'

view of the materials in an attempt to determine if unintended perceptions were encouraged by the materials.

COMPONENT DESCRIPTIONS

Component Descriptions:

The materials were organized in a way designed to facilitate acquisition of the objectives and subobjectives of the project. Below each major section is described, keyed to the objective and subobjectives to which it was directed.

Figure 2

Component Descriptions

Section	Description	Orientation or Objectives
INTRODUCTION	A short expository sequence that detailed objectives and purposes of materials and provided directions for use of materials.	Presented organization of materials.
I. DATA SOURCES FOR PROTOTYPE TESTING	Text, practice exercises, feedback, simulation exercises sandwiched between pre and posttests. Content treated principles of formative evaluation and recommended data sources: criterion performance on cognitive and affective indices; within program responses from written and observed behavior; use of questionnaire and interview procedures to determine full impact of materials and to diagnose deficiencies.	1. identify data sources 2. identify appropriate applications of prototype testing 3. Correct poorly designed prototype tests 4. design a prototype test
II. DATA SUMMARY FOR PROGRAM REVISION	Text, models, practice exercises and feedback. No tests. Treated simple tabular ways of reporting data by largest revisable unit. Demonstrated graphing multiple variables for the purpose of program improvement.	1. translate graphs into verbal summaries 2. Identify graphs as reflecting homogeneous or discrepant data
III. REVISING INSTRUCTIONAL PROGRAMS AND DESCRIPTION OF PROTOTYPE TEST PROCEDURES	Text, practice exercises, simulations and feedback. Pre and posttests. Exercises included description of target population, description of materials, criterion and sub-objectives, item form. Text focused on the logic of revision cycles emphasizing the research-base for decisions and offered specific guidelines for revisions given particular data patterns.	1. summarize raw data into graph form 2. interpret data and suggest revisions appropriate to data

EMPIRICAL TRYOUTS OF PRODUCTS

The empirical tryout procedures used in this project were informed by a number of considerations. First, the staff wanted to have usable data on a schedule that would permit subsequent inferences to be incorporated in product revision. Secondly, we wished to gather information on subjects that represented the range of target population to which we were directing the materials. We were partially successful in both needs.

The data gathering ventures are listed in Figure 3.

Figure 3

Empirical Tryouts for Evaluating Instructional Products

Date	Site	Type	Description of Subjects	N	Components Tested	Purpose
Jan. 1973	UCLA	Informal	Graduate students in instructional development	12	Concepts underlying program, orally presented; exercises, Section I.	To modify conceptual organization to determine viability of exercise organization.
Feb. 1973	Research Training Session, AERA, New Orleans	Formal	AERA members	70+	Revised concepts of program, orally presented exercises	same as above
July 1973	Tallahassee, Florida, A&M	Formal	Teachers, administrators	29	First drafts of texts Section I and III, Appendix Revision of Section I exercises. First draft exercises Section III and Appendix.	Determine adequacy of sequence, prerequisite assumptions, amount of practice.
Fall 1973	Sherman Oaks	Informal	Teachers	2	Rewrite of Section I and exercises.	Determine effectiveness.
Fall 1973	Univs. Indiana Oregon Arizona State	Formal*	Graduate students in instructional technology programs	13/42**	Sections I, II, III, Appendix	Test integration of sections, retest revisions of I and III, first draft of II.
Jan. 1974	Centinela Sch. Dist. (Calif.)	Formal	Curriculum Specialists and Teachers	11	Complete package	Product effectiveness and affect.
March 1974	Pomona Sch. District (Calif.)	Formal	Curriculum Specialists	14***	Complete package Rewrite of Section II	same as above

Apr. 1974	**UCLA**	Formal	Reading Specialists	24#	Total package excluding Appendix	same as above
Apr. 1974	Research Training Session, AERA, Chicago	Informal	AERA members	30	Total package excluding Appendix	Dissemination

* Formal: These tryouts were expected to be formal. In practice, however, instructors did not require them and data were returned on a schedule that made it impossible to use information in next revision.

** 13/42: Only 13 sets of materials were returned, mostly with incomplete data. Discussion with instructors reflected the fact that the semester completed before the topic of formative evaluation was thoroughly treated.

*** Although 22 sets of materials were distributed and administered, complete data were obtained for only 14 subjects.

\# Although 45 students took the pretest, the class size settled at 38. A number of teachers were absent the night of the tryouts because it was "Back to School Week" in Los Angeles. Data requirements and course schedule would not permit us to delay field tests.

Each tryout is briefly described and pertinent data presented. One might notice how the interval between trials compared with materials revised accelerated as the project drew to a close.

Tryout 1. UCLA

The conceptual tryout of the materials took place within a course context. Students had been introduced to the basic Practicum materials in the Fall quarter, but the course had ended before the topic of formative evaluation had been tested. The instructor (the present author) presented the notion of formative evaluation, data sources for prototype and field tests, the ideas of data display, and principles for revision in a two class session (approximately four hours). The presentation was interrupted by questions and comments by students. On the basis of such remarks, the instructor made some modifications in organization of concepts and added the idea of pretesting (which had been an unstated assumption). In addition, students tried out the initial set of fifteen practice exercises developed for Section I (data sources) for the project. Their role was to criticize material, detect ambiguities, and suggest corrections in exercises.

Tryout 2. American Educational Research Association

In a Mini-training session conducted under the auspices of the Committee on Research Training, the author and staff presented the concepts underlying the program to an enormous crowd, jammed in a small room. Each participant paid eight dollars to AERA for the privilege of attending the session. More than eighty individuals began the four hour period, but some left because of the scarcity of materials (only 50 were prepared) and crowded conditions. Drs. Evan Keislar and Merl Wittrock, of the UCLA program on Learning and Instruction and the Practicum staff and Dr. Edys Quellmalz, of the Southwest Regional Laboratory participated with the author in the tryout. After an introduction to evaluation, Drs. Baker and Quellmalz, presented the necessary information on data sources for formative evaluation of instruction. The participants then worked in small groups on certain of the exercises for Section I. Staff members observed and collected information regarding points of difficulty. The session met with apparent success as a second offering was scheduled for the next annual meeting (1974) of AERA. No formal data of an achievement sort was collected. A brief questionnaire was circulated to the group but reponses were irregular and focused mostly on the fact that the room situation was poor.

Tryout 3. Florida A & M University, Tallahassee

During the Spring, the project was contacted by Dr. Wally Cox of Florida A & M University, regarding materials for a workshop he was conducting on evaluation. Our staff agreed to supply him with about one day's worth of material in the area of instructional evaluation. His group was anticipated to be around 30 individuals, much larger than what we would have selected for a first formal tryout. However, we agreed to the project.

Dr. Arlene Fink, of the UCLA Center for the Study of Evaluation, had planned to participate in the workshop using Center materials. She agreed to modify her schedule and to monitor our part of the tryout for us, thus saving travel funds. Dr. Fink prepared a report summarizing her activities, her observations of the class and suggested possible ways to interpret the data.

The 29 participants included 16 public school teachers and 13 administrators. All came from the Florida region and were distributed in elementary and secondary school posts. The procedures of the tryout were as follows:

Introduction, 15 minutes

Pretest, Part I, 15 minutes

Data sources, 1 hour

Posttest, Part I, 15 minutes

Pretest: Appendix (Data Analysis), 15 minutes

Appendix, 1 1/2 hours

Posttest, 20 minutes

LUNCH

Pretest, Part II, Revising Programs, 15 minutes

Part III, 1 hour, 50 minutes

Posttest, Revising Programs, 25 minutes

No more than three-minute schedule overruns were permitted and no additional breaks were given. Although materials were intended to be self-instructional, groups of subjects worked together to clarify written materials and to share problems and strategies in completing the practice exercises. Such cooperation was fostered because the subjects were crowded and the previous three days of the institute had accustomed participants to working together. In addition to completing the pre and posttests the participants filled in a self-description sheet regarding background information and a post instruction attitude instrument. They were also encouraged to write notes and make comments in their notebooks.

Materials were presented in a notebook format, with each section color coded. Because of time pressure, material was reproduced by spirit master and was **loathsome purple-print.**

Results

Subtests for each section contained six items. In Section I students were

asked to identify appropriate data sources, in Section II to solve simple computational problems in data summary, and in Section III to determine whether revision was necessary or not when presented with product descriptions and data to analyze.

Table 1

Pre and Post Test Scores: Florida A & M

	PRETEST	POSTTEST
Section I	18%	63%
Section II	30%	56%
Section III	37%	33%

The strange, and disappointing results, particularly for Section III had some explanation. Student participation and completion of exercise material was reasonably good for Section I, 21 of 29 students completing the materials. For Section II, 13 of 29 completed the materials, and for Section III, only 8 people completed the last exercises. Subjects also commented profusely on the posttest items for Section III, with indications that they did not understand what was expected of them. Dr. Fink also noted that the compressed treatment time was a problem for many of the students who expressed the need to think over some of the material presented rather than charge through the notebooks during one day.

Attitude information collected was also puzzling. Given poor posttest performance, the attitude data was unusually good. All 29 individuals indicated that they felt the materials would be useful in their job situation. Of the 23 subjects completing the attitude instrument, 21 felt Section I was helpful, 19 thought Section II (Appendix) was helpful, and 20 thought Section III was helpful. Negative comments centered about the large number of practice exercises and confusing directions and language in the text. Most frequent comments of a constructive nature focused on the need for more time, and the suggestion that group discussion be interpolated with reading and exercise material.

Revision

Following Trial 3, the following revisions were instituted:

1. Content related to field testing was dropped; emphasis was now on instructional prototype tests, as examples of formative evaluation with the most serious and interesting problems for the potential audience of the book.

