DOCUMENT RESUME

ED 063 330   TM 001 347

AUTHOR   Russell, H. H.; Leithwood, K. A.
TITLE   Evaluation and the Diffusion of Educational
              Innovations.
PUB DATE   Apr 72
NOTE   48p.; Paper presented at the annual meeting of the
              American Educational Research Association (Chicago,
              Illinois, April 1972)

EDRS PRICE   MF-$0.65 HC-$3.29
DESCRIPTCRS   Academic Performance; Communications; Cost
              Effectiveness; Criterion Referenced Tests; Curriculum
              Development; *Diffusion; Educational Accountability;
              Educational Change; *Educational Innovation;
              Educational Objectives; Educational Research;
              *Evaluation; *Goal Orientation; Interaction;
              Interinstitutional Cooperation; *Models; Norm
              Referenced Tests; Post Testing; Pretests; Program
              Development; Reliability; Teacher Responsibility;
              Validity
IDENTIFIERS   *Trent Valley Center

ABSTRACT
         The role of evaluation in a working model for the
diffusion of educational innovations is the central topic of this
paper. The working model is a product of the Trent Valley Centre
(TVC), a "regional laboratory" type extension center of the Ontario
Institute for Studies in Education. The first objective of the
program is to generate the working model for the diffusion of
educational innovations, both process and product, and the second
objective is to implement the emerging model in a school system that
is characterized by the absence of markedly atypical features. The
major components of the model include: a climate for change;
academic-practitioner interaction; roles for evaluation; program
development strategies; interschool cooperation; county wide
communication networks; and teacher responsibility for change. Stages
in the model include: agreement to begin; establishment of an
organization; selection of problems and goals; study of available
solutions; pilot trials; adopt, adapt, reject decisions; and field
trial. The major contribution to new knowledge that this model makes
has to do with the ways in which the components interact to produce
validated educational change. (Author/DB)

AERA Annual Meeting

Chicago 1972

Evaluation

and

The Diffusion of Educational Innovations

The Trent Valley Centre
of the Ontario Institute
for Studies in Education

H. H. Russell
K. A. Leithwood

The role of evaluation in a working model for the diffusion of educational innovations is the central topic of this paper. The working model in question is a product of the Trent Valley Centre (TVC) which is a "regional laboratory" type extension centre of the Ontario Institute for Studies in Education. Since the time of its establishment in September 1969 the TVC focussed most of its resources on a program that aims to identify general procedures for improving student performance on school objectives through the application of educational research manpower and information. The first objective of the program is to generate the working model for the diffusion of educational innovations, both process and product, and the second objective is to implement the emerging model in a school system that is characterized by the absence of markedly atyp.cal features. The third and fourth objectives of the program, concerning field trials of the model, and comparisons of the model with competing models, are not in an operational sta.e and hence remain beyond the scope of the present paper.

The Trent Valley Centre is located in Peterborough, a city of nearly 60,000 population approximately ninety miles northeast of Toronto, Ontario. During the initial year of operation of the TVC there were five schools participating directly in the TVC Program. The number of participating schools expanded to seventeen during the second year, and now in the third year there are approximately thirty schools involved in the TVC Program. A full report of the second year's activities is available in "Educational Change and Innovation: A Case Study" (Russell, Leithwood and Baxter, in press).

Rationale for the Trent Valley Centre Program

The questions why evaluate, why innovate, and why plan the diffusion of innovations in education, have answers which seem self-evident, and yet there are many issues raised by these questions that remain unresolved in most school settings. School innovations are typically introduced to better achieve existing school goals or objectives. Under such circumstances, it is necessary to conduct a careful evaluation to determine the exact extent of the advantage of the innovation over the previous program. It is reasonable also to introduce innovations for the purpose of achieving new and more desirable school goals. Again evaluation is an essential component in the total scheme, because in the absence of evaluation data it is possible that the new goals are achieved only superficially by the innovative program. Another reasonable purpose for innovating is to increase the proportion of students who achieve objectives that are previously achieved by too small a proportion of students. In each of the above cases of gains, there is the possibility of a cost factor which may be prohibitive. The issue of cost raises the possibility of innovations that are designed primarily for the purpose of reducing the dollar or manpower costs. In such cases even a zero gain in student performance may be acceptable. In the event that any of the above types of benefits to society can be made or are made in a school setting, extension of the benefits over a wider range of schools and students is warranted. The TVC study assumes that planned diffusion can be more effective than either unplanned diffusion activities or standard dissemination procedures. This is a testable hypothesis which may be accepted or rejected on the basis of empirical verification at a later date.

There are very basic questions that remain unanswered in the rationale considerations raised above. Who selects educational objectives and who decides what innovations will be adopted or rejected? When officials of a national government select objectives and when accountability at that level is under consideration, then the gross national product is the most convenient indicator of attainment. In the past, the implicit goal of education, so far as national leaders and the GNP are concerned, has been to increase the amount of time spent in school per student, to increase the number of students benefiting from school, and as well to increase the proportion of students who attain graduation status in the higher levels of education. Economists characteristically deal with these variables and they have reasonable evidence to show their relationship to the GNP (Denison, 1964, pg. 23).

The present view of the Trent Valley Centre staff is that GNP is an inadequate measure of improvement within a society (the argument is basically that proposed by Galbraith, 1969, pg. 124). Furthermore, economists in education and educational administrators should accommodate their views to the position taken by Denison (1962, pg.67),

"It (GNP) can deal only with changes in the amount of formal education received by members of the labour force, it cannot take into account changes (presumably improvements) in the quality of the day's schooling."

Denison also asks,

"can these schools teach as much in less time or at less cost through attention to the critical objectives being achieved, than through innovative means to maximize student attainment of the objectives?"

The Trent Valley Centre has been influenced by such considerations and its attention to precise assessment of student performance on educational objectives is an essential ingredient in the TVC Program.

When national leaders have the responsibility for setting educational goals, either explicitly or implicity, it seems that the accountability issue is relatively clear and that attention to specific student performance objectives will probably lead educational researchers and educational administrators in the right direction. However, there are problems with educational goals for a large unit of jurisdiction, whether supernational, national or subnational. Even the act of setting goals over such a large unit of jurisdiction may be dysfunctional. Dewey has commented on the effects of nationally or externally imposed goals in Democracy and Education (Dewey, 1969, pg. 8),

> "The vice of externally imposed ends has deep roots. Teachers receive
> them from superior authorities; these authorities accept them from
> what is current in the community. The teachers impose them on
> children. As a first consequence, the intelligence of the teacher
> is not free; it is confined to receiving the aims laid down from above.
> Too rarely is the individual teacher so free from the dictation of
> authoritative supervisor, textbook on methods, prescribed course
> of study, etc., that he can let his mind come to close quarters
> with the pupil's mind and the subject matter. This distrust of
> the teacher's experience is then reflected in lack of confidence
> in the responses of pupils. The latter receive their aims through
> a double or treble external imposition, and are constantly confused
> by the conflict between the aims which are natural to their own

experience at the time and those in which they are taught to

acquiesce. Until the democratic criterion of the intrinsic

significance of every growing experience is recognized, we shall

be intellectually confused by the demand for adaptation to

external aims."

The Trent Valley Centre staff recognizes the basic significance

of Dewey's comment and accepts some responsibility for influencing school

people to consider aims of education that are rooted in the students and

the people of the school community. By accepting such a view it is not

necessary to exclude aims of education for larger jurisdictional units: in

fact, the need for such aims is obvious. Such aims must be sufficiently

general that individual students and teachers may be free to identify and

pursue specific local and individual goals.

A. N. Whitehead (1956) provides even more precise guidelines for

the identification of the person or persons who should have responsibility

for identifying educational objectives. Whitehead says,

"But the first requisite for educational reform is the school as a

unit with its approved curriculum based on its own needs and evolved

by its own staff. When I say that the school is the educational unit,

I mean exactly what I say, no larger unit no smaller unit."

(Whitehead, 1956, pg. 21).

If Dewey and Whitehead have provided good advice on this issue, then the

"who" question raised earlier must be answered in a way that provides a

high degree of school autonomy and individual freedom.

