DOCUMENT RESUME

ED 223 728 TM 820 876

AUTHOR Hall, Gene E.
TITLE Evaluation of the Delivery of Services: A Concern-Based Perspective For the Design of Evaluations. R&D Report No. 3126.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Aug 80
NOTE 26p.; Paper presented at the National Conference on Longitudinal Evaluation of Bilingual Programs (August, 1980); Some tables are marginally legible due to small print.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Bilingual Education; Change Strategies; *Educational Innovation; *Evaluation Methods; Measurement Techniques; Program Descriptions; Program Development; *Program Evaluation; Program Implementation; Time; *Validity
IDENTIFIERS *Concerns Based Adoption Model; Innovations Configuration; Levels of Use of the Innovation

ABSTRACT

Change process research had developed some practical tools and concepts which can assist in the determination of essential variables for an evaluation design. The Concerns-Based Adoption Model (CBAM) Project at the Texas Research and Development Center in particular has developed specific measures which can be applied directly: "Levels of Use of the Innovation" measure evaluates to what extent a program is actually being used, while "Innovation Configuration" describes what adaptations, if any, have been made in the program. The need to view any program within a time frame necessary for it to be considered fully installed has also been a valuable contribution from the change process research. The two CBAM measures can also be used to establish what this time period is. Taking into account the three dimensions of time for implementation, levels of use, and description of use, evaluation designs can be far more specific about what they portend to evaluate and can attribute effects of this specifically described program with a greater degree of validity. (Author/PN)

EVALUATION OF THE DELIVERY OF SERVICES: A CONCERNS-BASED PERSPECTIVE FOR THE DESIGN OF EVALUATIONS

Gene E. Hall

Research on Concerns-Based Adoption
Research and Development Center for Teacher Education
The University of Texas at Austin

R & D Report No. 3126

Paper presented at the
National Conference on Longitudinal Evaluation of Bilingual Programs
August 1980

Abstract

Bilingual education is gaining more attention as many new programs are being tested. Policy decisions on whether to continue, alter or terminate bilingual programs are often based on evaluation studies of them. The evaluation of a program as it is supposed to be is one thing, but the evaluation of the program which is actually being used is another matter entirely. Programs are very often changed by the user, sometimes so much that they are not recognizable as the original model. Evaluation designs must first of all describe exactly what program is being studied before its effects can be accurately correlated. Furthermore, evaluation designs must determine if a program, whether "pure" or "adapted," is actually being used in order to know if, indeed, there are any effects to be assessed. Unfortunately, these determinations are often not made before an evaluation is conducted, with results thus being susceptible in invalid assessment and policy decisions being made on the basis of misleading information.

Change process research has developed some practical tools and concepts which can assist in the determination of these essential variables for an evaluation design. The Concerns-Based Adoption Model Project at the Texas Research and Development Center in particular has developed specific measures which can be applied directly: Levels of Use of the Innovation measures to what extent a program is actually being used, while Innovation Configuration describes what adaptations, if any, have been made in the program. The need to view any program within a time frame necessary for it to be considered fully installed has also been a valuable contribution from the change process research. The two CBAM measures can also be used to establish what this time period is. Taking into account the three dimensions of time for implementation, levels of use, and description of use, evaluation designs can be far more specific about what they portend to evaluate and can attribute effects of this specifically described program with a greater degree of validity.

EVALUATION OF THE DELIVERY OF SERVICES:

A CONCERNS-BASED PERSPECTIVE FOR THE DESIGN OF EVALUATIONS[1]

Gene E. Hall
Research on the Improvement of Practice in Schools and Colleges
Research and Development Center for Teacher Education
The University of Texas at Austin

Program development has significantly increased in sophistication and effectiveness during the last fifteen years. Program development in bilingual education clearly reflects this general trend. There is new sophistication in terms of concepts, theories, and models for bilingual programs and the different operational programs around the country provide a rich diversity of applications of these concepts and theories to practice. During this period there has also been increasing sophistication in theories, models, and approaches to evaluation. However, evaluation as an applied science still is facing a great challenge in being able to handle effectively the complexities of these sophisticated programs.