2. Clarification of directions for exercises and tests.

3. Clarified and reduced numbers of practice exercises for Sections I and III.

4. Moved Data Summary section to Appendix and deleted pre and posttests.

5. Prepared test for Section II.

6. Edited and simplified language throughout.

7. Revised all pre and posttests.

8. Improved attitude survey so responses would be more specific.

Trial 4. Los Angeles

In anticipation of planned tryouts at the four sites where the Practicum project was being tried out, some of the revisions were tested on a small sample (N=2) of teachers in the Los Angeles area. These teachers went through the materials, took all tests and commented profusely on the notebooks. Inferences from the data were difficult to make in that one individual did extremely well throughout, while the other had continued difficulty. Neither of the individuals appeared to like the materials, and comments related to boredom were scrawled on the materials. Participants seemed to feel there were too many examples included in the text materials and that too many practice exercises had been produced. Each of the teachers had read materials in one unrelieved sitting and thus feelings of boredom might have been exacerbated by the compressed treatment and lack of socialization. Nonetheless, the staff was considerably depressed at this point, and introduced into the subsequent versions, the directions for subjects to bypass exercises when they felt they understood the concepts.

Trial 5. Oregon, Indiana, Arizona

These trials were anticipated for students in the Practicum courses being simultaneously offered at the University of Oregon, Arizona State University and Indiana University. Revision of materials and necessary reproduction in preparation for such tryouts occupied a great deal of time. Materials were mailed as scheduled well in advance. Unfortunately, due to the exigencies of course organization, the materials were apparently not seriously required at any site. Most devastating for us, however, was the fact that they were not returned promptly. We expected these trials to provide us with high quality information. Instead, we received only partial data from individual students. We also received a number of individual letters from students indicating that they felt positively about the materials and would like to have copies of the materials made available to them at a later time. However, the tragedy for the project was that after considerable expenditures, we did not have reasonable information about whether the materials, as revised,

worked any better.

Notebooks were distributed to the Practicum sites to be administered by the professor in his regular course. Data were returned from only seven of the 35 notebooks prepared. Results for four of six individuals were received from the Indiana trial, but only three individuals returned materials from Arizona and no usable information came from Oregon.

COMPONENT	INDIANA			ARIZONA		
	N	PRETEST (18)*	POSTTEST (18)	N	PRETEST (18)	POSTTEST (18)
Data Sources	4	6.25 (22)	16.00 (24)	3	4.33	15.00
Revising Programs	3	0.00	20.66	1	no data	no data

* number possible.

A follow-up questionnaire yielded only three responses. Although the results are especially scanty, there is evidence that the programs were controlling responses from pre to posttest. However, dependency on such sparse data was suicidal and we decided to test again.

Trial 6. Lawndale, California

In desperation, a local teacher institute was set up and taught by a staff member at a nearby school district. The focus of the institute was on evaluation and the materials provided a basis for the course offering. In January, the materials were used by eleven secondary school teachers. Their results, by sections are presented below:

Table 2

Performance of Lawndale Teachers

	PRETEST	POSTTEST
Section I	24%	71%
Section II	NO TEST	NO TEST
Section III		
Select	5%	63%
Implement	7%	24%

Attitude information was also collected for this trial. In interviews, all participants indicated that they could use again the concepts they

learned, especially the revision principles. The revision section was considered the most useful part of the material and only one of the eleven did not feel competent to plan the evaluation of instruction. Program strengths were reported by participants in terms of the value of practice and feedback exercises, the concepts taught and appendix. Weaknesses identified related to writing style and a lack of sufficient examples (the opposite criticism from Trial 4). The observers noted that boredom was expressed during Section I. A desire for group practice on exercises was observed during this trial. The most serious problem was in maintaining teacher's attention at the time of day the workshop was offered (3 to 5 p.m.) after a full teaching day.

Revision

Some difficulty emerged in the revision section again. The material was seen to be of value to participants and they were able to identify principles for application based on the analysis of data. However, a difficulty related to the extent to which they could provide adequate descriptions of how such principles would be implemented. One might rationalize this difficulty based upon the fact that they were given very abbreviated descriptions of hypothetical materials and thus their implementation might tend to be general rather than specific. The test itself was scrutinized, and a thorough analysis of pre to posttest items on the revision section demonstrated that there was unequal difficulty between the two tests. The pretest provided a situation with two objectives and two subobjectives for analysis, while the posttest asked for three objectives and three subobjectives. On the pretest examples, all data were honogeneous across data sources, while on the posttest a number of conflicts were introduced. On the pretest, ten of the eleven subjects were unable to suggest even one correct revision activity, whereas, after instruction, all were able to suggest at least one, and most three or more appropriate principles of revision.

Following the data analysis, the program was substantially revised.

1. Section II was expanded.
2. Pretests and posttests were made parallel.
3. Directions were clarified.
4. Answer sheet formats were changed.
5. Text was heavily edited and simplified.
6. Text section was visually reformated.
7. Photographs with captions (for affect) were added.
8. Glossary was added.
9. Teacher guidelines were prepared.

Trial 7. Pomona

Seventeen curriculum specialists provided data from a trial conducted by a graduate of the UCLA doctoral program. Twenty-two subjects participated in the tryout, eleven reading and eleven mathematics specialists. Only 17 participants completed the workshop. The workshop was scheduled on two half day periods. Report of the results is presented below:

COMPONENT	N	PRETEST (18)	N	PRETEST (18)
Data Sources	17	2.93	17	13.35
		(28)		(38)
Revising Programs	17	17.82	14-16	31.38

Affective data were obtained through the use of a questionnaire. Participants were apparently satisfied with the materials, a departure from pervious tryouts. Of the 15 responses to this questionnaire, the average rating was 2.82 on a 1-4 scale for the attention holding properties of the materials. Positive responses (2.91) were also obtained for the Appendix section. Participants rated the fairness of the tests and clarity of feedback especially high (3.13). The average rating across all items in the questionnaire was 2.83. A copy of the questionnaire is included in the Appendix.

Trial 8. UCLA

Was this tryout the criterion test? We had anticipated that it was. On the other hand, we felt that with minor revisions, the product might be substantially improved. To that end, we conducted a tryout in April that would simulate an alternative form of product use. Students were pretested on the materials in two class sessions, allowing about 20 minutes per session for responses. Notebooks were distributed to students with directions to complete all the activities and to be prepared to take a posttest on the material. We attempted to make it clear that no grading contingency was related to performance on the test. At the next class session, after a one week interval, students completed the posttest and questionnaire.

Students in the tryout were teachers studying to obtain a reading specialists credential and a master's degree. The results from the tryout are presented below in Tables 3, 4, and 5.

Table 3

Performance of UCLA Students in Final Tryout

PART I. DATA SOURCES	PRETEST $\bar{X}$	POSTTEST $\bar{X}$
Data Sources (4)*	.98	3.84
Subject Characteristics (2)	.28	1.56
Instruments (5)	.67	4.44
Administrative Problems (7)	.73	4.36
PART II. REVISION		
Graphing (3)	.06	(4) 4.00
Identifying Principles of Revision (10)	1.24	7.35
Implementing Principles (1)	.00	.66

* number possible.

Students were also asked to keep records of their reading times and were encouraged not to attempt to assimilate the entire set of materials at one sitting. Results of reported reading times are presented in Table 4.

Table 4

Reading Times - UCLA Final Tryout

	TEXT $\bar{X}$	EXERCISES
Data Sources	47.95 minutes	50.62 minutes
Revision	55.43 minutes	80.00 minutes

Estimates obtained for section averaged about 1 hour. Therefore, the entire time required to complete the materials at the demonstrated level of performance was under five hours.

Table 5

Results of Attitude Survey

	$\bar{X}$
Overall rating of product (4)	3.1[illegible]
Concepts clearly presented	3.04
Posttests fair	3.24
Directions for practice clear, easy to follow	3.32
Information will be useful later	3.24

From an inspection of the data, one might infer that the staff was pleased, as well as relieved by the results.

PERSONNEL AND PROJECT COSTS

Project personnel have been limited to the principal investigator and a small group of graduate students or research assistants. Early components of exercises were prepared by Judy Safford and Samellyn Wood during the Winter of 1973. Deborah Feinberg was responsible for the review section on statistical procedures. Both she and Lynn Smith prepared practice exercises and early versions of the criterion tests. The majority of exercise writing, revision and data analysis was executed by Aleta Saloutos. Shireen Powell assisted in data analysis. The burden of coordination of materials and field tests was Ms. Saloutos' responsibility, although Ms. Smith assisted her in her task. Unfortunately, the staff assignments were thinly dispersed over the period of funding. Lynn Smith has been with the project over the period of June through May, although only three months of half-time employment was compensated. Aleta Saloutos joined the project half-time in September and continued until April 30 in that role. The discontinuity of assignments and moderate proportion of compensated time represented minor problems. Other roles tended to interfere with the project's activity, especially responsibilities associated with the Practicum and other job commitments. Nonetheless the project held together, apparently through loyalty of the junior staff.

Staff costs were small. From expenditures reported as of June 1, we estimate that less than $13,000 were spent on the design, development and testing of these materials. When we include supplies, reproduction costs, overhead, and travel for dissemination purposes, the project was completed for considerably under $20,000. A gross transformation to costs by hour of "validated" materials would be under $2,000 for each hour of instruction. As a comparison, the project by Baker and Quellmalz (1972) required approximately $22,000 in personnel costs for a package of 12 hours duration. Apparently, in the face of inflation, development is becoming slightly more efficient.

The project was benefited by the low overhead rate from UCLA Extension, as well as the "free" work voluntarily performed by the staff. The relatively low cost figure may encourage the funding of discrete (rather than large scale, programmatic) development efforts at university sites.

DISSEMINATION

One pervasive concern of developers is the problem of distribution of materials designed for specialists rather than widely distributed groups such as teachers. It is not always easy, even when materials have been subjected to rigorous evaluation to pursue the usually desirable route of commercial publication. Ideally, commercial publication would be the ideal destination for these materials.