Independently of the issue of compatibility between school goals

and national goals, there is good reason to study carefully school goals,

and to accurately assess the degree of attainment of these goals and their
specific subgoals or student performance objectives. This Is the business
of evaluation and, furthermore, student performance data can serve as a
reasonable starting point for the consideration of school change and the
introduction of innovations. The identification of innovative programs
that are compatible with the students' needs and the community's ability
to pay the cost is a sophisticated business that sometimes leads to the
adoption of rigid packaged programs and at other times to the preparation
of school-made, school-unique programs. Whatever the innovation, the
adopt-adapt-reject decision must be related to precise evaluation data,
and in cases where positive benefits are the clear result of an innovative
program, the implementation of an effective diffusion program can magnify
and multiply the benefits to a larger population of students.

The Trent Valley Centre working model consists of both components
and stages. The components, seven in number, are features found to be
critical to change which appear in one or more of the stages of change.
The stages are seven temporally sequenced points on a continuum beginning
with a decision to begin working toward change and ending at the stage of
field trial, from which point recycling through some earlier stages is
still likely to occur. The major components of the model include: (a) a
climate for change; (b) academic-practitioner interaction; (c) roles for
evaluation; (d) program development strategies; (e) interschool cooperation;
(f) county wide communication networks; and (g) teacher responsibility for
change. Stages in the model include: (a) agreement to begin; (b) establish-
ment of an organization; (c) selection of problems and goals; (d) study of

available solutions; (e) pilot trials; (f) adopt, adapt, reject decisions; (g) field trial. The model may be thus conceptualized as a components X stages matrix (Figure 1) with a total of 49 potential cells some of which are obviously important, some empty and many the significance of which has yet to be determined. Those components and stages that relate to evaluation are elaborated in some detail whereas the remainder are given only the amount of attention necessary to provide a meaningful overview of the model.

Stages \ Components	1. Climate for Change	2. Academic School Interaction	3. Roles for Evaluation	4. Curriculum Development Strategies	5. Inter-school Cooperation	6. Communication Network	7. Teacher Responsibility
1. Agree to Begin	11	12	13	14	15	16	17
2. Establish Organization	21	22	23	24	25	26	27
3. Select Problems, Goals	31	32	33	34	35	36	37
4. Study Available Solutions	41	42	43	44	45	46	47
5. Pilot Trials	51	52	53	54	55	56	57
6. Adopt Adapt Reject	61	62	63	64	65	66	67
7. Field Trials	71	72	73	74	75	76	77

Figure I - - - A matrix showing possible interactions of components and stages in a school change model.

Climate for Change

The climate for change is an elusive component to analyze, primarily because of the variety of forms it takes in different school contexts and the consequent difficulty in identifying common underlying features. Classification of the type of change referred to, and illustration of several situations where such a climate pervades, probably yields the best possible understanding of this component.

Types of change have been classified by many authors (Maguire, 1970) using a variety of criteria, one of the most powerful being the relationship between a change agent and client system (Bennis, 1966). All of these classification systems, however, include a category labelled "planned change" and the variety of definitions offered for such change suggests that it is congruent with the change being discussed here. Chin (Maguire, 1970) suggests that such change is ". . . a deliberate and collaborative process involving a change agent and client system that are brought together to solve a problem or to plan and attain an improved state of functioning in the client system by utilizing and applying valid knowledge (p.11)." In the present model such a definition needs to be tempered by some of the characteristics of what Bennis describes as "technocratic change." This type of change relies on the client's (teacher's) definition of his problem, the agent's knowledge of strategies leading to solution and the collection and interpretation of data to facilitate and validate that solution. Such planned technocratic change is endorsed as a systematic procedure for efficiently carrying out rationally preconceived alterations in the educational enterprise. Because such alterations in the curriculum area, as well as many other areas, are continuous, the roles of the client and agent, in this model, are not clearly distinct. The client must be given

the opportunity, in his initial encounters with the agent, to acquire enough of the agent's motivations and skills so that agent and client are eventually dual roles played largely by the same personnel with decreasing external inputs. Such welding of client-agent roles into the person of the educational practitioner is designed to overcome serious criticisms levelled by Herzog (Maguire, 1970) at typical approaches to the concept of planned change. These criticisms include (a) viewing schools as objects to be manipulated and (b) failure to recognize that most people are engaged in activities because they see value in those activities not because they are resistant to change. A third criticism - that planned change is too often naively profession-centred is compensated for in the TVC model by systematic involvement of the community in school goal specification.

While the type of change (planned-technocratic) has now been established, it is still not clear what the term "climate for change" really means. Planned-technocratic change implies a process of conscious, systematic and scientific response to rationally identified educational needs or goals. Such a process requires large resource allocations on the part of those involved and, hence, a clear understanding and substantial commitment to the need for change on the part of practitioners engaged in planning the change. The climate for change is the behaviourally represented indicators of such understanding and commitment. Such a climate seems to exist (a) when teachers wish to meet with parent groups to cooperatively engage in school goal setting, (b) when teachers meeting as a group openly discuss their real problems and cooperatively plan toward solutions, (c) when a principal is willing to assume a facilitative rather than directive role with teachers, (d) when a school staff desires to rearrange timetables to

free blocks of cooperative planning time, (e) when a teacher feels secure
enough to oppose, on rational grounds, changes suggested by the principal.
All of these indicators of what is meant by a climate for change are found
In the schools in which the change model is developing and many more could
be cited. Suffice it to state that the climate for change is a pre-condition
to planned change or even planning for change, and without it subsequent
actions would probably be largely ineffective. It is, in fact, the failure
to establish such a general climate for change that has doomed many curriculum
development projects to failure before they have begun.

The ways in which a climate for change is established vary greatly
across schools, but in most instances the principal is responsible for its
Initiation. Several appendums to this statement should be noted immediately.
First, while the principal is a centre of communication (see page 29), the
teacher has ultimate responsibility for effecting classroom change (see page 29)
and hence the initiation of change by the principal must be as a stimulant to
the teacher's assumption of responsibility for change. Such initiation cannot
be forced nor should it reduce the teacher's important decision-making
responsibilities, but only make clearer how the principal can be used to
facilitate the changes envisioned by the teacher. Most important, the
teacher must be made aware that his actions are endorsed and supported by
the principal. Second, it would be misleading to suggest that the methods
used by principals in the project so far, to initiate change, have much or
any generalizability beyond the schools in which they were used. Principals
who encountered difficulty in stimulating their staffs to assume responsibility
for change when the TVC Program began are, almost to a man, still having a
great deal of trouble, in spite of the benefit of consultation with fellow
project principals having more success in this regard. The establishment of

a climate for change is a highly complex endeavour involving an optimum
mixture of principal support, staff readiness, group dynamics, problem
visibility, community characteristics and a host of other, less readily
identified variables. When the right mixture exists, a possibly false
issue such as low school scores on a standardized test in mathematics
can be sufficient to elicit a commitment to change which will carry those
involved through the labour of all seven of the model's stages of change.
A more authoritative evaluation of school programs using more valid data
is clearly a preferred means of beginning.

Academic-Practitioner Interaction

The academic-practitioner interaction issue has remained virtually
unresolved over most of the history of education. As a consequence, the
number of solutions generated by educational researchers to problems of
concern to practitioners is woefully small. Several reasons for this state
are readily evident. Educational researchers are too often engaged in
finding solutions to problems identified by themselves and irrelevent to
school people. No matter how significant the solutions thus generated,
their impact will be minimized by the lack of need for such solution. Re-
search activities, to be productive in school contexts, must be organized
around issues considered critical by practitioners. This means (a) creating
two-way communication links between the researcher and practitioner and
(b) utilization of expert manpower to solve the emerging issues even if
that means engaging researchers in projects which are not of natural concern
to them. The idea of the researcher, using his scientific skills in pursuit
of knowledge vital to the interest of others and not necessarily himself, is
largely foreign and probably distasteful to many educational researchers.
What is being suggested is the adoption of a model for client-agent interaction

which more nearly approximates the industrial rather than the university
model with its record of brilliant yet often disjointed and ineffectual
achievement. Perhaps the best of both worlds is possible with the proper
matching of a large pool of expert manpower resources with problems judged
relevant by external, school sources.

