DOCUMENT RESUME

ED 040 963 SP 004 012

AUTHOR	Baker, Eva L.
TITLE	An Experimental Appraisal of the Generalizability of Empirical Revision Procedures for Curriculum Materials.
PUB DATE	Mar 70
NOTE	10p.; Paper presented at annual meeting, AERA, Minneapolis, 1970
EDRS PRICE	EDRS Price MF-$0.25 HC Not Available from EDRS.
DESCRIPTORS	*Material Development, *Programed Materials, Programing

ABSTRACT

An experiment was conducted to determine if the use of rules to guide revisions has generalizable effects across programs and programers. The hypothesis was that pupil achievement would be greater for instructional programs revised according to specified rules than for unrevised programs. Each of 10 programers produced a self-instructional sequence for fifth graders. The 50-frame programs including pretest and 10-item posttests were to take approximately one hour to complete. After each was administered to two fifth graders the first draft programs with technical performance data were randomly distributed to the programers with a set of rules for use in revision. Then 64 fifth graders were randomly assigned to receive either a first draft or revised program. Scores were averaged for each condition, first draft or revision, for each program and the mean for each program per condition provided the unit of analysis. A t test based on the paired observation model was employed using the S's mean performance per condition, matched by program, as the entry. An omega square value to test the strength of association was computed. It was inferred from the results that empirical revision based on rules pays off modestly, i.e., instructions can be provided so that the efforts expanded in revision will have some effect on the obtained levels of pupil performance. Confidence limits for the difference were calculated so that cost of revision could be assessed. [Not available in hardcopy due to marginal legibility of original document.] (JS)

Su

PROCESS WITH MICROFICHE AND PUBLISHER'S PRICES. MICROFICHE REPRODUCTION ONLY.

PAPER PRESENTED AT THE ANNUAL MEETING OF THE AMERICAN RESEARCH ASSOCIATION, MARCH 2-6, 1970
MINNEAPOLIS, MINNESOTA

An Experimental Appraisal of the Generalizability of Empirical Revision Procedures for Curriculum Materials

EVA L. BAKER
Graduate School of Education
University of California, Los Angeles

Educational practice is increasing its dependence upon replicable instructional products for use in the schools. Without waiting for research findings to filter haphazardly into teaching routines, products can substantially alter instruction by translating research into immediate application and solution to a variety of learning problems. Replicable products derive their impact in that they are based upon learning principles analyzed in laboratory settings. For example, the use of confirmation, practice and prompting is frequently observed, and each of these principles has a growing literature of support.

A primary tenet of product development (Markle, 1967) is the employment of empirical revision for program improvement. The use of revision assumes that effects of the usual first draft effort of a given program will not be sufficient to allow the student mastery. Validation of the characteristic iterative of revising practice has not been adequately studied. Silberman, et al., (1964) investigated factors associated with program revision processes and identified particular attributes of program revisions related to pupil learning. Robeck (1965) found that developmental testing improved the effects of a subsequent draft of a program. Gropper and Lumsdaine (1961) conducted an early study in which changes based on data from a first trial resulted in improved performance on a revision. A difficulty with the investigations is that they lack generalizability of effect. In most cases, the investigators themselves made the revisions on relatively intuitive grounds. Effects may be attributed to artistry rather than the power of a variable. Even where the revision rules were eventually articulated, as in Silberman's study, there was no test of the applicability of the rules to other situations and programs.

In order to determine if the use of rules to guide revisions has generalizable effects across programs and programmers, an experiment was conducted. The hypothesis was that pupil achievement would be greater for instructional programs revised according to specified rules than for unrevised programs.

PROCEDURE

Developmental Phase

Ten programmers involved in a graduate course in programmed instruction formed the sample. Each programmer produced a self-instructional sequence designed to be administered to a child at the fifth grade level. Programmers were told to develop their programs in accordance with instruction (Markle, 1964) received in class, providing practice, confirming, response and so on. Each program was to include at least 50 frames and take approximately an hour to complete. Any topic might be chosen by the programmers, although directors were given to select objectives which required application of principles rather than simple recall. Each programmer was also required to prepare a pretest measuring stated entry and enroute behaviors, and a posttest consisting of at least ten items. Ten first draft programs[2] were produced and each was individually administered to two different fifth grade students. Data from this developmental testing were summarized, and the report submitted by each programmer included entry skills, error rates, and criterion performance of the two learners.

Revision Treatment

First draft programs with technical performance data were then randomly distributed to the programmers; no programmer received his own program to revise. A set of rules was provided for the programmers to use in attempting a revision. Samples of such rules are listed below:

A. WHEN ERROR RATES ARE LOW AND CRITERION PERFORMANCE POOR:

1. Check to see that the desired behaviors are practiced in the program.
   ACTION: ADD PRACTICE

2. If the behaviors are practiced, check to see that the frames are not overprompted and that prompts are faded, so that practice in the criterion behavior is provided.
   ACTION: FADE PROMPTS

[2]Topics of the programs were: Sensory Imagery, Roles and Role Conflict, Good Teeth Care, Reflexive Pronouns, Hebrew, Digestion, Advertising Techniques, Navigation, Simile and Metaphor and Rules for Tutors.