There are some recent conceptual and methodological breakthroughs that can significantly contribute to the power of evaluation studies, some of which are from disciplines other than evaluation itself. The interdisciplinary nature of evaluation makes possible the integration of knowledge from other fields, and it is to their credit that bilingual program evaluators have been willing to incorporate approaches and findings from other areas. Research on the change process and on implementation represents one area which has been able to make valuable contributions to evaluation processes. This paper briefly explores some

---

[1]The research described herein was conducted under contract with the National Institute of Education. The opinions expressed are those of the author and do not necessarily reflect the position or policy of the National Institute of Education. No endorsement by the National Institute of Education should be inferred.

of the developments from change research which can be of help to evaluators. Two major concepts are focused upon, that of Levels of Use of the Innovation and that of Innovation Configuration. Illustrations and sample data "from the field" will be used to explain these concepts and their relative measurement techniques. The paper also explores some implications of these perspectives on change to the evaluation of bilingual programs.

Perspective

Recent research on change places emphasis on viewing change as a process, not an event. Practitioners, researchers, evaluators, and policy makers are showing increasing sensitivity and awareness to this assumption. Three key parameters of the change process research that are important to consider in the design of an evaluation study are:

1. implementation is a phase of the change process;
2. innovations are usually adapted during implementation; and
3. attention must be given to individual users and nonusers of an innovation.

Addressing these parameters has been crucial to understanding, managing, and assessing the effects of change efforts.

The selection of designs and measures for education studies should also reflect awareness of these parameters. Both summative, or product, and formative, or process evaluation have valuable roles to contribute, but the appropriate approach must be used at the appropriate time. Exploring these three variables can help define whether summative or formative designs are to be used, and which variables should be emphasized.

Change occurs over time. In a key synthesis of the literature, Fullan and Pomfret (1975) articulate the distinctions between development, dissemination,

the decision to adopt, and implementation. As they point out, implementation is the phase of the change process that occurs after the program is introduced into the classroom, building or school district. They emphasize, however, that full use of the program is not instantaneous and is not achieved without problems. Formative evaluation is particularly relevant during this phase. Program implementors need information about how the change effort is progressing. This feedback is essential if the implementation of a program is to more quickly and easily become a routine pattern of uses. Contrary to much evaluation practice, the implementation phase is a poor time to be conducting summative evaluations.

The term innovation refers to the new program or process being implemented or to key changes being made in an existing program. Recognition that the innovation can, and probably will, be adapted by different users to fit different contexts is another major finding of the change process research that has implications for evaluation. Perhaps the most well-known reference on adaptation of the innovation is the Rand change agent study (Berman & McLaughlin, 1975) in which the authors propose the concept of "mutual adaptation." The innovation is adapted to fit the local context and the local users rather than being implemented exactly as prescribed by its developers.
At the same time, the users adapt somewhat to fit the requirements of the innovation. Thus, both the innovation and the users adapt to more closely approximate the requirements of the other. The implication of this concept for evaluation is that evaluators cannot assume that the innovation as implemented at each site is a close formulation of the innovation as espoused by the program developer, or that the innovation is implemented in the same way in each site.

Individuals involved in change are a key focus of research based on the Concerns-Based Adoption Model focuses on (Hall, Wallace & Dossett, 1973). The CBAM model recognizes implementation and innovation adaptation, and then brings

an additional emphasis to the understanding of change as a process by focusing on the individual. From the CBAM perspective each individual will adopt the innovation in their own way, and in that respect each individually is considered to be either a user or a nonuser. CBAM experience suggests, that in general, administrators and other decision-makers are not reliable sources of information as to what actual classroom practice is for each individual. Therefore, evaluators should be encouraged to assess directly whether each individual is a user or nonuser.

In summary, recent change research focuses on the individual, acknowledgement of the adaptation of innovations, and the recognition of the implementation phase of the change process. Each of these parameters has important implications for the design of program evaluations. Evaluation studies cannot consider a program as an event which automatically assumes full functioning simply by being placed in the classroom. Instead, use or nonuse of the innovation must be directly assessed for each teacher. The operational form of the innovation must also be assessed individually. Furthermore, evaluation studies must take into account change as a process, thereby utilizing both formative and summative evaluations and conducting studies over time, rather than just gathering data at one or two points in time and issuing a single report.