Even when academics have dealt with school relevant problems,
the strategies for operationalizing change, as often as not, have been
impotent. Responsibility for this condition rests evenly with academics
and practitioners. Academics have typically adopted dysfunctional tactics
characterized by a large proportion of discussion of problems and solutions
at a general level, and a very small proportion of expended effort in the
task of change implementation. It is erroneous to suggest that the
generation of solutions and their implementation are separate activities.
In fact, only through the labour of attempted implementation does a
solution acquire that ultimate richness which qualifies it as a worthy
achievement. Part of this unfortunate situation must be attributed to the
implied wishes of practitioners, however. An invitation issued to an academic
"expert" to speak at a professional development day on some related occasion
is one of the best possible ways of confusing an issue and building in a
resistance to change on the part of a majority of those who attend such an
occasion. Change will only occur when an organization is established to
facilitate communication on a continuing basis. If inspirational speeches
were in fact effective, every school would be a model of planned change
since all staffs have been subjected to such inspiration many times. The
lecture strategy, therefore, may serve to give the appearance of change
activity and avoid the potential trauma associated with the realities of
change.

Roles For Evaluation

Scriven (1967) has distinguished between goals and roles for
evaluation. Although there is some disagreement, (Stake, 1967), the goal
for evaluation in this model is singular, as Scriven suggests, that being
to judge the merit or worth of an educational variable. Ths most con-
troversial issue here, in fact, is not whether there are other goals for
evaluation but who will be the judge. Consistent with the concept of
planned-technocratic change, discussed above, in which there is gradual
welding of client-agent roles, and contrary to the expressed preference
of Scriven (1967), the judge in this model is the client-practitioner.
The initial change agent assumes consultative responsibilities with regard
to data analysis, interpretation and research design, where they are
important, but the adopt-adapt-reject decisions are exclusively in the
domain of the practitioner who must ultimately implement the decision and
be held accountable for its consequences.

The roles of evaluation are many including prediction, selection,
national assessment, diagnosis, curriculum assessment and many others.
These various roles have been classified as either formative or summative
depending on the uses made of the resulting data (Scriven, 1967). Summative
evaluation provides information to enable a potential adopter of an educational
process or product to determine the relative costs, advantages and dis-
advantages of an innovation prior to making a commitment. When packaged
programs are selected by schools, this is the use to which evaluative data
are put. Formative evaluation, on the other hand, proposes to supply
information to the developer to assist in the further refinement of process
or product. As a result, formative data may often have to be very detailed

and particularly multi-dimensional in order to be useful. Furthermore, they must be available when they will be useful to the developer. Where the teacher is also curriculum developer, data from the previous year's program must be available before beginning the new year's program, and data from each segment of the program must be available as the subsequent application of the program segment is planned.

Within this formative role, although not exclusively, evaluation serves three distinct functions in the change process. These three functions are the promotion, facilitation and validation of change. Evaluation data often function to promote change by stimulating attention to problems with the status quo and providing a basis for making decisions about program adequacy with regard to current educational objectives. A decision to evaluate present curricula also leads through a process of goal clarification and refinement which virtually guarantees that the program finally evaluated will not be the original program. This is one of the effects of experimenter intervention that is to be avoided at all costs in controlled research, but on the other hand it is to be greatly encouraged in the process of planned educational change. Evaluation has thus begun to facilitate change and can do so further by enabling developers to identify program objectives being achieved least well, as a focus for the initiation of systematic curriculum development. Such identification of initial focus has the advantages of:

(a) diagnosing weaknesses teachers may be able to improve on immediately in the classroom;

(b) scaling down the size of the curriculum development task by avoiding work on objectives already being well achieved.

Both of these features are especially attractive when teachers are also

developers since they are likely to feel the press of daily classroom needs, and have only enough time to work on the most urgent curriculum problems.

The third function of evaluation in change, that of validation, speaks directly to much of the available change literature which appears to imply that change is to be valued in its own right without regard to the consequences of that change. The TVC model diverges most severely with the literature in taking considerable pains to assess the effectiveness of change and in providing opportunities to adapt or reject ineffective changes. Many dimensions of innovative products, chosen for or developed in the project, have been measured to date. They include student, teacher, teacher aide and parent attitudes, description of treatments, student performance and the relationship of student performance with baseline descriptions including IQ, socioeconomic status and standardized achievement scores (see Chapter 6, Educational Change and Innovation: A Case Study). Particular attention has been paid to student performance data on the assumption that an innovation to be worthwhile must result in student achievement that is higher, different and more valuable, or less costly than in the previous program. (An unlikely but possible exception to this would be an innovation which had a positive effect on attitudes and produced the same amount and type of student performance as the previous program at the same cost.)

The assessment of student performance has moved the focus of measurement within the model away from classical norm-referenced measurement techniques into criterion-referenced measurement. This shift has taken place not because of a profound disagreement with classical measurement methods but because of a difference of purpose. Classical measurement methods are designed largely for selection and prediction. Because of this, test development procedures tend toward making such instruments unidimensional.

Assessments, using such techniques, indicate where a testee stands in relation to other testees with respect to some identified measure or some clearly prescribed summation of measures. However, the threefold purposes of evaluation in the TVC model require information of a very different sort.

Within the TVC model the student variables being measured may be heterogeneous rather than homogeneous, and each component part may require separate assessment. This appears to be the case in practice where the dual purposes of formative evaluation as used in the model include both diagnosis and curriculum evaluation. In each case, student achievement of specific educational objectives is the information being sought and selection and prediction decisions do not enter. In spite of the basically different functions, the same pool of test items may serve both classical and criterion-referenced measurement purposes. The differences in purpose will dictate differences in procedures for item selection from the common pool.

The shift in emphasis from norm-referenced to criterion-referenced data has some additional positive features that were not taken into account during the initial discussions, that have led to the shift. A renewed emphasis on the individual rather than society, and a new fascination with the negative effects of competition and aggression seem evident not only in current literature but also in official government policy statements such as the Commission on Post Secondary Education (1972). Criterion-referenced data do not eliminate competition, nor do they inevitably lead to a focus on the individual rather than his group. However, such data do not depend on, nor do they emphasize quantitative discriminations among students. Also

they more readily lend themselves to heterogeneity In goals among Individuals and to the use of data by the Individual who Is measured.

Process evaluation Is considered to be as Important as product evaluation but It Is more difficult to develop an appropriate methodology for It. Several promising plans have recently been Initiated In cooperation with OISE sociologists Michael Fullan and Glenn Eastabrook. These plans Involve Initial data collection of school characteristics and modes of operation related to Innovativeness. Subsequent Intervention programs will be undertaken, when project schools desire It, to alter, where possible, characteristics which detract from optimum change activity. To date, the collective processes Imbedded In the change model appear to be reasonably effective, but the process evaluation of the next two years will be necessary to Identify specific strengths and weaknesses of these processes.

Student achievement In the project curricula have been, or will be In the near future, assessed employing one or more of three experimental designs. Using Campbell and Stanley's (1963) terminology these designs Include (A) the "one-group pretest-posttest" pre-experimental design, (B) the "nonequivalent control group" quasi-experimental design and (C) an extremely useful adaptation of the pre-experimental "static group comparison."

This year the "one group pretest-posttest" pre-experimental design was used to evaluate student achievement In all of the project components. This design yields Information limited by a number of extraneous variables that can jeopardize Internal validity. Those variables Include history,

maturation and the effects of testing. The effects of history and maturation
will be greatest in the case of two packaged programs since a greater time
elapsed between pre and posttesting. Results from teacher-built programs,
because of shorter elapsed time are less influenced by history and
maturation, but more affected by the tendency of students to do better
on the second administration of a test or alternate form of that test. As
a result of the possible inflationary effect of these threats there is
interpretation of results. Nevertheless, as already discussed, because
two teacher-built curricula are in the formative stages of development,
the inadequacies of this design are to be preferred to a more rigorous
design which might have discouraged further program development - the
promotion of which is a major role for evaluation in this context. While
inadequate resources are a major reason for employing this same design
with two highly developed packaged curricula, the benefits of the design
remain compelling. These have to do with the freedom teachers are given,
in a non-comparative setting to develop not the already fixed program
components, but the necessarily school specific techniques required to
make the program optimully successful. Further, the pressure of
implementing a new curriculum in the first year could, under some cir-
cumstances, be added to by the knowledge of comparison.

