USER ORIENTED PRODUCT EVALUATION*

By

Marvin C. Alkin and Joseph Wingard

CSE Report No. 81

September, 1972

Evaluation Theory Program
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California

*Based on a paper presented at a symposium entitled "Evaluation of Instructional Systems: Current Problems Facing Developers," American Psychological Association, Honolulu, Hawaii, September 3, 1972.

Introduction

Instructional product development as a technology and commercial enterprise is the progeny of the programmed instruction and curriculum development movements of the late fifties and early sixties. To say that the field has expanded since then demeans the art of understatement. Based on published listings in the Educational Product Information Exchange (EPIE), including only the more popular products, there has been an estimated 48% increase in the number of commercially available products, and an estimated 88% increase in the number of commercial producers from 1967 to 1970. Academically, we have not only witnessed the establishment of over 20 regional laboratories and R&D centers in recent years, but the educational literature is annually deluged with articles describing various schemes and systems models for product development procedures (e.g., Banathy, 1968; Briggs, 1970; Flanagan, 1967; Glaser, 1966a, 1966b; Mager & Beach, 1967; McNeil, 1968; Stowe, 1969). The complexity of such models vary considerably from general three- and six-stage schemes (e.g., Gilbert, 1962; Schutz, 1970) to comprehensive step-by-step procedural checklists (e.g., Borg & Hood, 1969).

The primary focus of much of the past and current R&D activity has been on refining and improving product development technology. This quest for clearly defined and documented development procedures has undoubtedly resulted in the more efficient production of improved instructional products. However, an important area of the R&D effort, the specification of generally accepted standards for reporting the results of these products to potential users, has been neglected. Consequently, there is a paucity of conveniently assembled and readily interpretable information that would enable users to make accurate and informed evaluations of different, but comparable, instructional products.

Types of Validation Data

In order to further clarify the problem, it is important to identify the types of validation data that should be reported as a minimum to users in order to assist them to intellectually select among alternatives instructional products. One document addressing itself to considerations of this type is a monograph by Tyler, Klein, and Michael (1971). This publication contains a set of recommendations labeled "desirable" or "essential" for use by those involved in the selection of curriculum and instructional materials. The key distinction between the intent of that document and the focus of this paper is that Tyler, Klein, and Michael address themselves to the full range of curriculum and instructional program decisions. This range extends from the initial choice of objectives for educational programs to the consideration of appropriate channels for dissemination. Thus, while their monograph provides important input to this paper, its range of concern is much broader and there is ambiguity as to the audiences for whom the publication is intended. That is, there is some question as to whether the documents is a guide to classroom teachers, school district selection committees, instructional materials producers, instructional materials disseminators, or all of them. In this paper, we are concerned only with the problem of what should be included as a minimum in an evaluation report to users of instructional products.

Common sense and experience suggest that the potential user will be most interested in validation data which delineates the instructional context of a product so that he can determine whether it is appropriate or inappropriate for his particular situation. The instructional context refers to the product's purposes, target population, effectiveness, and efficiency of installation and management. Each of these aspects of the instructional context is discussed in greater detail below.

Product Purpose

Perhaps the first question a potential user will ask about a product is, "What can my students learn from these materials?" Stated another way, the user wants to know what the product's criterion goals and objectives are so that he can determine whether they are congruent with those previously identified as important in his particular instructional situation. Thus a product which contains a set of precisely stated objectives is more valuable to the potential user than one which consists only of broad statements of philosophical or instructional intent.

Precise statements of instructional objectives not only permit users to assess whether a product is appropriate, but they also enable him to judge the extent to which it can fulfill the requirements of a given instructional situation. For example, suppose a school district wanted to purchase a product that was designed to instruct students to write simple, compound, and complex English sentences. A careful study of a product's objectives might reveal an emphasis on the structure and form of sentences rather than on punctuation. If the district were also concerned with the punctuation, it might have to consider the purchase of supplementary resources to ensure that all important objectives would be taken into account. Without the specific statement of objectives, the district could not determine the true purpose and utility of the product.

It is also important to the user that the product include descriptions of the types of instruments that are used to assess attainment of the criterion objectives. There is a definite advantage to knowing how and to what extent student success will be assessed for a given product. In addition, it is important to consider whether such assessment involves commercially produced tests that can be hand-scored and locally analyzed, or require machine-scoring facilities.