Two Basic Questions

The literature on change emphasizes the importance of understanding what happens in individual classrooms through direct assessment. It is not possible to make interpretations about how the process of change and implementation of new programs is unfolding or to draw conclusions about the effects of new programs, without knowing first-hand what is happening in each classroom. There are two basic questions that must be asked before an evaluation study can pro-

ceed, for either process or product evaluations. The simplest way to ask these questions regarding the program being examined is:

For each teacher, school and district,

1. IS IT BEING USED?
2. WHAT IS IT?

The first question asks simply whether or not each teacher who is supposed to be using the innovation in fact is. Incidently, the first question must also be asked of any comparison groups being used in the study. The second question addresses the issue of innovation adaptation. This question asks exactly what operational form of the innovation is being used by each "user." In the Concerns-Based Adoption Model (CBAM) research, direct efforts have been made to determine use or nonuse of the innovation through a concept called Levels of Use of the Innovation. Also in the CBAM, research attempts have been made to describe the various operational forms of the innovation which could be implemented in different classrooms. The concept of Innovation Configurations has been developed to describe the phenomenon.

Is It Being Used?

Regardless of whether a process or product evaluation is being done, "is the program being used?" is the first question that evaluators need to ask with regard to any program. Although this question seems obvious and straight forward, it is alarming to review evaluation and research studies in which it has been neither directly nor indirectly asked or answered. The concept of Levels of Use has proven to be very useful in answering this question.

Levels of Use of the Innovation (Hall, Loucks, Rutherford & Newlove, 1975) describes what individuals and groups are doing or not doing with an innovation. The concept focuses on whether or not and how they are using the innovation.

The first major breakout that Levels of Use (LoU) reveals is to distinguish operationally between users and nonusers. A further breakout is possible into eight distinct Levels of Use, including three levels of nonuse. The eight Levels of Use are summarized in Figure 1.

More specifically, the nonuse levels range from the person at a Level of Use 0, Nonuse, who is not doing anything relative to the program, to a Level of Use I Orientation person who is considering or exploring use of the program but has not decided to use it, to a Level of Use II Preparation person who has made a decision to use the program and is preparing for first use. For each of these nonuse levels, behavioral indicators have been identified and a set of categories can be used to further describe the actual behaviors.

The five use levels are as follows: Level of Use III persons are at a Mechanical Level of Use--this tends to be a time that is typically observed in early users of an innovation. Use is disjointed and the person has a very short-term focus in relation to their use of the program. Later on we find that persons move on to being at a Level of Use IVA Routine. These persons have an established pattern of use and are making few, if any, adaptations in their use of the program. Some individuals move on to a Level of Use IVB Refinement where they are making adaptations which are for student benefit. Level of Use V Integration people are making adaptations in use of the innovation by collaborating with other users of the innovation so that collectively they can have greater impact. The rare Level of Use VI Renewal person is one who is making major modifications in the program or considering an alternative to the present innovation.

These eight Levels of Use have been defined in the Level of Use chart and a measurement procedure has been developed. This measurement procedure which has been demonstrated to be reliable and valid, uses a focused interview (Loucks,

Figure 1

LEVELS OF USE OF THE INNOVATION

LEVELS OF USE	DEFINITION OF USE
0 NONUSE	STATE IN WHICH THE USER HAS LITTLE OR NO KNOWLEDGE OF THE INNOVATION, NO INVOLVEMENT WITH THE INNOVATION, AND IS DOING NOTHING TOWARD BECOMING INVOLVED.
Decision Point A	Takes action to learn more detailed information about the innovation.
I ORIENTATION	STATE IN WHICH THE USER HAS RECENTLY ACQUIRED OR IS ACQUIRING INFORMATION ABOUT THE INNOVATION AND/OR HAS RECENTLY EXPLORED OR IS EXPLORING ITS VALUE ORIENTATION AND ITS DEMANDS UPON USER AND USER SYSTEM.
Decision Point B	Makes a decision to use the innovation by establishing a time to begin.
II PREPARATION	STATE IN WHICH THE USER IS PREPARING FOR FIRST USE OF THE INNOVATION.
Decision Point C	Changes, if any, and use are dominated by user needs.
III MECHANICAL USE	STATE IN WHICH THE USER FOCUSES MOST EFFORT ON THE SHORT-TERM, DAY-TO-DAY USE OF THE INNOVATION WITH LITTLE TIME FOR REFLECTION. CHANGES IN USE ARE MADE MORE TO MEET USER NEEDS THAN CLIENT NEEDS. THE USER IS PRIMARILY ENGAGED IN A STEPWISE ATTEMPT TO MASTER THE TASKS REQUIRED TO USE THE INNOVATION, OFTEN RESULTING IN DISJOINTED AND SUPERFICIAL USE.
Decision Point D-1	A rountine pattern of use is established.
IVA ROUTINE	USE OF THE INNOVATION IS STABILIZED. FEW, IF ANY, CHANGES ARE BEING MADE IN ONGOING USE. LITTLE PREPARATION OR THOUGHT IS BEING GIVEN TO IMPROVING INNOVATION USE OR ITS CONSEQUENCES.
Decision Point D-2	Changes use of the innovation based on formal or informal evaluation in order to increase client outcomes.
IVB REFINEMENT	STATE IN WHICH THE USER VARIES THE USE OF THE INNOVATION TO INCREASE THE IMPACT ON CLIENTS WITHIN THE IMMEDIATE SPHERE OF INFLUENCE. VARIATIONS ARE BASED ON KNOWLEDGE OF BOTH SHORT- AND LONG-TERM CONSEQUENCES FOR CLIENTS.
Decision Point E	Initiates changes in use of innovation based on inout of and in coordination with what colleagues are doing.
V INTEGRATION	STATE IN WHICH THE USER IS COMBINING OWN EFFORTS TO USE THE INNOVATION WITH RELATED ACTIVITIES OF COLLEAGUES TO ACHIEVE A COLLECTIVE IMPACT ON CLIENTS WITHIN THEIR COMMON SPHERE OF INFLUENCE.
Decision Point F	Begins exploring alternatives to or major modifications of the innovation presently in use.
VI RENEWAL	STATE IN WHICH THE USER REEVALUATES THE QUALITY OF USE OF THE INNOVATION, SEEKS MAJOR MODIFICATIONS OF OR ALTERNATIVES TO PRESENT INNOVATION TO ACHIEVE INCREASED IMPACT ON CLIENTS, EXAMINES NEW DEVELOPMENTS IN THE FIELD, AND EXPLORES NEW GOALS FOR SELF AND THE SYSTEM.