3. If the behaviors are practiced and prompts faded, check to see that difficulty of content appears to be equivalent in both program practice and criterion test.
ACTION: PROVIDE PRACTICE WITH EQUIVALENT DIFFICULTY

B. WHEN ERROR RATE IS HIGH AND CRITERION PERFORMANCE POOR:

1. Re-analyze the task into component skills to make sure that en route and entry behaviors are not omitted.
ACTION: ADD FRAMES TO PROMOTE MASTERY OF RELEVANT SUBTASKS.

2. Make sure student has adequate directions for the tasks expected. Add instructional cues, or statements of the rule guiding the kind of response the student is to make.
ACTION: ADD INSTRUCTIONAL CUE
ADD STATEMENT OF INSTRUCTIONAL OBJECTIVES AT OUTSET OF PROGRAM

3. Review program to determine if the programmer has attempted to "motivate" the learner.
ACTION: ADD "REINFORCERS" DIRECTLY TO THE PROGRAM OR AS A CONTINGENCY

Programmers were told that revisions would subsequently be tested.

Experimental Phase

An experiment was conducted to assess the differential learner performance of fifth grade students subjected to first draft or revised programs.

Subjects: Sixty-four students in fifth grade classes at two schools were the subjects of the experiment. Subjects were randomly assigned to receive either a first draft or revised program. At least three children were treated by either version of each program.

Procedure: Students were assembled in a large classroom and programs were distributed at random. Following a forty minute instructional period, the children were posttested on the behaviors taught in their programs. Children exposed to programs on the same topic, e.g., sensory imagery, received the same posttest.

Criterion Test: Each programmer had been directed to supply at least ten criterion items to serve as a posttest. When the test supplied exceeded ten items, superfluous items were randomly discarded from the test, using a table of random numbers. A single criterion test was used for each pair of programs.

Analysis: The scores were averaged for each condition, first draft or revised, for each program. The man for each program per condition provided the unit of analysis. Programs which reached a 90% criterion level in the first draft form were dropped from the analysis, since there would be no reason to revise them. Three programs, with mean performances of 97%, 100%, and 90% were thus deleted. A t test based on the paired observation model was employed using the S's mean performance per condition, matched by program, as the entry. Data are reported in Table I below:

Insert Table I about here

As recommended by Hays (1963) an omega square value to test the strength of association was computed

---

Insert Table II about here

The relatively large omega square value might be expected given the sample size so small. From these results one could infer that empirical revision based on rules pays off modestly, that is, instructions can be provided so that the efforts expended in revision will have some effect on the obtained levels of pupil performance. Since the size of the difference attributable to revision is critical to assess the cost of revision, the confidence limits for the difference were calculated.

Insert Table III about here

Even though the absolute size of the difference is itself small, the difficulty of attaching a value to any increment of pupil learning must be faced. By converting the difference to standard deviation units, one can see that the mean increment obtained is equal to approximately .68 standard deviation units on the first draft distribution. If such a stable improvement could be predicted across subject matter fields, based on the application of these or a similar set of rules, the use of revision cycles could be evaluated in terms of their cost.

REFERENCES

Glass, Gene V. and Hakstian, Ralph, "Measures of association comparative experiments." American Educational Research Journal, Vol. VI, No. 3, May, 1969, pp. 403-414.

Gropper, G. L. and Lumsdaine, A.A., The Use of Student Response to Improved Television Instruction: An Overview. Report No. 7, (Summary of six prior reports), Studies in Televised Instruction, American Institute for Research, Report No. AIR-C13-61-FR-245 (VTI) Pittsburgh, 1961.

Hays, William L., Statistics for Psychologists. Holt, Rinehart and Winston, New York, 1963, p. 327.

Lumsdaine, A.A., Handbook of Research on Teaching. Chapter 12, N.L. Gare, Editor, Rand McNally, 1963.

Markle, Susan M., "Empirical testing of programs," National Society for the Study of Education Yearbook on Programmed Instruction. Editor: Phil Lange, 1967.

Markle, Susan M., Good Frames and Bad, John Wiley, New York, 1964.

Robeck, M., "A study of the revision process in programmed instruction." Unpublished Masters Thesis, University of California, Los Angeles, 1965.

Silberman, H. Coulson, Melaragno, J.R., and Newmark, G., Use of Exploratory Research and Individual Tutoring Techniques for the Development of Programming Methods and Theory. Final report NDEA Project 7-14-0000-181. System Development Corporation, Santa Monica, California, 1964.

Table I

t Test for Differences Between Pairs of First Draft and Revised Programs

Revision		First Draft		
$\bar{X}$	s	$\bar{X}$	s	*t*
5.37	2.341	3.74	2.42	4.464*

*$p < .01$

Table II

ω^2 Value for Strength of Association

$N_1 = 6$

$N_2 = 6$

$t = 4.464$

$\omega^2 = .5928$

Table III

95% Confidence Interval for the Difference

M_D	s_{M_D}	$\underline{t}_{(95/2;6)}$	Interval
1.63	.365	2.447	$\pm$ .893