Target Population

An adequate description of the target population for the product should specify the background characteristics of learners who might be expected to achieve the goals and objectives of instruction. These characteristics could include age, sex, prerequisite skills, and socio-economic status. For example, in a bilingual school district with a large number of Spanish-speaking students, it would be advantageous to know the linguistic demands of a commercially produced social studies product before it is purchased and installed. Similarly, if ethnicity is a serious district concern, it would also be beneficial to determine in advance the relevance of the objectives of different products to the students' cultural interests.

Information about the selection of media for instructional products is of interest to the user, but not as a basis for the comparison of different media produced findings. Earlier research focusing on specific comparisons of film or videotape X versus book Y has generally yielded little that would substantially assist the user in selecting among different products. In fact, Edling (1968) concludes that "there is little evidence to support the concept that given media, qua media, contribute to or better learning than other media." The utility of media selection information to the user is primarily to clarify equipment requirements for successful product installation and management. For example, procurement and general maintenance of the appropriate audio-visual equipment necessary to sustain a film-based instructional product would be an important financial consideration for a district with limited resources.

Another important consideration related to the efficiency of installation and management is the exportability of the instructional product.

Potential users would benefit significantly from information about the product's ability to yield reliable student outcomes given specified instructional procedures, but assuming some inevitable deviations from these procedures in many school situations. The latter circumstance is particularly important because, regrettably, instructional programs are rarely implemented exactly as planned. Thus, it would be advantageous for the user to have some idea of the product's flexibility prior to its purchase and district-wide installation. For example, a user with an eye toward delays in equipment deliveries and school-wide flu epidemics might want to know how much time can be cut from the product's instructional units without interfering with the attainment of desired student outcomes.

Product Effectiveness

Once the user has determined whether the product aims at achieving significant goals for students who are like those for whom he is responsible, he wants to know how effectively the product performs. Basically, he needs data about the success of the product. Such data should be displayed and summarized so that it is readily interpretable. For example, three types of data might be provided:

1. The students' average score for each objective. This will provide the user with relatively precise information about the product's success.

2. The number (or percentage) of students who achieve a given objective. This will provide general information.

3. The number of students who achieved all or a given percentage of objectives. This will provide information about the overall success of the product, including some information about the rate at which the objectives were achieved.

Product Efficiency

Efficiency considerations demand that a commercially developed product should either demonstrably increase the number of students who are successful by current standards of success within a given time period, or drastically reduce the cost of maintaining the current success rate (E. Baker, 1972). For example, if 60 percent of all sixth graders in a given district are successful with the current math curriculum, then the district might require that a new instructional product in math achieve 85 percent student success. Alternatively, if the current 60 percent success rate costs the district $200 per student per annum, then it might demand that the new product achieve the same 60 percent rate of success, but at an annual expenditure of $100 per student. While it is impossible to provide precise cost data for districts in which a product has never been used, it is feasible to make accurate cost-estimates based on field test and experimental installation of the product in other comparable districts. As additional districts purchase and install a product, more of this kind of information would be made available which developers should compile and present to other potential users as evidence of the product's proven utility.

The preceding discussion of the types of validation data needed to define a product's context is not intended to suggest that product developers are totally negligent in compiling and presenting information to potential users. An examination of the product development literature indicates that much of these validation data are currently available. However, they are simply not collected and reported to users in an easily interpretable form. For example, while many producers of instructional

products do provide extensive data on product performance, these typically take the form of technical reports or working papers that describe field test and experimental installation results and procedures (e.g., Lipson, 1967; Niedermeyer & Ellis, 1970; Research for Better Schools, 1968; Resta & Hanson, 1971; Scharf, 1968; Scott, 1970). To most teachers and district administrators, this format for presenting data is anything but palatable.

Two partial but notable exceptions to this practice of reporting incomprehensible data are the systems of Quality Assurance and Integrated Instructional Management (IMS) developed by the Southwest Regional Laboratory (R.L. Baker, 1972). The former provides enroute information on various indicators of student performance and pacing, while the latter is designed to aggregate and synthesize input-output information in a manner that is understandable, comprehensive and consistent with the informational requirements of students, teachers, principals, parents, curriculum supervisors, district administrators, and development personnel. The distinct advantage of Quality Assurance and IMS is their tailored comprehensibility for different user interest groups. Their most significant disadvantage, however, is that the information is supplied for formative or developmental purposes, so that the instructional product must be purchased and installed without finalized description or explanations of the data. This means that potential users are de facto field test subjects! If a concerted effort comparable to that which produced Quality Assurance and IMS were exerted to provide information to users before the purchase of instructional products, much of the problem of providing the user with comprehensible validation data would be eliminated.