As far as more informal sorts of advertising, some interest in the materials has already been created and demonstrated by letters and other requests for them. First, the external field testing of the materials has developed an interest in both instructors and students at those sites. The absolute numbers are small, but since field tests were conducted at institutions with active programs in instructional development, the ultimate consequences of their continued use at these sites are encouraging. Second, two Mini-Training sessions, sponsored by the American Educational Research Association were conducted during 1973 and 1974 annual meetings. Evan Keislar, Merl Wittrock, and Eva Baker of UCLA and Edys Quellmalz of the Southwest Regional Laboratory for Educational Research and Development participated in the training, conducted on the topic identical to the materials under development. Over one hundred people participated in these sessions, many of whom expressed interest in using the materials to train students or staff in formative evaluation skills.

The materials have also been offered as potential training vehicles for development projects in the area of career education. Their greatest utility, however, may depend upon their promotion for use by teachers in school settings. Especially, as teachers' roles and functions change with the advent of new instructional systems,[1] the importance of regularized and local formative evaluation of instruction will be recognized. Thus, while the principles espoused in training were planned for instructional development activity, learning resource specialists in school settings might be able to apply the same procedures to evaluation efforts at the building or classroom level. Certain reactions from our tryout populations (most of whom were teachers or curriculum specialists) suggest such a conclusion.

We have confidence that these materials meet a need in the field, a need that is presently unsatisfied. Hopefully, the outcome of our activity may be made available beyond the limited numbers of copies we are able to print.

---

[1] Report of Panel 8, Conference on Teaching, National Institute for Education, Washington, D.C., June, 1974.

SUMMARY

This project produced an integrated set of materials that are appropriate for use in graduate training, research and development settings, and public school personnel use. The materials were developed in the course of repeated tryout and revision cycles and now meet the expectations of their designers. The development history of the materials was limited because of format expectations and concerns for dissemination. However, the present form of the materials is apparently appropriate.

No catastrophes marked the development process of Evaluating Instructional Programs. The relationship of this materials development effort to the larger Practicum model training project introduced some unavoidable difficulties, particularly in staff utilization. However, positive side effects were also stimulated by this project. The activity supported in part five different graduate students, one of whom was influenced to pursue doctoral work as a function of her responsibilities on the project. The opportunity to work on these materials also suggested some alternative ways of conceptualizing research on instructional roles, methods which will be explored in planned research by this author.

INSTRUCTOR GUIDELINES

Evaluating Instructional Programs

- Directions to Administer Program
- Observation Forms
- Pre and Posttests (with answers)
- Glossary

I. DIRECTIONS FOR ADMINISTERING PROGRAM

GENERAL DIRECTIONS FOR THE INSTRUCTOR
(Used in Tryouts)

Evaluating Instructional Programs is designed to be primarily self-instructional but can also be used effectively in a workshop or discussion setting. The program consists of three major sections and an Appendix:

Section I Data Sources for Prototype Testing

Section II Data Summary for Revision

Section III Revising Instructional Programs

Appendix Elementary Data Summary Procedures

The materials in each of these sections, with the exception of Data Summary for Revision, are accompanied by pre and posttests. Regardless of the method (self-instructional or teacher-mediated) in which you present the materials, you will want to

1) remind participants to complete the Personal Inventory preceding the Introduction; collect the Inventory forms,

2) remind participants to write their names on the pretest at the beginning of each section; collect the pretests after they have been completed,

3) remind participants to write their names on the posttests at the end of each section; collect the posttest,

4) instruct participants to reply anonymously to the questionnaire provided at the end of the text; collect the questionnaires.

After the participants have completed Section II, they should be asked to review the Table of Contents in the Appendix. The Appendix provides supplementary instruction in the summarization of data, construction and interpretation of histograms and frequency polygons. One must be able to apply these techniques in order to do the exercises in Section III. If there are topics in the Table of Contents with which an individual is unfamiliar, he or she should then read the appropriate part of the Appendix and complete the practice problems that accompany the explanatory material. If participants are familiar with all of the topics in the Table of Contents, there is no need to attend to the Appendix.

In order for this program to be of maximum effectiveness to participants, modifications should be considered in two areas of the instructional sequence:

1. terminology: Unfamiliar language used in development and evaluation should be explained to the participants. This might be done on an "as needed" basis by the instructor, or in written form by duplicating the accompanying Glossary.

2. relevance: Special attention should be given to pointing out the utility of the program to specific participant groups. Teachers, for instance, will benefit from learning a) to select materials that offer evidence of preferred development procedures, b) to measure empirically the effectiveness of instructional materials with their own students, and c) to utilize research-based principles for improvement of materials. It is highly desirable that the Instructor take advantage of any opportunities to emphasize the utility of the materials for the particular group of participants with whom he/she is working.

We would greatly appreciate your completion of the Instructor's Observation Checklist, and any additional remarks which would help us increase the effectiveness of the program. Your cooperation in participating in the use of these materials is sincerely appreciated by the development staff.

INSTRUCTOR'S OBSERVATION CHECKLIST

Evaluating Instructional Programs

Section I

	Start	Finish	Elapsed Time
A. Administration of Materials			
1. Time taken to complete pretest	_____	_____	_____
Remarks:			
2. Time taken to read expository material	_____	_____	_____
Remarks: (ability to proceed without instructor aid, confusion re: terms, concepts)			
3. Time taken to complete practice exercises	_____	_____	_____
Remarks: (ability to complete without aid, frequency of referral to text)			
4. Time taken to complete posttest	_____	_____	_____
Remarks: (frustration level, clarity of directions)			

B. Attitude Toward Materials

1. Participant comments about difficulty level of material:

2. Participant comments about utility of material:

INSTRUCTOR'S OBSERVATION CHECKLIST

Evaluating Instructional Programs

Section II

	Start	Finish	Elapsed Time
A. Administration of Materials			
1. Time taken to complete pretest	_____	_____	_____
Remarks:			
2. Time taken to read expository material	_____	_____	_____
Remarks: (ability to proceed without instructor aid, confusion re: terms, concepts)			
3. Time taken to complete practice exercises	_____	_____	_____
Remarks: (ability to complete without aid, frequency of referral to text)			
4. Time taken to complete posttest	_____	_____	_____
Remarks: (frustration level, clarity of directions)			

B. Attitude Toward Materials

1. Participant comments about difficulty level of material:

2. Participant comments about utility of material:

INSTRUCTOR'S OBSERVATION CHECKLIST

Evaluating Instructional Programs

Section III

	Start	Finish	Elapsed Time
A. Administration of Materials			
1. Time taken to complete pretest	____	____	____
Remarks:			
2. Time taken to read expository material	____	____	____
Remarks: (ability to proceed without instructor aid, confusion re: terms, concepts)			
3. Time taken to complete practice exercises	____	____	____
Remarks: (ability to complete without aid, frequency of referral to text)			
4. Time taken to complete posttest	____	____	____
Remarks: (frustration level, clarity of directions)			

B. Attitude Toward Materials

1. Participant comments about difficulty level of material:

2. Participant comments about utility of material:

II. PRE AND POST TESTS WITH ANSWERS

NAME_____

PRETEST: EVALUATING INSTRUCTIONAL PROGRAMS

Prototype Testing

Directions: Read the description of the material which follows then answer the questions below in as few sentences as possible.

Materials Description

An experienced master teacher is developing a series of materials she hopes will teach elementary school children to read music more easily than existing state adopted texts. The components she has planned are as follows:

- a filmstrip and accompanying tape which depicts musical notes as talking cartoon characters.
- a pupil booklet which includes simple sheet music, a short text, and plastic overlays displaying various types of musical notes and symbols.
- a teacher's guide keyed to both the filmstrip and booklet.
- all materials are presently in rough form.

ASSUME IT IS YOUR JOB TO "PROTOTYPE" TEST THESE MATERIALS. THE INFORMATION GATHERED FROM THE TEST WILL PROVIDE THE BASIS FOR REVISION OF THE MATERIALS. ANSWER THE FOLLOWING QUESTIONS AS IF YOU WERE PLANNING SUCH A TEST.

1. What information about the materials do you wish to acquire during the test?

_____

_____

_____

_____

2. What kinds of subjects will you use?

_____

_____

_____

NAME_____

Prototype testing (cont.)

3. What measures or data sources will you use to obtain the information you desire?

4. What preparations will you make in advance to make sure that your test runs smoothly?

PRETEST CONFIRMATION SHEET

Data Sources

The responses suggested herein are designed to be representative rather than exhaustive examples of adequate ways in which each question might have been answered.

Question 1

All responses should focus on acquiring information from or about students rather than teachers.

SAMPLE RESPONSES:

a) collect criterion data
   Did students attain the objective(s)?

b) collect practice (within-program) data
   How well did students perform on en route objectives?
   How well did students perform on practice exercises?

c) collect affective data
   Did students feel they understood the instruction?
   Did students like the program?
   Were materials too difficult or simple?

d) collect observer data
   Did students use the materials in the prescribed manner?
   Was there evidence of frustration or other difficulties in using the materials?

Question 2

A small number of students (no more than 10-12) selected from all grade levels in which the materials are designed for use.

Question 3

Pretest data is not necessary if baseline data has been previously collected. Responses should focus on students.

SAMPLE RESPONSES:

a) practice exercises within the program completed by student

b) criterion test (posttest)

c) attitude survey

d) interview forms

e) observer notes

Confirmation Sheet (cont.)