The quasi-experimental "nonequivalent control group" design was
used with pre and posttesting in one instance this year and its use will be
expanded next year to include four programs thus far evaluated. Socio-
economic status (as well as school size) was used as a basis for control
because of both its potency in relation to achievement variables (Ireton,
1970) and the availability of such data without large resource expenditures.

Depending, to some extent, upon the equivalence of groups on pretest

scores this design controls the main threats to internal validity of

history, maturation, testing and instrumentation since these factors

should influence both control and experimental groups similarly. Re-

gression effects are not likely to occur since there is no need to be

concerned with extreme groups.

An adaptation of the "static group comparison" design has been

made which makes the design inexpensive, convenient and practical to employ,

as well as yielding unusually dependable comparative data for evaluation

purposes. Typically, this design compares the posttest results of a group

which has experienced a treatment with one which has not for the purpose

of establishing the effect of the treatment. The comparison group consists

of the pupils in the same grade, same school with the same teacher the prior

year. For example, an innovative kindergarten program being introduced in

September 1971 is to be evaluated with this design, so in May 1971 the grade

1 children (last year's kindergarten class) were posttested and these re-

sults compared with the May 1972 results of the class taking the innovative

program. The assumption of pretest group equivalence cannot be accepted

as readily as would be the case had comparison and control subjects been

assigned to groups randomly. Yet history is virtually the only confounding

variable to a highly effective group match given the relatively stable

population from which the subjects are drawn. In many cases the children

even have the same parents. Certainly, group measures of SES, IQ, and

general information as well as measures of school environment and community

environment are almost the same for both groups.

This design has a number of important advantages for development

work, the major one being Its economy. Comparison groups are always difficult to find and non-innovating schools are understandably reluctant to participate in studies for which they perceive little payoff to themselves. Using this design, both control and treatment groups can be drawn from the innovating school. In order to provide both groups using randomization techniques, half of the target population must forgo the innovative treatment in order to serve as comparisons, or the treatment can be offered to half of the target population for the first half of the year and the remainder of the target population in the second half of the year. This latter arrangement necessitates innovating during only half of the school year. None of these requisites for randomization are convenient for school people and none are necessary with the design modification suggested. The main need for caution with the proposed design concerns the degree of instability of the school population and environment. Rapid changes can invalidate the design and thus it becomes infeasible in spite of its apparent strengths.

Several other issues related to the operationalizing of experimental evaluation design are appropriately dealt with here. The first concerns procedures for data collection, specifically the collection of criterion-referenced data. As discussed earlier, collection of such data ideally include comprehensive assessment of student achievement of all program objectives, resulting in many test items. In most school settings the amount of time required to test all students on all items becomes prohibitive. Three solutions to this problem are most readily apparent and two have been tried in the TVC Program, to date. One is to incorporate the testing into the ongoing program using the items for teacher-diagnostic

as well as overall evaluative purposes. A second solution involves random assessment of item samples to participating students. The third solution depends on an emerging analytical technique that requires a minimum of items to focus on a student's particular domain of learning. The first solution has the advantage of providing immediate feedback to the teacher and is difficult to surpass as a formative evaluation strategy when the teacher is also the curriculum developer. The disadvantages, however, include inconsistent testing procedures since there are many testers, as well as placing the responsibility for carrying out the testing schedule largely on the shoulders of teachers. When the data are to be used for summative purposes, experience indicates that in spite of relatively open communication channels, the necessity of carrying out the testing schedule and procedures exactly as planned may not be fully understood. This severely jeopardizes the reliability of data.

A more satisfactory compromise now seems to be the provision of diagnostic test items for use in the formative mode as well as a formal testing program carried out by R and D personnel. In order to undertake the latter, techniques have to be found to both dramatically reduce testing time while at the same time assessing most measurable program objectives. One method of doing this, that has been satisfactory in the TVC experience, is to randomly divide the total item pool into a number of sub-tests, each requiring about the same period of time to complete, and then randomly assigning sub-tests to students within a given class. This strategy results in a mean score for each item, for each class, rather than a score for each student. Such a result is particularly attuned to the purposes of criterion-referenced measurement since the adequacy of the program in achieving each

of Its objectives becomes the focus of evaluation rather than the adequacy of the student.

The third solution is under study by the TVC during Its development by OISE staff (see Olivier 1971). Through the careful preparation of dependency trees, Olivier has been able to select test items that in themselves provide information concerning the many dependent items from lower positions on the dependency tree. At the present time the Olivier system is operational in a computerized remedial mathematics program, but a hard copy product of the program can be used effectively without access to a computer facility.

Frequently criterion-referenced measures consist of large numbers of performance items which place heavy demands on the time of the evaluator if formal testing procedures are to be carried out. In the project the time of testers has been severely limited resulting in a search for adequate methods of expanding evaluation without a reduction in the quality of the data. In one program now under evaluation (to be reported next year) a highly selected group of volunteer parents were trained in the administration of individual performance test items in science. From the R and D evaluation point of view, the data they collected were highly reliable due to the extensive training and monitoring of testing procedures. In fact, because parents were chosen who had children the same age as the subjects being tested, the rapport of tester and testee often surpassed the rapport a professional tester would be capable of establishing. From the point of view of the principal implementing the program that was being evaluated, the use of parents as testers had several added positive features. The parents, having become thoroughly familiar with the school operation, in general, and the new program, in particular, acted as goodwill ambassadors to the community for the school. Exposure to the school under structured conditions

also added considerable impetus to this volunteer parent program.

Curriculum Development Strategies

In spite of the importance of processes for change, the criterion against which they must be judged is the outcome or product of those processes. It would be difficult, however, to defend an absolute distinction between process and product since a product like "student achievement," as it can be measured, is only a static and, therefore, artificial record of continuous learning and performance. An operational distinction can be made where classroom treatments such as IPI or Mathematics A are defined as products, activities preparatory to such treatments such as teacher planning and in-service training are processes, and student achievement is the outcome criterion against which both product and process are judged. When the product is defined as classroom treatment, the limitations of both product and process evaluation become immediately evident. Such evaluation, concerned with student achievement, is an assessment of unique, non-repeatable treatments when those treatments are each considered as a unit. There are critical components of those units that are repeatable, however, and the curriculum development processes or strategies in this model focus on these repeatable components.

The strategies employed in program development in the model include four dimensions (Figure II), these being the identification of educational objectives, the gathering or generation of materials to be used in achieving those objectives, the choice of instructional techniques for manipulating those materials in an effective manner, and evaluation of the achievement of objectives, materials and techniques.