From: The LoU Chart. Austin: Research and Development Center for Teacher Education, The University of Texas, 1975.

Newlove & Hall, 1976). The interview is generic in terms of its process and therefore can be applied to different programs and processes. The interview takes the form of a general conversation with a teacher who is asked to describe behaviorally what they are doing or not doing with the innovation. The interview is constructed using a branching format based on the operational definitions of Levels of Use. Training manuals and systems have been developed for training evaluation staff to certifiable levels in terms of their proficiency in conducting Levels of Use interviews and rating interview data.

There are several interesting implications for evaluation that can be derived from the Levels of Use research. For example, the proportion of individuals that are found at each Level of Use is not equal (See Table 1). In a stratified sample of innovation users and nonusers based upon amount of experience with an innovation, it is commonly found that the largest sample of users are at LoU IVA Routine. Lou V and VI individuals are relatively rare in most stratified samples. Another pattern that is typically observed is that 60 to 70% of the first-year users of an innovation will be at LoU III Mechanical Use (Hall & Loucks, 1977).

One implication of this last finding is that it can be assumed with something as complicated as a bilingual program that many, if not all, first year users of the program will be at a Mechanical Level of Use. This is a time for conducting process or formative evaluations and not a time for conducting a summative evaluation. Persons at a Mechanical Level of Use may be doing a more awkward and disjointed job with the new approach than they were previously with their past practice. It is not a fair time for the individual or the new program to be assessed in terms of outcomes and ultimate effects.

A useful application of the Levels of Use concept, especially for formative or process evaluation, is to provide a benchmark for the rate and degree of

implementation which is occurring. A study using the LoU technique to monitor a three-year effort involving a revised science curriculum provides an example. A summary of this study's data is presented in Table 2. Levels of Use were assessed five times during the three-year period, while a concerns-based staff development effort was also underway (Loucks & Pratt, 1979). As part of the staff development plan, the LoU data were immediately fed back to the staff developers so that staff development experiences could be adapted according to the Levels of Use of teachers in various parts of the school district. The LoU data were also used to illustrate to district policy makers that implementation was not an event and that use was moving ahead.

Looking at the data in Table 2, several trends can be readily discerned. First of all, the number of individuals, at LoU 0 Nonuse rapidly decreases as the implementation effort progresses in time. There is also a definite trend towards use (LoU III - VI) of the revised curriculum materials. One of the interesting findings relative to the process of implementation is that there was still a large proportion of individuals at a Mechanical Level of Use after three years. This led the staff developers in this particular situation to acknowledge that implementation of any new curriculum takes longer than is typically assumed; and, furthermore, staff development to assist teachers in moving to higher Levels of Use requires a different plan than does a staff development effort to assist teachers in moving from nonuse to an initial stable pattern of use.