Question 4

Suggested preparations for a smoothly administered test include the following:

a) secure permission to administer test

b) secure a location with a proper testing environment

c) train observer(s)/interviewer(s)

d) provide extra copies of test materials

e) double check materials for errors

f) ensure names are included on products or written materials, design coding method if names are omitted

g) review directions with all persons involved in administering the test

PRETEST

Revising Instructional Programs

Directions: You are provided with the following development information:

1. description of target population
2. description of materials
3. terminal and sub-objectives
4. terminal objective item form
5. raw prototype test data from four sources

YOUR TASK:

1. Plot criterion test and subobjective data using frequency polygons.
2. Identify data patterns for terminal and subobjectives.
3. Suggest revision decisions consistent with research-based principles and product specifications.
4. Describe how you would implement your recommendation(s) for revision.
5. Be sure to write your name on the answer sheets.

PRETEST

"Discovering Your Philosophy"

PRODUCT SPECIFICATIONS

Target Population: "Discovering Your Philosophy" is an "applied philosophy" course to be offered to senior high school students. The course is designed for students who wish to clarify their own values on controversial issues and be exposed to various philosophical viewpoints.

Description of Materials:

1. Hidden Feelings--a short programmed book with guidelines and 25 exercises to train students in identifying value terms and phrases.

2. "What Do You Think?"--a pamphlet containing a series of issues about which students must write a one paragraph value statement.

3. Four Great Philosophers--four short booklets providing thumbnail analyses of the major philosophical ideas of Plato, Rousseau, Aristotle and Marx. Each booklet contains ten exercises in which students must discriminate one philosopher's assumptions from those of other philosophers.

4. An attitude questionnaire.

Terminal Objectives:

Given a political, social or economic question about which a policy decision must be made, the student will

1) identify the value issues of the policy question
2) state his/her own values
3) indicate an appropriate historical or philosophical referent for each of the values mentioned.

Subobjectives:

1. Student will identify value terms included in position statements.
2. Student will write a short position paper recommending a policy decision on a given issue, using connotative words and phrases.
3. Student will identify value terms included in position statements.

Terminal Objective Form:

Response Description: The student will write an essay or tape record a response which specified:
1) three value issues of the policy question
2) a summary of personal values with regard to the policy
3) historical or philosophical referents for each of his/her value statements.

Content Limits: The learners will be given a list of policies for a) the American economy, b) personal moral behavior, c) international relations, d) political actions, e) penal reform, f) abortion and euthanasia, g) treatment of dissident groups, and h) treatment of minority groups. Student will select one policy from the above upon which to base his/her response.

Criteria: Identification of value issues means the learner will list any courses of action and/or questions that include judgmental terms. A summary of personal values will include an assertion of what is right, desirable, important, necessary or the converse. Historical and philosophical documentation for these assertions should be drawn from a) philosophers discussed in class, b) prominent spokesmen for certain causes drawn from the press or national T.V., or c) any published interpretation of historical events. Documentation must be given for at least five assertions.

Outline of Prototype Test Procedures:

Four high school students participated in prototype testing procedures. Students read and completed practice exercises for each of the Great Philosophers booklets, then discussed their reactions and responses to the exercises in class in order to obtain feedback for their tasks. "What Do You Think" and Hidden Feelings were completed as homework assignments, while the rest of class time was devoted to discussion of personal values as well as those of the philosophers presented in the materials. Students were encouraged to introduce ideas of philosophers and historians not covered in class. Finally, students completed the criterion task. Only one student chose to record, rather than write, responses to the exam. Upon the completion of instruction, students completed an attitude questionnaire and then held a "round table" discussion about the materials. The instructor kept notes of pertinent remarks made during this dialogue.

PROTOTYPE TEST DATA

Criterion Test Data

Terminal Objective #1
(Identify 3 value issues)

Student	Score
Clancy	1
Siegfried	3
Mathilda	2
Yvonne	1

Terminal Objective #2
(State own values)

Student	Adequate Response	Inadequate Response
Clancy	x	
Siegfried	x	
Mathilda	x	
Yvonne	x	

Terminal Objective #3
(Indicate value referents for at least 5 assertions)

Student	Score
Clancy	1
Siegfried	2
Mathilda	4
Yvonne	1

Subobjective Data

Subobjective #1
(identify value terms --25 pts. possible)

Student	Score
Clancy	20
Siegfried	17
Mathilda	12
Yvonne	15

Subobjective #2
(write position paper)

Student	adeqt. respon.	inadeqt. respon.
Clancy	x	
Siegfried	x	
Mathilda	x	
Yvonne	x	

Subjective #3
(identify sources --40 pts. possible)

Student	Score
Clancy	30
Siegfried	27
Mathilda	18
Yvonne	22

Results of Questionnaire

	Agree	Disagree
I enjoyed the sequence of values	4	
I found Four Great Philosophers easy to use	3	
I thought "What Do You Think?" was helpful in clarifying my own values.	4	
Hidden Feelings helped me learn about words that imply values.	4	

The following comments are representative of comments made during the round table discussion.

"I really enjoyed this program."

"I'd like to do it again."

"Even though I didn't do well, it was fun."

"Some of the exercises were too difficult."

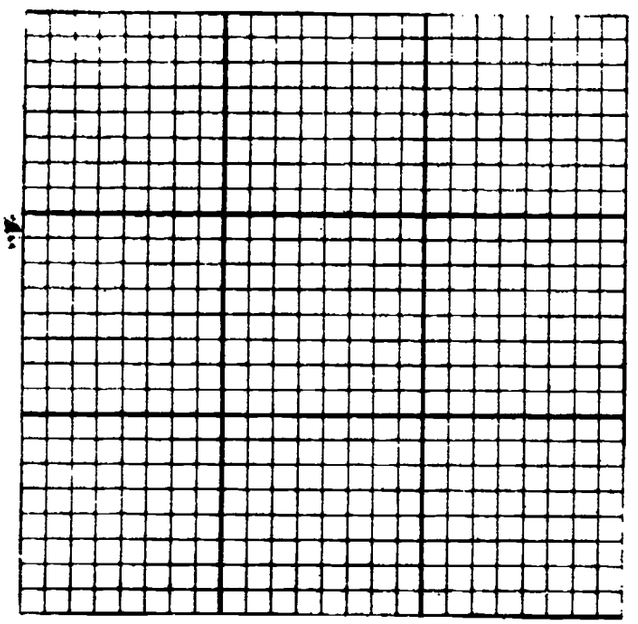
"I never seemed to quite catch on."

NAME_____

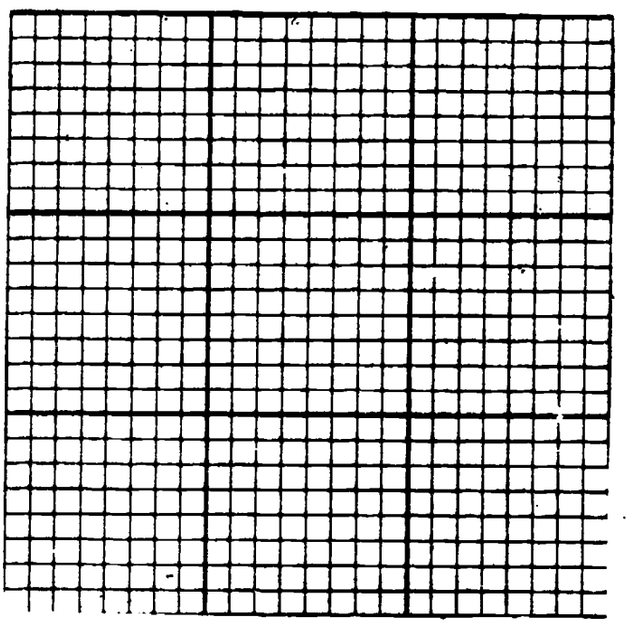
PRE-TEST ANSWER SHEET

Directions: Plot terminal objective data, using frequency polygons, in the spaces provided below. Describe data patterns on the line below each graph. Plot subobjective data in the same manner but on the next page. BE SURE TO LABEL EVERY GRAPH. After you have finished, go on to the following pages to suggest revision and implementation procedures.