The sorts of formative and summative evaluation required and the

- 25 -

selection strategies exercised on materials and techniques necessitates the

```
        ┌────────────────────────────────────────────────────────────────┐
        │   ┌──────────────────────────────────────────────────┐          │
        │   │         ┌──────────────────────┐                  │          │
        ▼   ▼         ▼                      ▼                   ▼          ▼
┌──────────────┐ ┌──────────────┐ ┌──────────────────┐ ┌──────────────┐
│  1.          │ │  2.          │ │  3.              │ │  4.          │
│Specification of│Accumulation of│ │Choosing Instruc-│ │             │
│Objectives    │ │Materials     │ │tional Techniques│ │Evaluation    │
└──────────────┘ └──────────────┘ └──────────────────┘ └──────────────┘
        ▲                ▲                  ▲
        └────────────────┴──────────────────┘
```
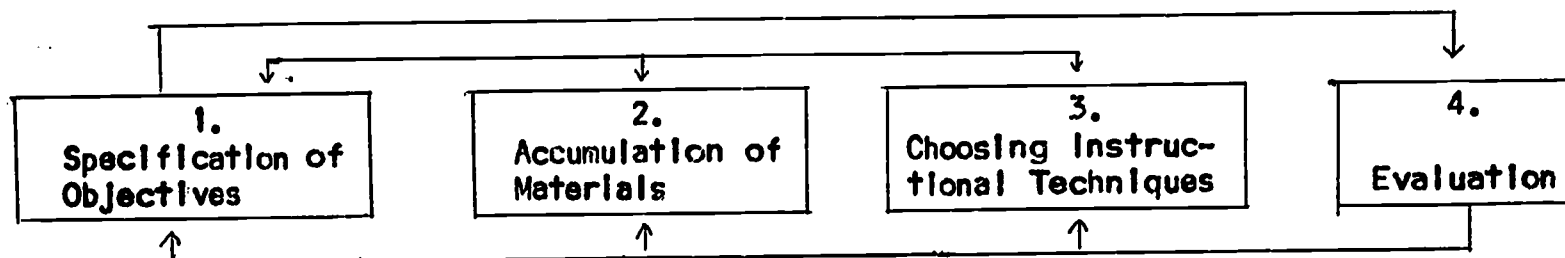
Figure II - - - Components within curriculum development strategy.

specification of objectives to be in operational, behavioural or student performance terms. While many arguments have been raised against such specification, most are patently invalid (Leithwood, Russell, in press; Popham, 1970). The position being forwarded really is no more complex than indicating ones purposes so that one can determine how to go about achieving those purposes and assess when or if those purposes are achieved.

The order in which the components appear in the curriculum development process vary greatly with the needs and characteristics of the developers. In many instances the shortest route to total curriculum development task "completion" (the task is really never done) may begin with the accumulation of materials component. Beginning with selection of materials is useful since it leads to innovations in the classroom very quickly and generates the need to move concretely to issues of objectives as a basis for decision-making about selection of materials. Beginning with evaluation has the advantage of diagnosing areas of greatest present weakness and treating those areas of greatest need first. It necessarily involves beginning where one is at present and moving from there - - a very sound policy. Beginning with the objectives component is logically appropriate but objectives specification can be a long and difficult task requiring considerable patience before the impact of such work becomes visible in the classroom.

One of the most promising ways of reducing the size of objectives specification task Is through the use of objectives pools like the Instructional Objectives Exchange (Popham, 1971) or, even better, objectives and Items pools (Horn and Russell, 1971; Lelthwood and Russell, 1971). Such pools enable teachers to select student performance objectives, which they feel are appropriate for their curricula, without the need of generating such objectives on their own. When the objectives have accompanying Items, two of the four components of curriculum development, objectives specification and evaluation, are greatly facilitated. The curriculum developer's major concerns can then be focussed on objectives selection and the materials and techniques components of curriculum development. It might also be possible for materials'developers to relate closely to such pools and help create teacher resource centres where, not only objectives and test Items, but associated materials were catalogued (Lelthwood, 1971). This Is not to suggest that the same materials cannot be used to achieve many objectives but some are obviously Inappropriate and a few extremely useful for Instructional purposes related to a specific educational objective.

The arrows In Figure II Indicate that curriculum development can begin with any component and move In a number of directions. It Is not possible, however, to arrive at the evaluation component without going through the specification of objectives first. Having arrived at evaluation, however, the resultant data potentially feedback Into all four components for further revision and refinement. Suppose, for example, that the results of evaluation Indicate no student performance gain on a given objective

after exposure to the curriculum treatment. Fifteen possibilities potentially are available to account for such lack of program ineffectiveness. These refer to decisions about each of the components in Figure II considered separately as well as all possible combinations of such components including:

1. The objective was inappropriate, unachievable or otherwise poorly selected for the program and requires revision or elimination;

2. The program materials designed to achieve that objective are ineffective and require revision or change;

3. Instructional techniques need review;

4. Items used to measure objectives achievement are invalid;

5. to 10. Problems with double combinations of components including 1 and 2, 1 and 3, 1 and 4, 2 and 4, 3 and 4;

11. to 14. Problems with triple combinations of components including 1, 2 and 3, 1, 2 and 4, 1, 3 and 4, 2, 3 and 4;

15. Problems with all 4 components of the curriculum development process.

Because each of the four components requires its own set of skills to carry out adequate development, and because of the large number of possible decision points (15) requiring data of some sort, two observations seem notable here. First, the complexity of the curriculum development process, considered without reference to a larger framework of school change, has been grossly underestimated by school people generally, and perhaps by academics as well. Second, the availability of multi-dimensional evaluation data is vital. Both of these observations seem, on the surface, to suggest that asking the busy teacher to be a part of curriculum development is very unrealistic. On the contrary, expecting anyone but the teacher-developer

to have access to the range of primary, secondary, objective and subjective, sometimes impressionistic, informally gathered data necessary to cope with the 15 decision points outlined is even more unrealistic. The task is simply beyond the scope, resources and technology of the professional curriculum developer and evaluator. What seems to be most realistic is substantial support for the teacher in his role both through extensive in-service training in a meaningful context, and facilitative and consultative personnel and agencies prepared to act on needs identified by the teacher.

Teacher Responsibility

The rationale for primary teacher involvement in curriculum development and school change, generally, has been foreshadowed in the previous discussion of curriculum development and welding the roles of the change agent and client into the person of the client. The need for such responsibility and identification of the school as the critical educational unit have been attributed in part to Dewey and Whitehead in the opening paragraphs of this paper. Both philosophers knew that if educational objectives are to be implemented in classrooms they must be the schools own objectives. When this is not the case and objectives are imposed on the teacher, a conflict typically arises between dormant explicit objectives and very active implicit objectives with implicit objectives most frequently winning. This consequent suppression of operational classroom objectives can only be detrimental to systematic curriculum development, communication and student achievement.

Teacher responsibility for curriculum, however, does not imply that superintendents, subject-matter specialists, consultants and trained curriculum developers cannot lighten the teachers' load greatly without violating the principle of teacher responsibility. What is suggested

4

edd

11

- 29 -

Is that teachers should be encouraged to identify the needs they see in their classrooms, specify (at some level) the objectives they have for their students, feed other persons when necessary with information to guide the development of materials and strategies that will help meet those identified needs and critically evaluate materials and strategies in light of their objectives. It should be the teachers' prerogative to make the final adopt-adapt-reject decision (within the financial limitations imposed on administration) because it is the teacher who is ultimately held accountable for the performance of the students subjected to the curriculum.

Communication networks and inter-school cooperation

Two networks for communication have been elaborated in the model. The first is a communication network which relates educational personnel by constituent position to one another where the school principal is the hub of communication with respect to change and innovation. The central role of the principal in this network highlights his function as change agent having direct communication access to senior administration, teachers, students,
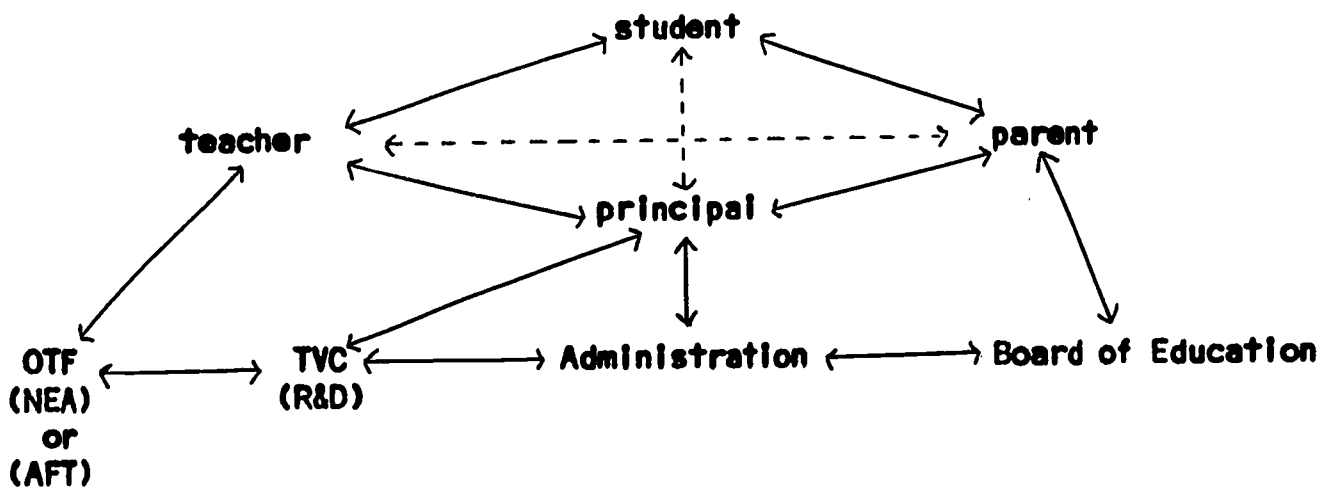

Figure III - - - The Central Position Of The Principal In Educational Communication

30

parents and outside agencies, although no direct contact with elected trustees in most cases. He is, therefore, in a position to initiate and facilitate change in his school and the school, as already suggested, is the primary educational unit.