The Problem Clarified

In the preceding discussion we have seen that certain types of validation data, those which delineate the instructional context, would be particularly beneficial for potential users trying to decide among different instructional products. The lack of adequate, comprehensible validation data virtually guarantees that less accurate and less informed comparative judgments will be made by users about alternative products. As any district administrator is well aware, the unfortunate result is that substantial expenditures of district, state and federal funds are annually wasted on the purchase and installation of instructional products that later prove inappropriate or ineffective. Considerable savings undoubtedly could be made if important standards for reporting results were developed and critical validation data disseminated to users while different products are under consideration.

There are several immediately obvious reasons for the severe dearth of information that we have observed. This lack of user-oriented validation data on instructional products is in part related to the nature of the evaluation process and the perceived roles and interests of the various participants. We will elaborate this point by initially establishing our conceptualization of the nature of evaluation.

We have in other papers referred to a definition of evaluation as the process of determining the decision areas of concern and then selecting, collecting, analyzing, and providing information of relevance to decision makers to help them choose between alternatives (Alkin, 1969). It seems perfectly clear to us that those performing the evaluation of product development projects are simply concerned with presenting their evaluation findings to a different class of decision makers than to potential users.

Now let us examine the nature of evaluation activities in order to further clarify this point. Nearly everyone is by now familiar with the distinction made by Michael Scriven between formative and summative evaluation activities. In the model of evaluation proposed by the Center for the Study of Evaluation (Alkin, 1971), we further distinguish formative evaluations into two activities referred to as implementation evaluation and progress evaluation. The purpose of an implementation evaluation is to determine the extent to which the program has been put into operation in the manner in which it was proposed for the group for whom it was intended. A progress evaluation is conducted to determine the extent to which the program's interim objectives have been attained.

In our view, the purpose of formative evaluation is to provide information to decision makers in order to enable them to modify and improve programs. The purpose of summative (or outcome) evaluation is to make final judgments about the worth of a program prior to its introduction, marketing, or further distribution. From this set of definitions, it is obvious that a large part of the activity of product development deals with formative evaluation. Indeed, the role of the evaluator is generally not unique or separate from his role in conducting the formative activities of product development. In fact, formative evaluation is an inherent part of the product developer's role. Because product development and formative evaluation go hand in hand, they are frequently performed by the same people. This of course greatly facilitates information provision and subsequent decision making.

Now, let us examine the role of summative or outcome evaluation in the product development cycle. Who is the decision maker and what are his information needs? One decision maker is the project director; another

is the director of the laboratory, center, or development agency. Another decision maker (or set of decision makers) are individuals at the Office of Education or at other funding agencies. Now, what is the kind of information that these decision makers require and demand? Note that the summative evaluation report, since it serves at least these three decision makers or decision audiences simultaneously, must therefore be addressed to those issues which are of interest to all three sources. Thus (1) The purpose of the evaluation is to demonstrate that the product developer has been successful. (2) The purpose of the evaluation is to demonstrate that the materials developed are in some way better for producing some intended set of objectives than competing materials, or that there are no competing materials available for the stipulated set of objectives. (3) The purpose of the evaluation is primarily to demonstrate the overall quality of a product rather than to specifically delineate the conditions under which the product is most useful.

Now let us consider another class of decision maker, namely the potential user -- the teacher in Honolulu or Pittsburgh, as well as superintendents, and those having the responsibility for materials selection. Suppose these individuals were to commission an evaluation report to examine the various instructional materials available and to indicate their appropriateness for the specific setting of a particular school district, school, or classroom. Would the previous evaluation report prepared for project directors or funding agencies achieve the decision purposes just described? Most certainly not. The nature of an evaluation is determined by the decision context and the kinds of decisions that will be made on the basis of that evaluation report.

Product developers in regional laboratories and R&D centers, and in other educational institutions do have in their possession much of the necessary validation data. Surely one would say that much of the data available

in other evaluation reports would be relevant and useful to potential instructional product users. Indeed, as we have said, this information represents an integral aspect of the product development process (see Markle, 1967). However, the overriding and myopic concern of most R&D efforts is to improve and refine product development as a technology, not to collect and disseminate interpretable information to users. Thus, summative validation data clearly delineating the instructional context of an educational product are neither synthesized nor reported in any single place or in any comprehensible form.