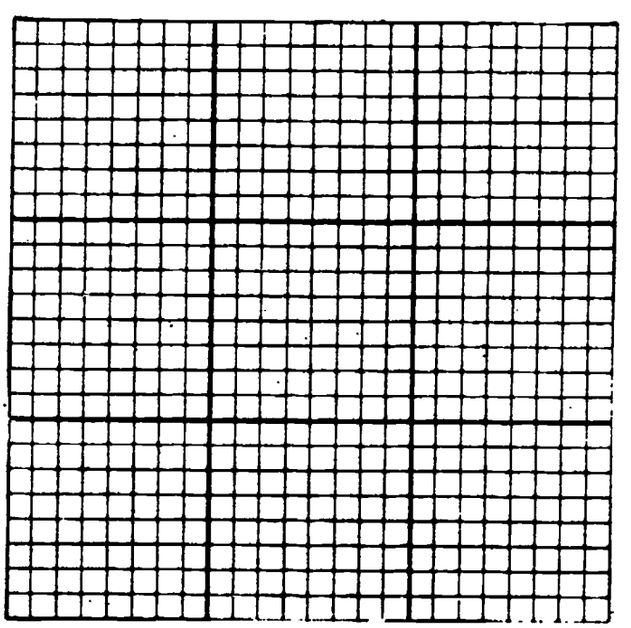
CRITERION TEST RESULTS



Data Pattern:_____

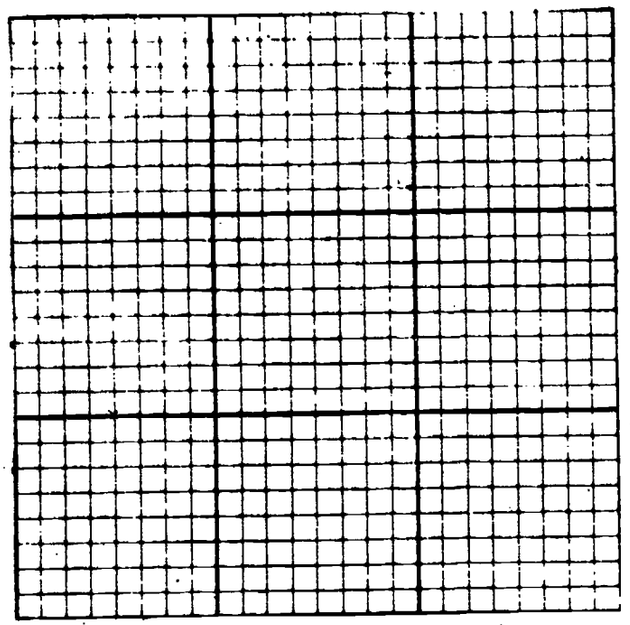


Data Pattern:_____

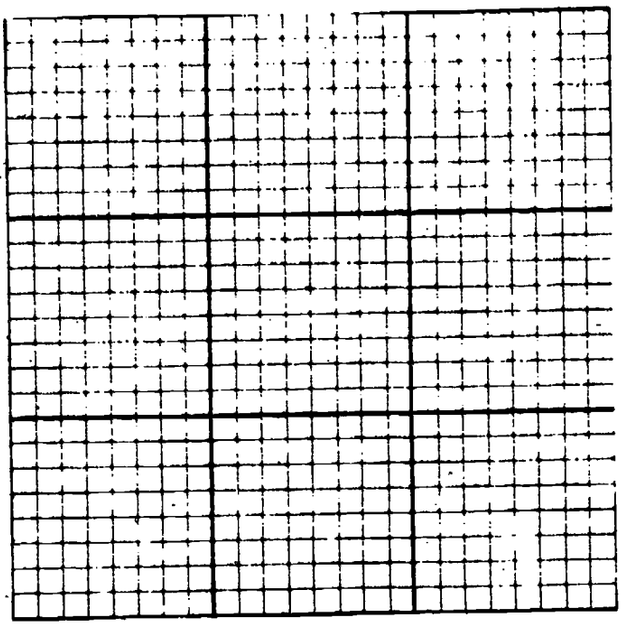


Data Pattern:_____

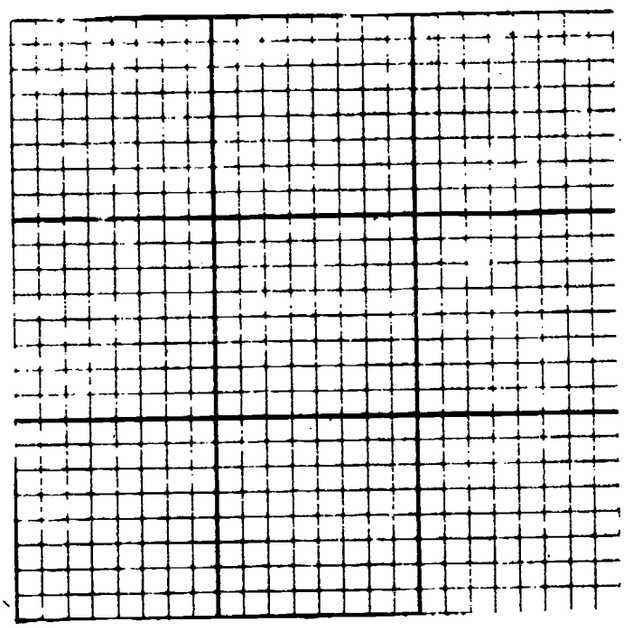
SUBOBJECTIVE RESULTS



Data Pattern:_____



Data Pattern:_____



Data Pattern:_____

NAME_____

Pretest Answer Sheet (cont.)

2. Which research-based revision principles would you use to modify this instructional product? Mark an X by all revisions that apply.

Terminal Objective #1, Subobjective #1

____ add practice
____ review task analysis
____ verify unprompted practice
____ pool and redistribute practice and posttest items
____ add feedback
____ provide task description (assure learner understands task)
____ provide motivational stimuli
____ carefully delete irrelevancies
____ leave everthing alone
____ *change format

Terminal Objective #2, Subobjective #2

____ add practice
____ review task analysis
____ verify unprompted practice
____ pool and redistribute practice and posttest items
____ add feedback
____ provide task description (assure learner understands task)
____ provide motivational stimuli
____ carefully delete irrelevancies
____ leave everything alone
____ *change format

Terminal Objective #3, Subobjective #3

____ add practice
____ review task analysis
____ verify unprompted practice
____ pool and redistribute practice and posttest items
____ add feedback
____ provide task description (assure learner understands task)
____ provide motivational stimuli
____ carefully delete irrelevancies
____ leave everything alone
____ *change format

Please go on to the next page.

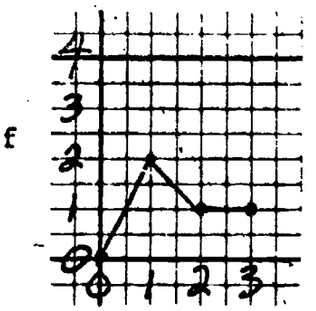
3. How would you modify specific instructional activities listed in the specifications to apply the principle(s) you have selected above? (e.g., "rewrite the directions in the Teacher's Guide")

PRETEST CONFIRMATION SHEET

Revising Instructional Programs

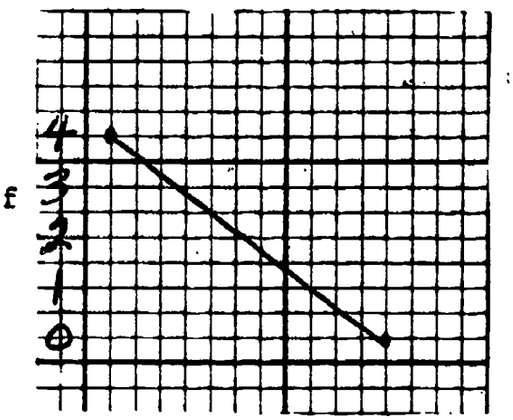
1. Criterion Test and Subobjective Data

Terminal Objective #1



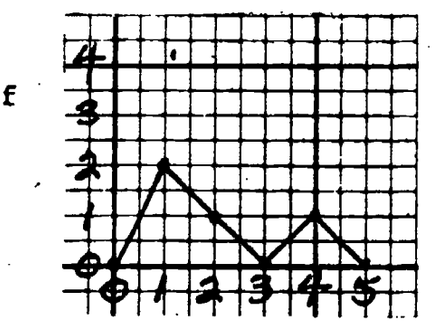
items correct

Terminal Objective #2



adequate inadequate

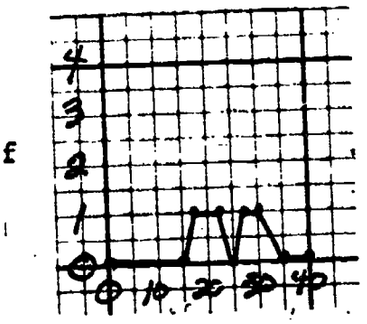
Terminal Objective #3



correct items

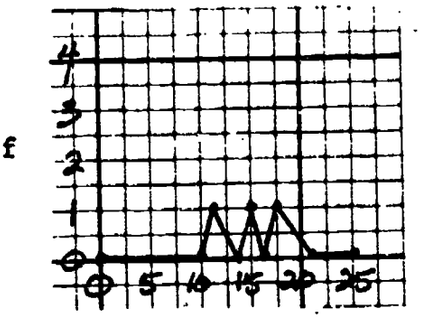
Pretest Confirmation Sheet (cont.)

Subobjective #1



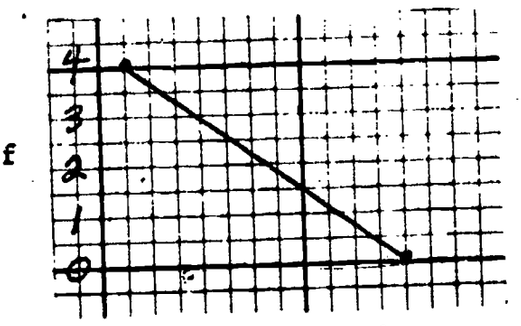
points attained

Subobjective #2



points attained

Subobjective #3



adequate inadequate

Pretest Confirmation Sheet (cont.)

2. The research-based revision principles which should be indicated are given below:

Terminal Objective #1, Subobjective #1

X	add practice
X	review task analysis
____	verify unprompted practice
____	pool and redistribute practice and posttest items
X	add feedback
X	provide task description (assure learner understands task)
____	provide motivational stimuli
____	carefully delete irrelevancies
____	leave everything alone
____	*change format

Terminal Objective #2, Subobjective #2

____	add practice
____	review task analysis
____	verify unprompted practice
____	pool and redistribute practice and posttest items
____	add feedback
____	provide task description (assure learner understands task)
____	provide motivational stimuli
X	carefully delete irrelevancies
X	leave everything alone
____	*change format

Terminal Objective #3, Subobjective #3

X	add practice
X	review task analysis
____	verify unprompted practice
____	pool and redistribute practice and posttest items
X	add feedback
X	provide task description (assure learner understands task)
____	provide motivational stimuli
____	carefully delete irrelevancies
____	leave everything alone
____	*change format

Pretest Confirmation Sheet (cont.)

3. Suggestions for implementing the revision principles in question 2 might be as follows: (responses are intended to be merely suggestive)

Terminal Objective #1, Subobjective #1

Add feedback to Hidden Feelings.
For example, the booklet should be returned as soon as possible following its completion by the student.

Add more practice and/or exercises to the Hidden Feelings booklet to ensure that the students have adequate opportunity to become familiar with the task.

Make sure students understand criterion task for the terminal objective of what is expected. This might be done orally by the instructor in class, or amended to the booklet and position statement papers.

Review task analysis of requisite skills.
Does Hidden Feelings omit any important training in helping students identify value terms and phrases? Is adequate practice given for the terminal objective?