The second communication network functions as a mechanism through which the principal and sometimes staff representatives performs many of their facilitating activities. This network links school principals formally involved in the innovative process through cooperative groupings of various sizes depending upon purpose. Those types of groupings are imbedded in this network. The first of these is labelled a "joint" group and consists of all innovative school principals in a county who wish to be a part of the change model. This group's function is to provide (a) a forum for general issues of common concern and (b) simple information relay.

```
                          ┌──────────────┐
                          │ JOINT GROUP  │
                          └──────────────┘

  ┌────────────┐
  │ Liaison    │              ┌────┐
  │            │              │ 1. │←
  │ Committee  │──────→       └────┘
  └────────────┘                            Special
                              ┌────┐
                              │ 2. │←        Interest
                              └────┘
                                              Groups
                              ┌────┐
                              │ 3. │←
                              └────┘

                              ┌────┐
                              │ N  │←
                              └────┘
```
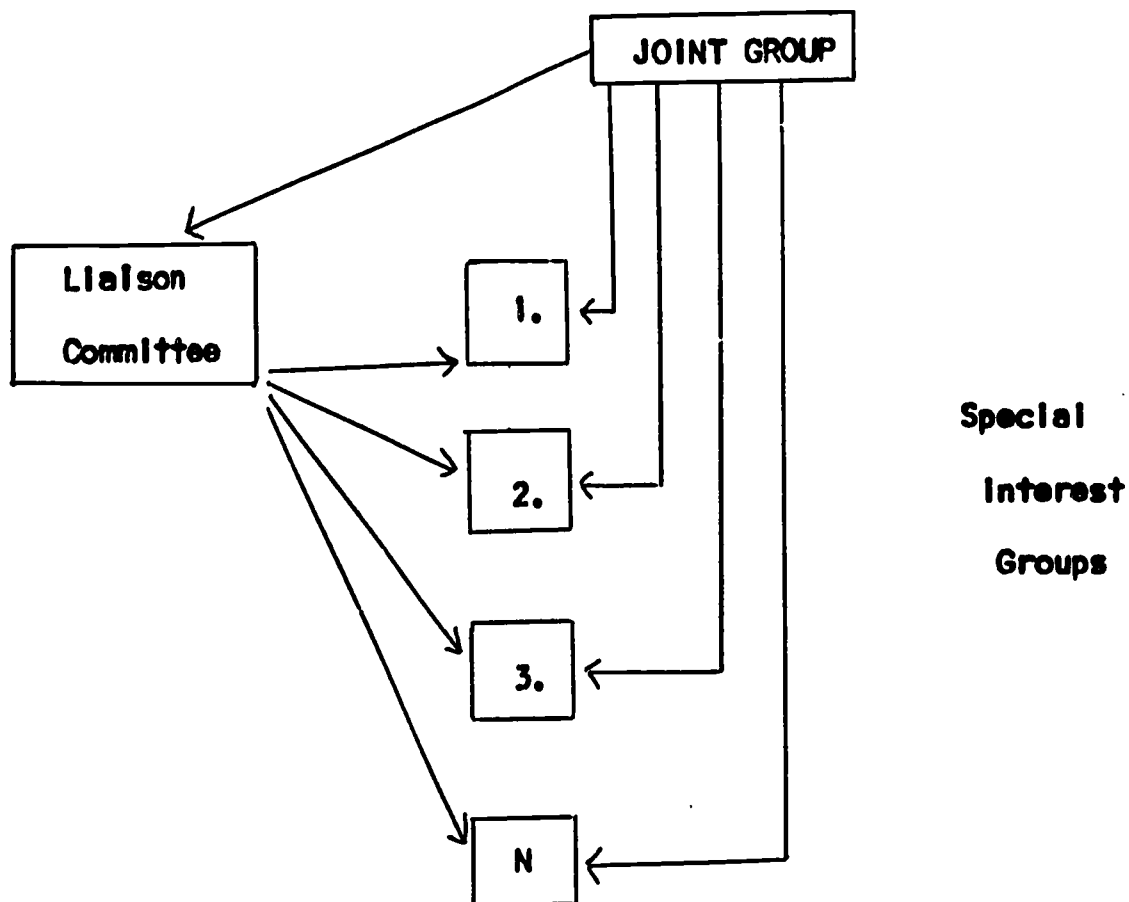
Figure IV - - - A School Interaction Network To Facilitate Cooperative Change.

A second group, the "liaison" group coordinates and facilitates the project work of the participating schools. More sub-groups are designed to deal with specific, substantive curriculum and other change issues. The group structure provides (a) the benefit of mutual experience (b) the dissemination of useful information (c) mutual support and (d) an efficient means whereby consultative assistance (e.g. Trent Valley Centre) can be mustered to deal with important problems of common concern. Furthermore, new schools wishing to join the project may do so at a relatively concrete level through a special interest group, although joining the project does carry with it the responsibility of studying change processes more generally. These new schools may be considered a special interest group focussed on change and representatives from the liaison committee play a consultative role in that group. The groups last only as long as it takes to solve the specific problem they were created to work on.

Some of the advantages of this organizational network are better appreciated when viewed in a broader perspective. An interesting paper by E. G. Bogue (1971) entitled "Disposable Organizations" provides such a perspective. Bogue contends that traditional organizational structures are often not capable of providing the fast acting response needed to deal effectively with contemporary change. Bureaucracies, with hierarchical systems, favour the status quo and contribute to inertia by reducing opportunity for change. One solution to this problem is organizational decentralization but county reorganization in Ontario education was a move in quite the opposite direction. The special interest groups in the network outlined above are what Toffler (Bogue, 1971) calls "throw away" or "disposable" organizations. Such organizations are (a) problem or issue

centred rather than function centred; (b) temporary with a built in self-
destruct mechanism activated upon problem resolution; (c) staffed so that
authority of competence replaces the authority of position and role. This
provides opportunity to utilize diverse specialists in a common venture;
(d) able to short-circuit channels of communication rather than follow
vertically structured paths.

Bogue cautions, however, that "most persons need a degree of
stability along with the challenge of change." Certainly the network being
discussed has a relatively stable, although flexible Joint Group and the
entire network operates cooperatively, rather than competitively within the
more familiar educational administrative structure. There seems to be
good reason, theoretically, therefore, for such a structure to be effective
in facilitating change.

Trent Valley Centre Model Stages

While the components of this model, as they have been discussed,
are important to consider in isolation, many (although certainly not all)
of their features are already well known and, indeed, have been part of
educational knowledge for many years. The major contribution to new
knowledge that this model makes has to do with the ways in which the
components interact in a dynamic way to produce validated educational
change. Educational literature abounds with treatise on objectives,
evaluation and communication. The school as the educational unit and
teacher responsibility for goal setting have been recognized many times
and by many leaders, and yet few attempts have been made to integrate
these ideas in a practical setting so that each component fits within a

larger framework or strategy that will eventuate in the kind of change that is really needed.