The LoU concept has been applied in a bilingual education program setting by Dominguez, Tumner and Jackson (1980). In their study of the implementation of bilingual education programs in urban and rural sites, they too found a distribution of Levels of Use ranging from LoU 0 Nonuse to LoU IVB Refinement. Because of the complexity of bilingual education being actually a "bundle" of

Table 1

OVERALL LoU DISTRIBUTION FOR STRATIFIED SAMPLES

LoU (IN PERCENTAGES)

	0	I	II	III	IVA	IVB	V	VI	N
TEAMING	7	3	4	9	61	13	2	3	397
MODULES	15	23	12	6	25	17	1	1	255

Table 2

Levels of Use Distribution

Percent at Each LoU

	0	I	II	III	IVA	IVB	V	VI	
Fall '76	45	9	36	3	5	1	0	0	N = 204
Spring '77	25	4	24	27	15	3	0	1	N = 202
Fall '77	6	10	25	28	22	2	6	2	N = 167
Spring '78	6	6	3	43	27	11	2	2	N = 157
Spring '79	9	1	1	29	45	12	1	3	N = 146

innovations rather than a single one, they assessed Levels of Use for various innovations within the bundle. For example, they distinguished between ESL, Spanish reading, and ESL Spanish math. Thus, they were able to provide more precision in distinguishing between users and nonusers of various components in their studies instead of simply referring to users and nonusers of a composite bilingual education program.

Another important application of the Levels of Use concept is to insure that a comparison or control group, does not have users of the materials within it. In other words, Levels of Use must also be assessed for the group that is assumed to be nonusers. In one study (Hall & Loucks, 1977), an evaluation was done of Individually Guided Education, an innovation bundle involving individualized instruction, multi-aged teaming and several other major innovations. This study observed that in the treatment schools in which the innovation of individualized instruction had supposedly been implemented for three years, only 80% of the teachers were in fact using individualized instruction. Use/nonuse data were then analyzed for the comparison schools, which were matched with eleven treatment schools. It was observed that 63% of the teachers in the comparison schools were in fact individualizing their instruction in reading! Both comparison and treatment schools had large proportions of users of a key component of the innovation.

In this type of evaluation study, it would be impossible to draw conclusions about the effects of the innovative program when in fact key specific innovations within the innovation bundle were in use in both experimental and control schools. Unfortunately, what too frequently happens in these types of evaluation designs is that rather than assessing use or nonuse, the inference is drawn based upon the statements of principals or district administrators that in fact one school is a treatment school and the other school a nonuser school.

When the achievement tests or other outcome data are collected and analyzed the conclusion is made that there are "no significant differences". Then a policy decision immediately follows that the new way is no better than the old and the new way is eliminated. When in fact, the new way and the old way have not received a fair test.

That too much is taken for granted in assessing use or nonuse seems to be an ongoing problem in the conduct of innovation studies. Moreover, there are too many evaluation studies that do not acknowledge the existence of the implementation phase. These studies continue to be based on the assumption that change is an event without the transition period involving Mechanical Level of Use. These kinds of studies further assume that once teachers have received program materials and a one or two day workshop in August, they then automatically become sophisticated users and are capable of functioning proficiently in the programs as of September. Although this may sound like a crazy approach to evaluation, it still goes on.

What Is "It"?

This basic question has also received attention in the change research. Obviously, in many instances the innovation is adapted during the adoption process. These adaptations are often appropriate and well within the scope of the original intent of the innovation as outlined by its developers. In other instances these adaptations in the innovation may be beyond assumed limits of the program. With some of these adaptations the program could conceivably have different effects than were originally intended. In CBAM, research of this phenomenon has been addressed through the concept of Innovation Configurations (Hall & Loucks, 1978). Innovation Configurations as a tool provides a means to describe what happens to the innovation as it is adopted by different users.

Again, the initial focus is upon the individual user, which can then be aggregated for the school or school district levels.