Terminal Objective #2, Subobjective #2

All materials should be left as they are. Although it would be technically correct to suggest deleting irrelevancies from "What Do You Think?" the subject matter would seem to make this task an exceedingly difficult one.

Terminal Objective #3, Subobjective #3

Add more opportunities for student practice in identifying the viewpoint of a given philosopher, e.g. more structured discussion practice, short quizzes, additional written exercises.

Make sure students receive feedback for philosophical referents. For example, discussion of responses to exercises contained in Four Great Philosophers may not provide sufficient feedback. One might add written confirmation to accompany the exercises then test again to see if the problem has been resolved.

Make sure students understand the task. Perhaps the directions in Four Great Philosophers are not well written. In addition to clarifying extant directions, a further description of what the students are to do might be given orally by the instructor during class time.

Review the task analysis for Four Great Philosophers. It is possible that the skills required in Objective #3 are built upon prerequisite training not provided in the instruction. Admittedly, given the subject matter of this material, task analysis may be a difficult procedure.

POSTTEST: EVALUATING INSTRUCTIONAL PROGRAMS

Prototype Testing

Directions: Read the description of materials given below. Use the worksheet to write an outline of the major elements of a prototype test plan you would design for the materials. Be specific in your use of data sources, measures, subjects and provision for administrative issues. Be sure to write your name on the worksheet.

Materials Description

A major difficulty experienced by students at the graduate level centers upon writing a dissertation proposal. Three professors in the Education Department are attempting to reduce this problem by developing materials which will help graduate students develop skills relevant to this sometimes traumatic task. Rough models of the materials are now ready for testing. The components of the instructional kit are as follows:

"The Harrassed Students' Guide to Literature Review" a self-instructional booklet which provides practice in the skills necessary for a thorough literature review

a film entitled "Scratching Your Intellectual Itch" which suggests human resources for review and improvement of a proposal

simulations (five) called "The Pitfalls of Procedures" which require participants to criticize quantitative data gathering techniques and suggest alternative strategies for given research situations

The objectives of the products can be inferred from the component titles.

On the next page, outline your proposed prototype test plans for these materials.

NAME _____

POSTTEST: EVALUATING INSTRUCTIONAL PROGRAMS

Worksheet

POSTTEST: EVALUATING INSTRUCTIONAL PROGRAMS

Revision

Directions: You are provided with the following development information:

1. description of target population
2. description of materials
3. terminal and subobjectives
4. terminal objective item form
5. raw prototype test data from four sources

YOUR TASK:

1. Plot criterion test and subobjective data using frequency polygons.
2. Identify data patterns for terminal and subobjectives.
3. Suggest revision decisions consistent with research-based principles and product specifications.
4. Describe how you would implement your recommendation(s) for revision.

BE SURE TO WRITE YOUR NAME ON THE ANSWER SHEETS PROVIDED.

POSTTEST: REVISING INSTRUCTIONAL PROGRAMS

"Ecology in Action"

PRODUCT SPECIFICATIONS

Target Population: American homemaker

Description of Materials:

1. Television broadcasts (in rough form) covering the following topics:
   a. Crises Past: An Overview
   b. Principles of Ecological Living
   c. Crises Present
   d. What Can We Do? (group action strategies)
   e. What Can I Do? (individual action strategies)
   f. Read the Writing on the Wall: Beating the Pinch
   g. What Have You Learned? (criterion examination)

2. Participants' self-instructional workbook, containing exercises keyed to each telecast.

3. Attitude survey forms for each participant.

Terminal Objectives:

Participant will apply principles of ecological living to family management by:

1. identifying in writing three imminent crisis situations
2. writing at least two original individual actions that would ease a specified crisis in his/her home.

Subobjectives:

1. identify five principles of ecological living
2. orally suggest two group strategies for coping with a given ecological crisis.

Terminal Objective Item Form:

Response Description: Given the conditions likely to precede three types of ecological crises, the participant will 1) identify the implied crisis and 2) write at least two original strategies for coping with each crisis at an individual level.

Content Limits: The participant will be shown portfolios of hypothetical news clippings, public service ads, and political cartoons pertaining to three crises: fuel shortage, protein supply shortage, pollution. Students will select one of the three portfolios.

Criteria: An adequate response contains these elements:

1) identification of probabl_ crisis by listing effects of current situation.
2) supporting evidence of effects taken from documents.
3) two strategies for coping with the crisis that have not been mentioned in class. A "coping" strate_y is one that
   a) utilizes solutions directly related to the cause of the crisis
   b) takes individual responsibility for alleviation of the crisis rather than depending on outside agencies for help.

Outline of Prototype Test Procedures

The broadcasts were shown to eight male and female volunteers in a homemaker's "ecology workshop" offered in an Adult Education Course. Participants viewed all broadcasts and completed workbook exercises for each. During the seventh broadcase, the participants completed an exam then responded in writing to a series of attitude survey questions about the usefulness and enjoyability of the series. Following each of the sessions, two different participants were interviewed by the instructor and gave their opinions of the evening's instruction. At the end of the course, all materials were turned in to the instructor and returned to the developers of the series.

PROTOTYPE TEST DATA

Criterion Test Results

Terminal Objective #1
Identify Crisis Situations

Participant	(3 possible) # correct identifications
Raquel	3
Christopher	3
Richard	3
Margaret	3
Elizabeth	2
Vicki	2
Dennis	3
Susan	3

Terminal Objective #2
Suggest Individual Strategies

Participant	(2 possible) # adequate strategies suggested
Raquel	2
Christopher	2
Richard	2
Margaret	1
Elizabeth	2
Vicki	2
Dennis	2
Susan	1

Subobjective Data

Subobjective #1
Identify 5 Principles

Participant	(5 possible) # principles correctly identified
Raquel	2
Christopher	3
Richard	4
Margaret	4
Elizabeth	2
Vicki	5
Dennis	3
Susan	4

Subobjective #2
Suggest Group Strategies

Participant	(2 possible) # adequate strategies suggested
Raquel	0
Christopher	1
Richard	2
Margaret	1
Elizabeth	0
Vicki	1
Dennis	2
Susan	1

Results of Attitude Survey

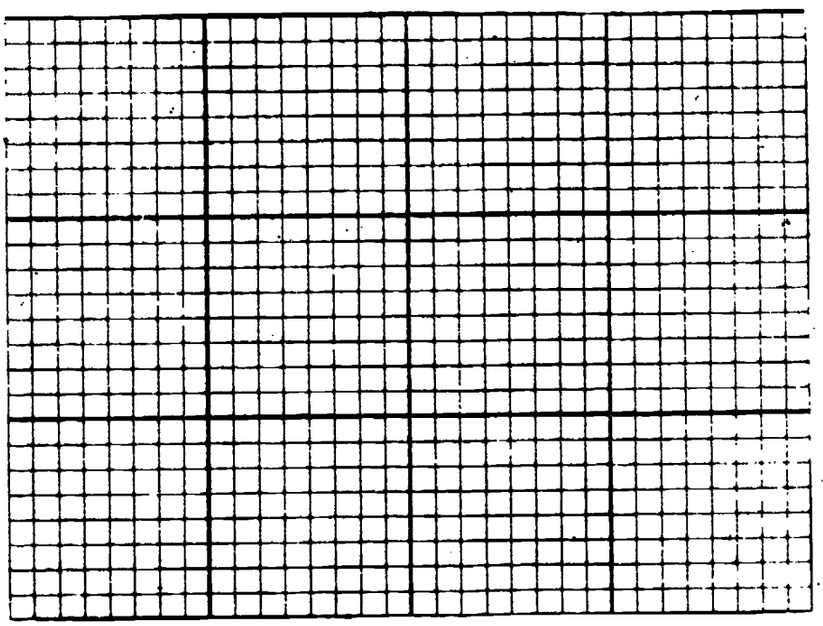
	Agree	Disagree
I learned specific ways to copy with probable crisis situations.	7	1
The programs held my interest all hour.	3	5
I would recommend the program to other homemakers.	2	6
I would like to see other programs developed by this company.	0	8

NAME_____

POSTTEST ANSWER SHEET

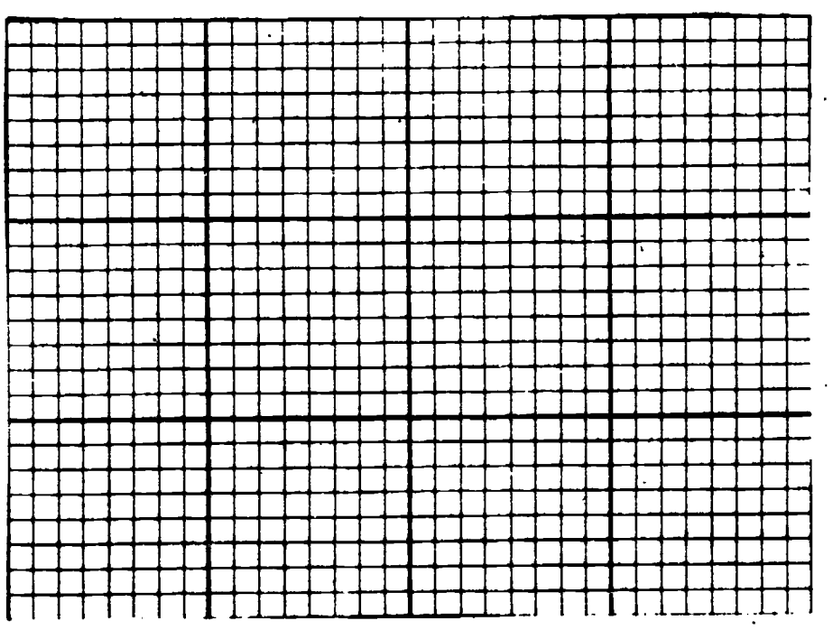
Directions: Plot terminal objective data, using frequency polygons, in the spaces provided below. Describe data patterns on the line beside each graph. Plot subobjective data in the same manner but on the back side of this page. BE SURE TO LABEL EVERY GRAPH. After you have finished, go on to the next page to suggest revision and implementation procedures.