But, if not the material producers, who else feels the responsibility for providing a full range of information for these users? Certainly not the publishing companies. Supplying such information to users of instructional products represents additional expenditures that are currently considered unnecessary and to some extent undesirable. Why raise a question about the validity of one's product for a variety of instructional contexts if you can get away without raising the issue at all? It simply is not good business to reduce profit margins by collecting and reporting information that the consumer himself has yet to ask for.

And that seems to be the case; consumers do not demand contextual validation data. Potential users of instructional products have not become sufficiently sophisticated to call for and expect comprehensive and understandable data before making their decisions. User decisions are most commonly based on the reputation of the commercial publisher or the presumed academic qualifications of the product's author. Should the latter include a doctoral degree and university or college affiliation, it is taken as evidence for the quality and effectiveness of the product.

To summarize, the lack of adequate reporting of validation data on instructional products is partially related to the myopia of product developers/cum evaluators who view the development process as the concern of decision makers who are not necessarily users, is partially a function of commercial publishers' priorities, and is partially due to a lack of consumer demand.

References

Alkin, M.C. Evaluation theory development. Evaluation Comment, 1969, 2(1), 2-7.

Alkin, M.C. A theory of evaluation. CSE Working Paper No. 18. Los Angeles: Center for the Study of Evaluation, August, 1971.

Baker, E.L. The technology of instructional development. Unpublished manuscript, UCLA Graduate School of Education, University of California, Los Angeles, 1972.

Baker, R.L. Measurement considerations in instructional product development. Unpublished manuscript, Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, California, 1972.

Banathy, B.H. Instructional systems. Palo Alto, California: Fearson Publishers, 1968.

Borg, W.R., & Hood, P. The twenty-seven steps in the development program. Unpublished manuscript, Far West Laboratory for Educational Research and Development, Berkeley, California, 1969.

Briggs, L. Handbook of procedures for the design of instruction. Pittsburg: American Institute for Research, 1970.

Edling, J.V. Educational objectives and educational media. Review of Educational Research, 1968, 38, 177-194.

Flanagan, J.C. Functional education for the seventies. Phi Delta Kappan, 1967, 49, 27-32.

Gilbert, T.F. Mathematics: The technology of education. Journal of Mathematics, 1962, 1, 7-73.

Glaser, R. Organizations for research and development in education. Paper presented at a conference of the American Educational Research Association, 1966. (a)

Glaser, R. Psychological bases for instructional design. AV Communication Review, 1966, 14, 433-449. (b)

Lipson, J.I. Individualized instruction in elementary mathematics. Reprint 22. Pittsburgh: Learning Research and Development Center, University of Pittsburg, 1967.

Mager, R.F., & Beach, K.M. Developing vocational instruction. Palo Alto, California: Fearon Publishers, 1967.

Markle, S.M. Empirical testing of programs. In Programmed instruction. Sixth Yearbook of the National Society for the Study of Education. Part II. Chicago: University of Chicago Press, 1967.

McNeil, J.D. A perspective of developmental projects. Paper presented to the Developmental Project Guidelines Conference, Minneapolis, June, 1968.

Neidermeyer, F.C., & Ellis, P.A. The development of a tutorial program for kindergarten reading instruction. Technical Report 26. Inglewood; California: Southwest Regional Laboratory for Educational Research and Development, 1970.

Research for Better Schools, Inc. A progress report: Individually prescribed instruction. Philadelphia: Research for Better Schools, Inc., 1968.

Resta, P.E., & Hanson, R.A. Installation requirements for the SWRL first-year communication skills program: Evaluation data, 1968-69. Technical Report 31, Southwest Regional Laboratory for Educational Research and Development, 1971.

Scharf, E. Implementation of individually prescribed instruction: Summary of problem areas and possible solutions. Philadelphia:Research for Better Schools, Inc., 1968.

Schutz, R. The nature of educational development. Journal of Research and Development in Education,1970, 3(2) College of Education, University of Georgia, Athens.

Scott, R.O. Program development through formative evaluation: The SWRL instructional concepts program. Technical Report 28. Inglewood, California: Southwest Regional Laboratory for Educational Research and Development, 1970.

Stowe, R.A. Case studies in instructional development. Bloomington, Indiana: Laboratory for Educational Development, Indiana University, 1969.

Tyler, L.L., Klein, M.F., & Michael, W.B. Recommendations for curriculum and instructional materials. Los Angeles: Tyl Press, 1971.