The simplest question to be asked here is, "what are the program users using?" The program will presumably contain distinctive features, which often include the use of certain materials and special roles which the students and teachers are to carry out. How does an evaluator determine the actual operational form of the innovation that is in place? Conceptually, the problem is fairly easily understood, but is a fairly complex problem methodologically. An

An innovation can be considered to consist of a series of major components of which users of the program can be assessed for thier use or nonuse. For example, a continuous-progress mathematics curriculum could consist of: 1) the components of instructional materials used, 2) the grouping of students, 3) and testing. Then in terms of use/nonuse, the simplest decision would be to judge as users of the program those teachers who were: 1) using the materials, 2) grouping students, and 3) testing them. However, the way that the materials are used, how students are grouped, and exactly how and what is done with test results can be important variations. Thus, in the concept of Innovation Configurations, the components are further divided into different "variations" of use which can be identified. This example can briefly illustrate how variations are identified within each component:

1. Instructional Materials Component

   a. Program materials only
   b. Program materials plus basic text
   c. Text only
   d. Teacher made

2. Grouping Component

   a. Large heterogeneous group
   b. Large homogeneous group
   c. Small groups
   d. Completely individualized

3. Testing Component

   a. Testing done once every six weeks with nothing done with testing results
   b. Testing done weekly with test results fed back to students
   c. Students test themselves as they complete each objective

In conducting an evaluation study, it would be the job of the evaluators to develop and complete an innovation configuration checklist (See Figure 2) that consisted of these components and their variations. For each subject, an innovation configuration checklist would be completed. A certain combination of responses across the components such as: (1-a) uses program materials only, (2-c) small groups of students, and (3-a) testing once every six weeks; would represent one "configuration" of the program. Any other combination of components would represent a different configuration. The evaluator could then summarize which configurations are in use in different classrooms. The concept of Innovation Configurations is thus used to define and describe the operational forms of the innovation. This entails describing: 1) what students and teachers are doing, 2) how materials and processes are used, and 3) what would be observed when the program is in actual use.

It is important to note that the concept of Innovation Configurations does not, however, describe implementation requirements. An implementation requirement would be possession of the textbooks. How the textbooks are used would be an Innovation Configuration component. This same rationale applies to outcomes. Outcomes and effects of a particular innovation are outside of the innovation configuration itself. They are the result of using a particular configuration.

These distinctions between implementation requirements, innovation configurations and outcomes become particularly important for the evaluator. An evaluator needs to be able to determine what effects are a result of use of the

Figure 2

Revised innovation configuration checklist for math curriculum study.

I.D. (Last 4 digits of SS#) _____ _____ _____ _____

Please check one choice under each of the five categories below that is the most descriptive of your math instruction.

1. Instructional resources used:

_____ Program materials (i.e., packets, worksheets) only
_____ Textbook(s) only
_____ Textbook(s) emphasized, program materials supplemental
_____ Program materials emphasized, textbook(s) supplemental
_____ Combination of text(s) and/or program materials with teacher-made materials and games
_____ Large variety of text(s), program materials, games, teacher-made materials, manipulatives, centers, labs, etc.

2. Grouping patterns:

_____ Teach whole class or two groups in a class with children ranging in abilities (i.e., heterogeneous)
_____ Teach whole class or two groups in a class with children of generally the same abilities (i.e., homogeneous)
_____ Teach 3 or more small groups that are fairly stable--the children in each group seldom move to a different group
_____ Teach 3 or more small groups that change continually--the children frequently move to a different group.
_____ Teach individuals only, no grouping

3. Clusters of objectives:

_____ Use program clusters largely in sequence as a framework for instruction
_____ Use program clusters largely out of sequence
_____ Do not use program clusters

4. Objectives:

_____ Use program objectives largely in sequence within the clusters
_____ Use objectives largely out of sequence
_____ Do not use program objectives

5. Kind of testing:

_____ Use posttest on objectives (either program-supplied or teacher-made) and cluster test (either program-supplied or teacher-made)
_____ Use cluster test (either program-supplied or teacher-made) only
_____ Use posttest on objectives (either program-supplied or teacher-made) only
_____ Use teacher judgement only
_____ Varies widely using posttests, cluster tests and/or teacher judgement

6. The use of test results:

_____ Each individual child is assigned work or activity depending on the results of the test given
_____ When most of the group passes a test the group goes on; those who fail are given special attention
_____ When most of the group passes a test the group goes on; those who fail will have another chance to learn later due to the spiral nature of the curriculum
_____ What is done with test results depends on the objectives being taught

ARE THERE OTHER SIGNIFICANT FEATURES OF YOUR MATH PROGRAM THAT WE HAVE NOT INCLUDED? (Please describe on back of this page)

innovation. In order to do so it must be determined first that the innovation is being used and second what configuration of the innovation is being used. As previously pointed out, it is conceivable that different configurations of the innovation will result in different outcomes. The process evaluator may want to present purely descriptive information about the different configurations that are in use so that the staff developers can plan specialized inservice which addresses the needs of teachers who are using different configurations. It is also conceivable that through early testing certain desired learning oucomes may be associated with certain configurations. These data could then been fed back so that staff development again could be targeted to facilitate implementation of those configurations which are associated with more of the desired outcomes. Hall & Loucks (1980); Pratt, Winters & George (1980) and George & Hord (1980), are papers that explore some of these implications further.