CRITERION TEST RESULTS



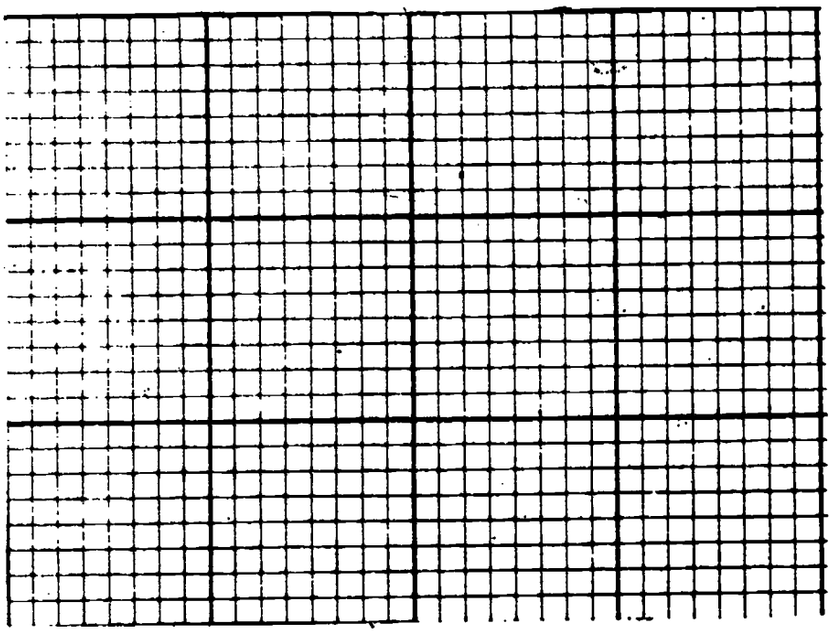
Data Pattern: _____

Data Pattern: _____

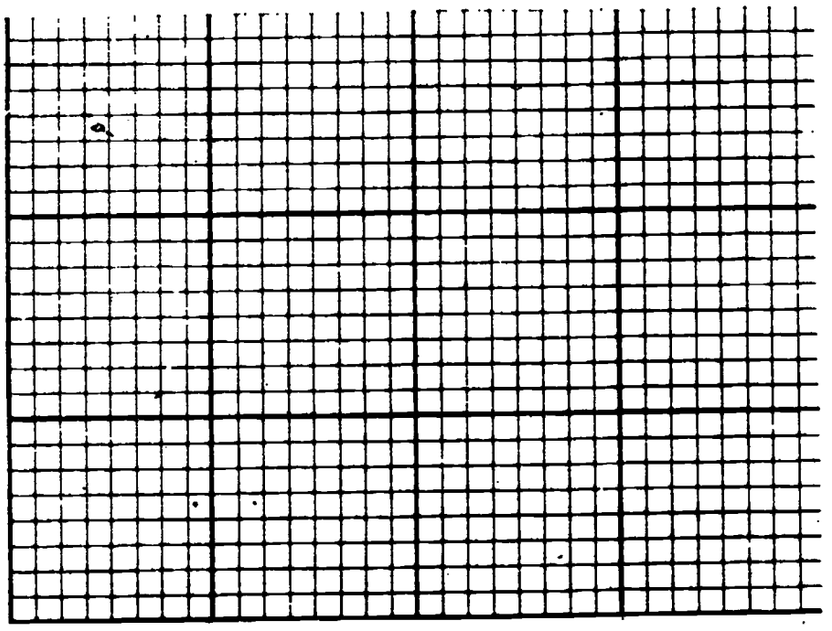


SUBOBJECTIVE RESULTS

Be sure to label graphs.



Data Pattern:



Data Pattern:

NAME_____

Posttest Answer Sheet (cont.)

2. Which research-based revision principles would you use to modify this instructional product? Mark an X by all revisions that apply.

Terminal Objective #1, Subobjective #1

____ add practice
____ review task analysis
____ verify unprompted practice
____ pool and redistribute practice and posttest items
____ add feedback
____ provide task description (assure learner understands task)
____ provide motivational stimuli
____ carefully delete irrelevancies
____ leave everything alone
____ *change format

Terminal Objective #2, Subobjective #2

____ add practice
____ review task analysis
____ verify unprompted practice
____ pool and redistribute practice and posttest items
____ add feedback
____ provide task description (assure learner understands task)
____ provide motivational stimuli
____ carefully delete irrelevancies
____ leave everything alone
____ * change format

3. How would you modify specific instructional activities listed in the specifications to apply the principle(s) you have selected above? (e.g., "rewrite the directions in the Teacher's Guide")

_____

_____

_____

_____

_____

_____

_____

(Use the other side of this page if necessary)

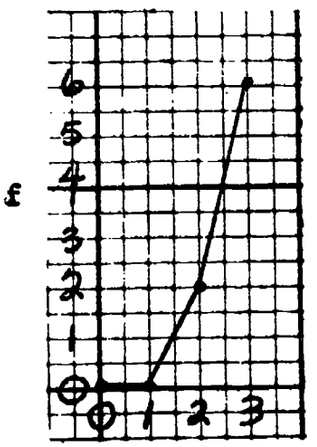
*not a research-based strategy for revision.

POSTTEST CONFIRMATION SHEET

Revising Programs

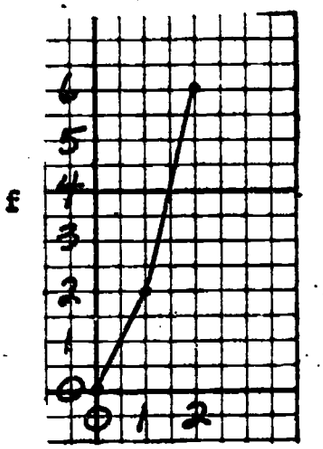
1.

Terminal Objective #1



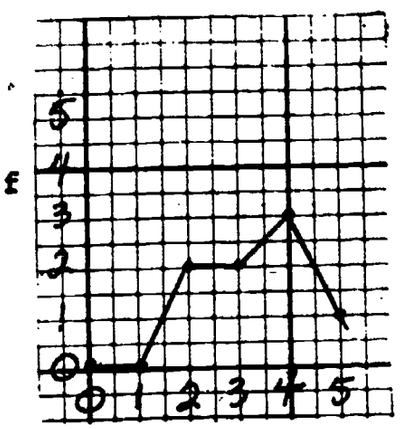
correct items
(positive data pattern)

Terminal Objective #2



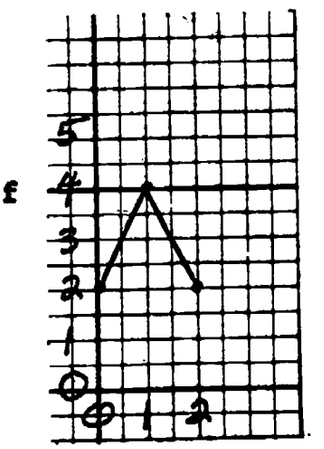
correct items
(positive data pattern)

Subobjective #1



correct items
(wide range data pattern)

Subobjective #2



correct items
(wide range data pattern)

Posttest Confirmation (cont.)

2. Although performance on Subobjectives 1 and 2 was less than outstanding, both Terminal Objectives were attained by most participants. Results of the attitude survey indicate an obvious lack of interest in the program. Therefore, your response should have been as follows:

   Terminal Objective #1, Subobjective #1

   __X__ provide motivational stimuli

   Terminal Objective #2, Subobjective #2

   __X__ provide motivational stimuli

3. Perhaps the presentation of the instruction is overly depressing due to the serious nature of the subject matter. One might attempt to add humor to the broadcasts wherever possible. Another way to approach the problem of motivation might be from the standpoint of group interaction. The use of the materials included no discussion among participants and may have caused them to feel isolated or bored. A "round table" follow-up to each audiovisual presentation might help alleviate disinterest in the program.

PARTICIPANT QUESTIONNAIRE

In order to improve the materials you have used, we would appreciate your candid responses to the following questions. Please feel free to write additional comments on the back of this page.

DO NOT WRITE YOUR NAME ON THIS PAPER

A. Circle the appropriate number.

1. Concepts were clearly presented in the text.	4	3	2	1
2. New ideas were illustrated by an adequate number of examples.	4	3	2	1
3. Directions for practice exercises were clearly written and easy to follow.	4	3	2	1
4. Confirmation sheets were accurate and complete.	4	3	2	1
5. Posttests were fair and of an appropriate level of difficulty.	4	3	2	1
6. The Data Sources (yellow) Section held my attention.	4	3	2	1
7. The Data Summary (buff) Section held my attention.	4	3	2	1
8. The Revising Programs (blue) Section held my attention.	4	3	2	1
9. I feel the information presented in this program will be useful to me at a later time.	4	3	2	1

B. Briefly answer the following questions:

10. In what ways do you think these materials can be improved?

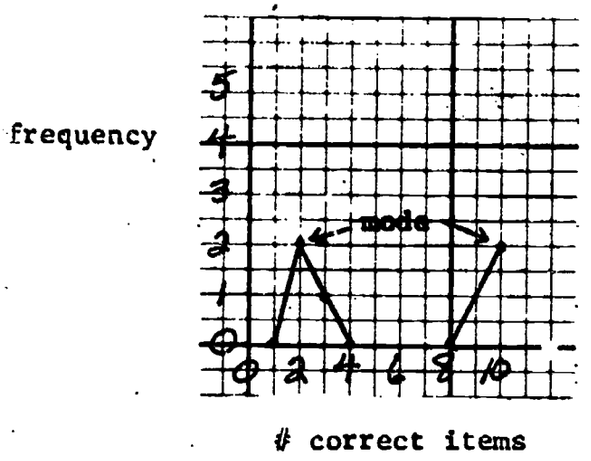
11. What do you consider to be the strong points of this program?