Figure V illustrates how the components of this model interact, as conceptualized to date, through seven temporally sequenced stages from initial agreement to change to the stage of field trials of new programs. The latter two stages are as yet not clearly defined and they represent a focus of research for the next two years. It should be noted also that some of the elements and orders within each of the other stages will change as work progresses to refine the model. The work began, in fact, as a "model of" change and now most refinement is in the direction of making it a "model for" change. In the remainder of this paper the stage-based component interactions will be discussed.
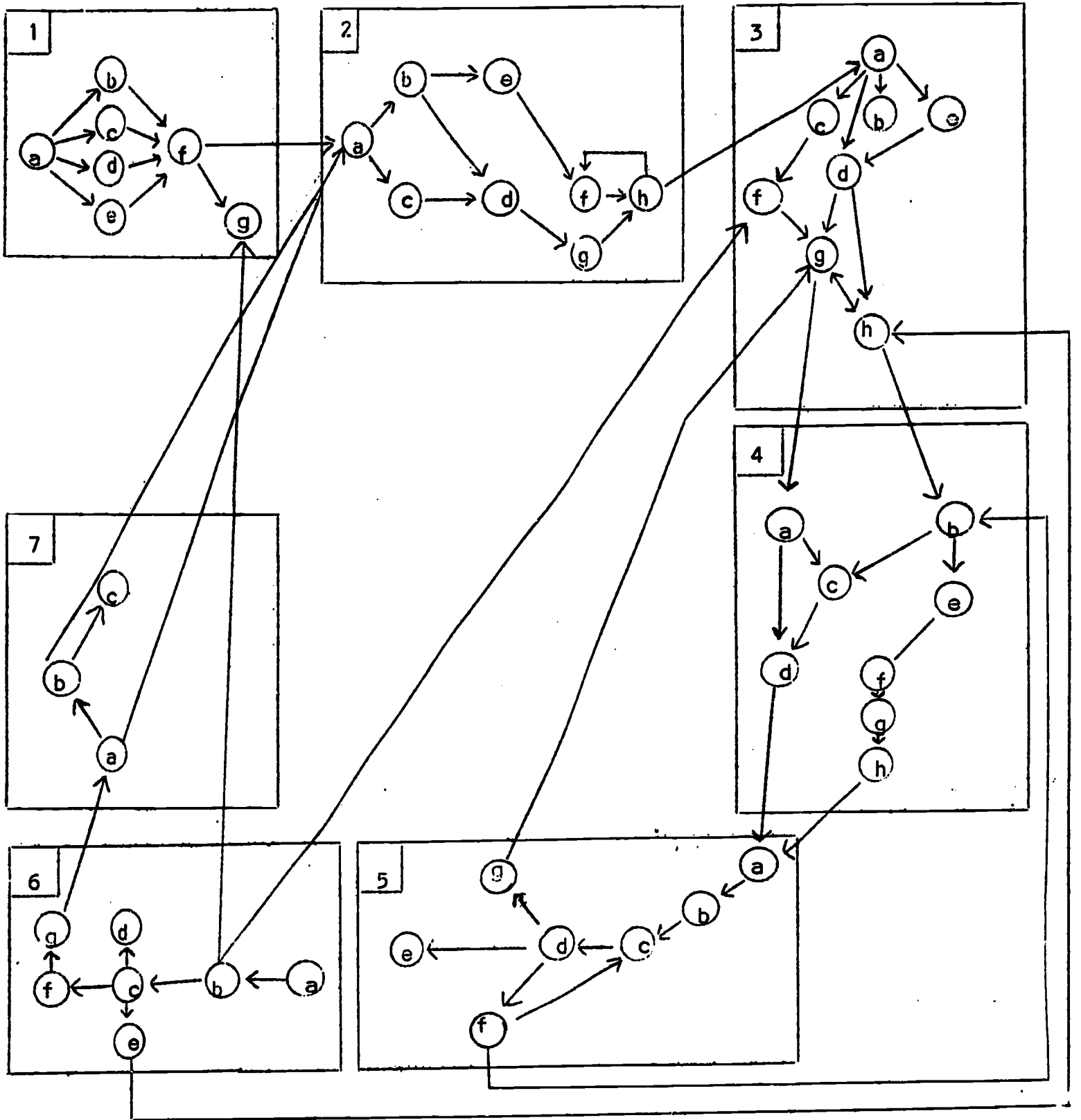
Figure Ⅴ- - - Stages leading to school change

Key to Figure Ⅴ

1. Agreement to begin

    (a) Meeting of representatives
    (b) Meeting of pr & T representatives
    (c) Meeting of R & D
    (d) Meeting of consultants
    (e) Meeting of other manpower groups
    (f) Specific agreement to proceed
    (g) Plan advisory committee

2. Establish organization

    (a) Select schools
    (b) School level meetings, T, pr, +
    (c) Meeting of pr, adm, R & D, consult.
    (d) Pr study change process
    (e) Select T within schools
    (f) T study change process within school during released time
    (g) Pr plans use of expert & volunteer manpower
    (h) T make go or no go decision

* Pt - parents
* Pr - principal
* R & D - Research & Development
* T - teacher
* St - student
* Adm - administration

- 36 -

3. Select problems and goals

(a) T study general school goals
(b) Pr acts as facilitator
(c) Pr study behavioural goal specification
(d) T study behavioural goal specification
(e) School community interacts on gen goals
(f) Pr study available solutions (consultants)
(g) T study available solutions (consultants)
(h) T generate or select instr. goals

4. Study available solutions

(a) T select inn. prog.
(b) T decide to dev. inn. prog.
(c) Pr seek authority to proceed
(d) T training + study
(e) T study materials available
(f) T select materials & organize
(g) T develop or select instructional objective
(h) T prepare auxiliary material

5. Pilot trial

(a) Design pilot trial (time sequence)
(b) Pre-test
(c) Begin trials
(d) Formative eval.
(e) Adopt decision ⎤
(f) Adapt   ⎬ Module
(g) Reject   ⎦

37

6. Adopt, adapt or reject

(a) Recycle prog. within school
(b) Report eval. data for prog.
(c) Adopt
(d) Reject
(e) Adapt
(f) Report summative eval. data
(g) Plan field trial

7. Field trial

(a) Design for eval.
(b) Design for diffusion
(c) Plan eval. of diffusion model

Stage 1: Agreements to Begin

This stage begins with a decision at the county level on the part
of some group of practitioners (a) that their schools could be better and
that they are prepared to investigate how such improvement could be
pursued. Having made this decision, the agreement to begin in a formal
manner (f) necessitates gaining the cooperation of all effected groups
Including principals, teachers, administrators, consultants and R and D
persons who may be able to facilitate the desired change. This gaining
of cooperation is the purpose of meetings (b), (c), (d) and (e). The
result of these meetings and formal agreement to begin is the planning
of an advisory committee (g) which has a broad educational community base
and is designed as a sounding board on which the innovators can test
their ends and means before and during operationalization of these means
and ends. Basic to decisions made at this stage are the concepts of
teacher responsibility and the school as the critical educational unit.
Five principals working toward their M.Ed. degrees were the original
initiators of change in the project from which the model is derived.

Stage 2: Establish Organization

The establishment of an organization for change begins with
(a) a selection of schools. Two of the most important issues here are
who does the selection and what are the selection criteria. Usually,
the preferred situation involves school self-selection on the basis of
evinced interest in change. When this is the case the process moves
directly to school level meetings (b) with and among teachers, principals

and whoever else the school unit feels would be helpful - possibly R and D
personnel and/or consultants. From this point two routes are possible
leading to the teachers' study of the change process in released time (f).
One of these routes involves the principal studying the change process first
(d) as a means of determining techniques for initiating an interest in his
teachers. This route would be followed when there was no strong initial
pressure from the teaching body to change. In such an instance part of
this study might lead him to plan for the use of expert and volunteer
manpower (g) as a way of stimulating interest. An alternate route, approp-
riate when the principal is ready to change and knows teachers are also
ready is to select the teachers who will begin (e) and initiate their
study of the change process perhaps in cooperation with him. Teacher
selection again is a critical issue and self-selection is vital where
possible. The task of change is a massive one, however, and for an entire,
larger staff to be involved simultaneously at the outset would present
problems which might be insurmountable.