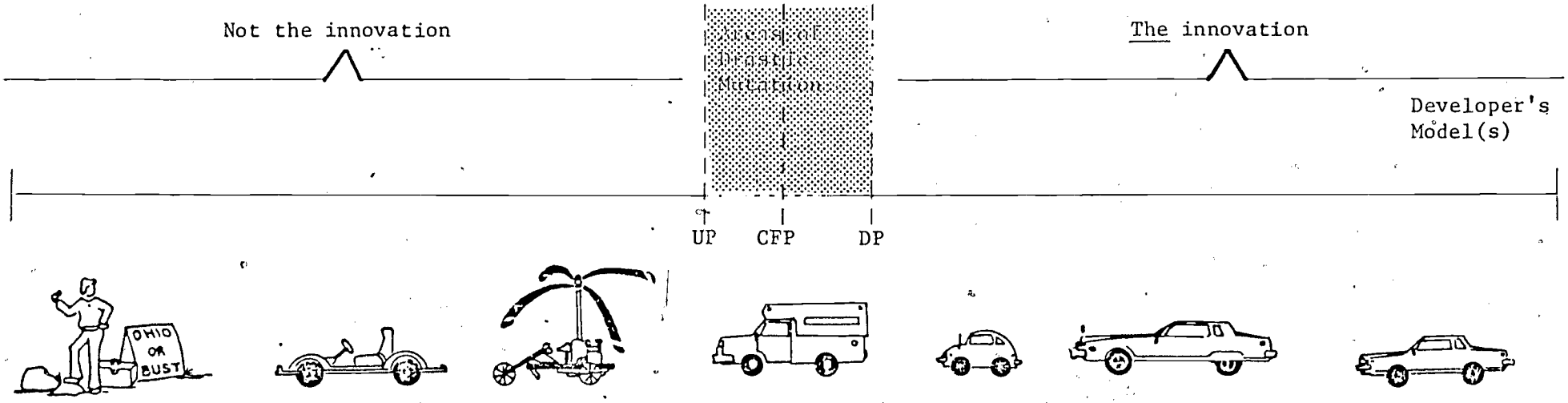
The Innovation Configuration framework has been applied in at least one study of bilingual education. Dominguez, Tumner and Jackson, (1980) identified a range of innovation configurations for bilingual education programs. Because of the elaborate nature of bilingual education, it was necessary to describe not only the configurations of bilingual education in terms of classroom practice, but also the school building and district level aspects of the configurations. Furthermore, they took advantage of the concept of an innovation configuration continuum (Hall & Loucks, 1978).

The Innovation Configuration continuum (See Figure 3) suggests that the various configurations of a program can be placed along a continuum that ranges from those that most closely approximate the developer's model, to those that are somewhat related to it and finally to those that are clearly deviant from the model as described by the developer. Using this procedure, Dominguez, et al., were able to distinguish between programs that were approximating the

Figure 3

Configuration Continuum

Using "Car" as the Innovation



Points of Drastic Mutation

UP -- User's Point

CFP -- Change Facilitator's Point

DP -- Developer's Point

ideals of bilingual education, (maintenance), those that were on their way towards it (transitional), and those that were clearly not within acceptable limits with regard to bilingual education theories and models (non-bilingual). (Dominguez et al., 1980). Many of the participants within the settings of the non-bilingual programs may have claimed, and even believed, they were using bilingual education materials. However, in terms of the criterion that were set up to describe ideal models of bilingual education, their configurations were not in the ballpark.

For the process evaluator, feedback about which configurations are in use and which components of innovations seem to be most clearly associated with desired outcomes can be of great value. The staff developers can fine tune and focus their staff development activities to address teacher needs in relation to certain components. Policy makers and administrators can understand better which resources and support to deliver so that the configurations put into use are those that are desired And all parties can be more concrete in their communications with, and understanding of each other.

In Summary

In this paper, I have attempted to lay out some issues that have come out of recent research on the change process that have direct implications for process evaluation. These implications are not limited to bilingual education but apply equally to other evaluation efforts. However, they clearly apply in bilingual education, especially given the complexity and lack of clear cut definition of what bilingual education is when it is made operational in classrooms, buildings and school districts.