12. Would you recommend these materials to another educator? Why or why not?

III. GLOSSARY

GLOSSARY

bi-modal distribution: a tally or graph that has two modes, that is, two scores which occur most frequently in a distribution. For example:



component: a part of an instructional product such as a text, game, film, teacher's guide, answer book, etc.

criterion test: a test given at the end of an instructional sequence to measure student performance on the objectives of that sequence. Items on the test are keyed to specified objectives stated in terms of student behavior.

data: information about the effectiveness of a program. Data may be quantitative, as in the form of test scores, or qualitative, as in student comments, written responses to questionnaires or observer reports.

data sources: the types of information acquired about a product.

data conflicts: information about the effects of a program that indicate an apparent discrepancy in the instruction. For example, posttest scores may indicate that student achievement is high, while attitude survey shows the students detested the program.

entry skills: those behaviors which a student must have mastered prior to the onset of an instructional sequence. Entry skills for a program that teaches students to multiply one digit numerals might be that students can count to 100 and add any numerals whose sum is less than 100.

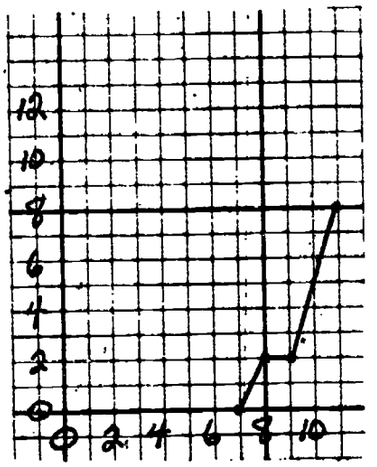
feedback: information given to a student about the adequacy of his/her responses. Feedback can be as simple as "that's right" or as complex as a detailed theoretical explanation of an algebraic postulate.

formative evaluation: gathering information about the adequacy of a program for the purpose of revising and improving the program.

Glossary (continued)

frequency polygon: a simple line graph which indicates the number (frequency) of correct student responses to a quantity of items. For example:

number of students (frequency)

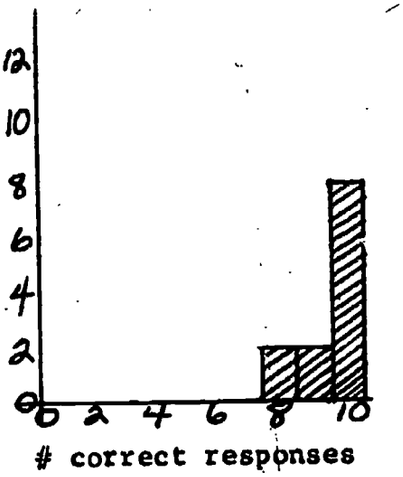


correct responses

2 students scored 8
2 students scored 9
2 students scored 10

histogram: a simple bar graph often used to display the same information as a frequency polygon. For example:

number of students (frequency)



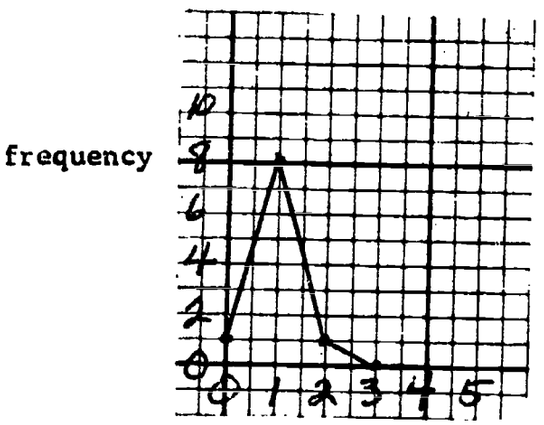
correct responses

2 students scored 8
2 students scored 9
8 students scored 10

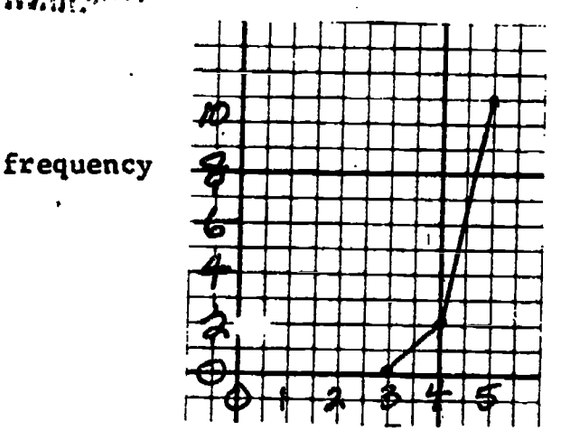
Glossary (continued)

homogeneous data: information about the effects of a program which is all of a similar nature. For example:

Graph A



Graph B



Graph A exemplifies homogeneously negative data.
Graph B exemplifies homogeneously positive data.

instructional development: a process of creating educational materials that is based on a cycle of tryout and revision. The framework of this process might be summarized as follows: 1) write a description of what the materials will teach and the methods to be used in achieving the objectives stated (specifications), 2) create a test to measure the objectives, 3) try out the test to ensure students cannot accomplish the objectives without instruction, 4) revise the criterion test if necessary, 5) generate a rough draft of the materials, 6) try out the rough materials to determine if they are effective in teaching the objectives, 7) revise the materials as needed, 8) try out the revised materials with a larger group of subjects to ensure they are effective with the type of student for which they are designed, 9) revise and retest if necessary, 10) disseminate.

item form: a model that shows how items on a criterion test are constructed so that another developer can create equivalent test items. Item forms include a description of an acceptable student response, the circumstances under which the student will be tested, the limits of the content of the item, and ways in which an instructor would differentiate correct from incorrect responses (criteria).

lean strategy for data collection: a process by which a product developer focuses on acquiring only that information which is absolutely necessary to the improvement of a set of materials, e.g., using small numbers of students during prototype testing, collecting only that data which will be relevant to revision, focusing on specific modification decisions to be made about the materials.

Glossary (continued)

mean: the average; the sum of scores divided by the number of scores.

measure: instrument used to obtain information about the effectiveness of a product. For example: tests, questionnaires, observation forms, interview blanks.

objective: a statement of what the learner is to do as a result of an instructional sequence.

en route objective: a statement of what a learner must be able to do as a prerequisite to accomplishing a terminal objective.

subobjective: same as an en route objective.

terminal objective: a statement of what a learner must be able to do as the final result of an instructional sequence

outcome: results of instruction, planned or unplanned.

practice: student trials of the skills required by the objectives of a program. Practice may take the form of written exercises, role-playing, oral recitation, etc. All practice should be either an analagous or identical to the task which will ultimately be performed by the student at the conclusion of instruction.

prerequisite skills: same as entry skills or entry behaviors.

product: any set of instructional materials with specified objectives, provision for learner practice and some form of criterion test.

product developer: a person who directs work on or creates instructional products.

prototype test: a tryout of rough draft products or parts of products (components) using a small number of subjects. The purpose of a prototype test is to determine whether the learning sequence that has been designed is functioning.

reactive pretest: a test given prior to instruction which affects later student performance. For example, a pretest which assesses student attitudes toward the use of drugs may sensitize participants to the information the developers wish them to acquire. Student responses in latter parts of the instruction may be affected by the students' attempts to respond in the way which is perceived to be "expected."

research-based instructional principles: generalizations about how people learn which are inferred from empirical (experimental) research.

replicable: able to be reproduced in another setting. Instruction products should be replicable, in that their effectiveness should not be tied to a single teacher or classroom.

Glossary (continued)

specifications: a description of the plans for an instructional product (see "instructional development."

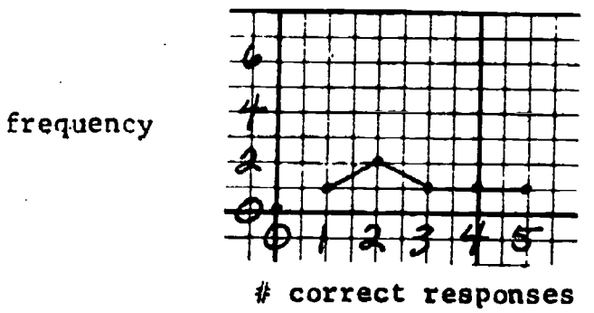
subjects: the group of learners with which an instructional product is tested. Subjects should be comparable in age and experience to the group for whom the materials are ultimately intended (target population).

summative evaluation: an assessment of the adequacy of an instructional sequence, with no intention of revision.

target population: the group of learners for which a certain set of materials are designed. A target population may be any group from "housewives" to "third year medical students specializing in endemic diseases."

task analysis: a process by which an objective is "broken down" into all the subtasks that might be necessary to master the final behavior. The subtasks ideally are verified by testing the sequence to ensure that no essential "steps" are missing.

wide-range data: information which suggests no single trend in results; the opposite of homogeneous data. For example:



No positive or negative trend is indicated by the scores--they are distributed across the graph.

within-program responses: answers provided by students during an instructional sequence as they practice for a terminal objective. Within-program responses may assume various forms, depending upon the nature of the practice provided, e.g., discussion questions, written exercises, etc.

References

Baker, Eva and Quellmalz, Edys. Research-Based Techniques for Instructional Design, National Center for Educational Research and Development, U.S. Office of Education, Washington, D.C., 1972.

Holland. James G., Doran, Judith, and Frezza, D.A. "The Use of Le rning Principles in Instruction." Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, 1974.

Report of Panel 8, Conference on Teaching, National Institute for Education, Washington, D.C., June 1974.