As previously stated, the change research indicates that evaluators need to answer two fundamental questions before either formative or summative evaluation efforts can proceed. "Is it being used?" and "What is 'it'?".

In the past, and all too frequently in the present, evaluation studies are conducted without having a clear cut answer to either the use question or the configuration question. Consequently, the conclusion "no significant differences" may not be valid, since the evaluation design may have masked differences by not separating users of acceptable configurations of the innovation from nonusers of the innovation or users who are using unacceptable configurations. Such invalid conclusions can lead to unfortunate policy decisions.

This is not meant to say that a "high-fidelity" perspective has to be taken in all implementation efforts. It is quite acceptable in certain situations for teachers and other users to be encouraged to create all sorts of adaptations in the innovation. In that case the process evaluators need to be able to document and describe the different component variations and configurations that result. Special staff development activities would have to be developed to address and encourage teachers in creating adaptations and diversions. In either case summative evaluation studies must clearly document with first-hand data that the innovative group does indeed have users in it and that any comparison groups are comprised of nonusers.

Another implication of the change research is that evaluation must be done over time. Change is a process, so longitudinal designs are essential. Evaluation studies cannot be conducted on just the first year of implementation. Formative evaluation data, which are descriptive of the implementation process, must be constantly returned to program developers, staff developers and others concerned with implementation. A one-time data collection will not provide the

required longitudinal perspective that reflects the growth or regression which can occur as the change process unfolds.

In this paper the concept of Levels of Use of the Innovation has been proposed as a way to answer the first question of "is it being used?" The concept of Innovation Configurations has been proposed to address the second question, "what is it?" The concepts and measurement procedures have been briefly described and sample data presented to assist the reader in thinking about these concepts and their possible implications for their setting. Clearly, formative/process evaluation is a very useful tool. In dealing with a program as complicated and multi-faceted as bilingual education, it is essential to develop as much clarity as is possible about the concepts and variables to be assessed. It is only through this careful thinking that the evaluator can provide information to program developers, staff developers, users, and policy makers which will help them in their task of improving education.

References

Berman, P. & McLaughlin, M. W. Federal Programs Supporting Educational Change, Vol. IV: The Findings in Review. Santa Monica: The Rand Corporation, 1975.

Dominguez, D., Tumner, W. E., & Jackson, S. L. Measuring degree of implementation of bilingual education programs; implications for staff development and program evaluation. Paper presented at the National Association for Bilingual Education Meeting, Anaheim, California, April 1980.

Fullan, M. & Pomfret, A. Review of research on curriculum implementation. Ontario: Ontario Institute for Studies in Education, April, 1975.

George, A. A. & Hord, S. M. Monitoring curriculum implementation: Mapping teacher behaviors on a configuration continuum. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

Hall, G. E. & Loucks, S. F. A developmental model for determining whether the treatment is actually implemented. American Educational Research Journal, 1977, 14(3), 263-276.

Hall, G. E., Loucks, S. F., Rutherford, W. L. & Newlove, B. W. Levels of use of the innovation: A framework for analyzing innovation adoption. The Journal of Teacher Education, 1975, 26(1), 52-56.

Hall, G. E., Wallace, R. C. & Dossett, W. A. A developmental conceptualization of the adoption process within educational institutions. Austin: Research & Development Center for Teacher Education, The University of Texas, 1973.

Loucks, S. F. & Pratt, H. A concerns-based approach to curriculum change. Educational Leadership, 1979, 37(3), 212-215.

Loucks, S. F., Newlove, B. W. & Hall, G. E. Measuring levels of use of the innovation: A manual for trainers, interviewers and raters. Austin: Research and Development Center for Teacher Education, The University of Texas, 1975.

Pratt, H., Winters, S. & George, A. A. The effects of a concerns-based implementation plan on the achievement of elementary science students. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1980.

Provus, M. Discrepancy evaluation for educational program improvement and assessment. Berkeley, California: McCuthan, 1971.

Scriven, M. The methodology of evaluation. AERA Monography Series on Curriculum Evaluation, 1967, No. 1, 39-83.

Stake, R. E. The coutenance of educational evaluation. Teacher's College Record, 1967, 68, 523-540.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. I., Merriman, H. O., & Provus, M. M. Educational evaluation and decision making. Itasca, Illinois: F. E. Peacock 1971.