Where self-selection does not occur the route from (a) to (d)
involved (a) meetings of principals, administrators and R and D consultants
(c) in order to facilitate selection. The gathering of sociometric data
has been contemplated (but thus far not employed) for assisting in such
selection. Two points need to be made clear here. First, the reaching
of the "go or no-go" decision point with regard to change after a study
of process is most fundamentally the reaching of a point where the teachers
decision takes priority. Second, there are two possible routes if the
"no-go" decision is made. One of these routes is back to further study of
the change process. The second route is to drop out of involvement in
change at all, a difficult or impossible thing to do.

Stage 3: Problem and Goal Selection

The decision to proceed with change leads to a study of general school goals on the part of the teacher at a high level of generality initially but at increasingly specific levels as study proceeds. The end result of this stage is either the generation and/or selection of specific program goals by the teachers or the selection of programs that speak to general teacher goals. In order to reach this stage, teachers must, and principals probably should, be involved in studying the technology of goal statements in student performance terms and studying available programs that potentially achieve these teacher-generated objectives. The school community might also be involved, as it has been in the project, at a general level, in order to ensure that the broad school goals reflected the goals of its most relevant society. Depending on whether or not a ready-made program can be found or a new teacher-generated program is to be developed, two routes into stage 4 seem possible.

Stage 4: Study Available Solutions

The route from stage 3 (g) is to stage 4 (a), the selection of the innovative program and this route can be a reasonably swift one involving principal search for authority to proceed and then teacher training and study of the program leading directly to a pilot trial at stage 5. If the innovators enter stage 4 at point (b) they must examine available materials, select appropriate parts of them, organize these parts and write auxiliary materials (e), (f) and (h). They must also design the instructional techniques to be used with these materials (g). They then are able to move into stage 5 at the same level apparently as those who chose the other route. This equality of stages may be misleading,

however, since those who take the route (b) to (h) may need to recycle
through stage 4 several times. If pre-selection criteria were adequate
this is less likely for those who chose the (a) to (d) route.

Stage 5: Pilot Trial

At this stage, the design of pilot trials using an appropriate
evaluation design, pre-test, trial initiation and the gathering of
formative evaluation data are common steps, but alternate routes become
available after this point. The decision as to route is based on results
of the formative evaluation and the possible routes involve program
adoption, adaption or rejection. If the data suggest adoption (e) the
route is directly into stage 6. An adaption decision (f) may suggest
recycling as little as simply beginning another trial with minor alterations
or as much as beginning back in stage. 4. The rejection decision takes the
innovator back to program or goal selection in stage 3.

Stage 6: Adopt, Adapt or Reject

A decision to adopt at stage 5 leads to stage 6 recycling of the
innovative program within the innovative school (a) and the gathering of
additional evaluation data. On the basis of this data adopt, adapt or
reject decisions are once again possible. An adopt decision leads to
reporting to other interested schools the results of evaluation and such
data then are considered to be used in the summative mode. This report of
data may lead to plans for program field trials in other schools who would
be at point (f) in stage 3. An adapt decision, depending on the size of
required adaption, may lead back to any point in stage 4 and even back to (h)
in stage 3, if problems are very fundamental. A reject decision might lead
out of the change process.

Stage 7: Field Trial

The steps in stage 7 are not clear at this time and probably will not be until school programs reach this stage - a stage only one program is now entering. It appears, however, that designs for evaluation and diffusion are integral parts of this stage as is a carefully planned evaluation of this diffusion model.

Conclusion

Evaluation as it has been elaborated in this paper is an integral part of a comprehensive program for school change. The Trent Valley Centre staff suggest that evaluation becomes most meaningful when it is set in a curriculum development or school change activity that has goals, procedures, and people that are to be served by evaluation. What appears to be an over emphasis on the TVC model for change should be interpreted as a further indication of the great importance that the TVC places on setting.

The Trent Valley Centre program which is elaborated as a diffusion model is in one sense a summarization of the activities that have been under-way within the TVC and its participating schools, over the past two and half years. In another sense it is a description of a pattern of interrelated tasks that other schools and school systems may use to advantage, knowing that each part of the model has been tested out a number of times in the Peterborough setting, and as the so called model is tried in other counties than Peterborough and as the Peterborough experience unfolds throughout the totality of county schools, there is some reasonable prospect that a more refined and detailed model will emerge.

Perhaps the most important feature of the Trent Valley Centre

program, insofar as its impact on evaluation procedures is concerned, is the central role of the teacher as a curriculum decision maker. There is no intention here to isolate the teacher as the sole decision-maker on curriculum matters, but on the other hand, by identifying the classroom teacher as a primary decision maker, there is a definite shift away from the common school practice of reserving major curriculum decisions for administrators and senior consultants. This downward decision-making shift has implications for evaluation, curriculum development and school organization. In the case of evaluation, the potential for school-specific and class-specific -- and possibly even pupil-specific -- objectives means that there will be a short feedback loop for data and rather than going to a county or a national office, they go back to the teacher and the student concerned. Decisions concerning school programs, class programs and individual programs are based on the data specifically related to these programs.

In cases where objectives are carefully elaborated for schools, classes and/or individuals, it is an easy step to generate criterion-referenced measures and another easy step to gather and report criterion-referenced data. These same test items that are used in a formative sense to make adopt-adapt-reject decisions at the classroom level can be used at a later point in time to make adopt-adapt-reject decisions about total programs or major sub-components of total programs. These same types of data used then in either the formative or summative sense to make adopt-adapt-reject decisions are in all cases gathered according to a design which is some instances involves a parallel control group operating simultaneously with the experimental group, while in other and perhaps

In most instances the control group is a previous year's cohort or class passing through the same grade with the same teacher in the same school, but under different program conditions.

At one point in the paper the issue of competitiveness was discussed and some hope was offered that the use of criterion-referenced tests, as opposed to norm-referenced tests, could, in fact, reduce competitiveness and at the same time preserve interest in the individual for himself as opposed with his role within the group. The fact that norm-referenced tests result in a ranking of students along a continuum seems to encourage competition among individuals, but on the other hand the criterion-referenced measures by virtue of their attention to specific goals could under some circumstances set the stage for even more games manship and competition. Perhaps those who want to reduce the amount of aggressiveness and competitiveness in students will have to find ways of achieving their end that are independent of the testing procedures introduced to assess student progress.

It is important to suggest, as a last point, that the Trent Valley Centre program has avoided the issue of invalidation of data through the creation of false incentives to increase student scores. The reference here is to the short feedback loops that result in teachers having access to student performance data for the purpose of making program decisions within their classrooms and within their schools. County and province wide decisions pertaining to schools and teachers are not made on the basis of these same data, nor are school grants or dollar incentives to teachers affected by these data. Any change in the incentive system can serve to invalidate the data and, as well, it may lead to the parrot-like learning

that students achieve readily under the pressure of a performance contract
that has payoff totally unrelated to the intrinsic value and the utility
value of the learning itself.

REFERENCE

Bennis, Warren G.   Changing Organizations, New York: McGraw-Hill, 1966

Bogue, E. G.   Disposable Organizations   Phi Delta Kappan   1971, 53, 94-96.

Campbell, D. T., Stanley, J. C.   Experimental and quasi-experimental
   designs for research on teaching.   In N. L. Gage (Ed.) Handbook of
   research on research on teaching.   Chicago: Rand McNally, 1963.

Dewey, John   Democracy and education.   New York: The Free Press, 1966.

Horn, D., Russell, H.   Evaluation materials package.   Development Proposals.
   Development proposal funded by OISE office of R. & D. 1971.

Ireton, H. et al   Infant mental development and neurological status,
   family socioeconomic class, and intelligence at age four.   Child
   Development, 1970,   , 937-945.

Leithwood, K. A.   Evaluating achievement of educational objectives.
   Orbit 9, 1971, 2 (4), 10-11.

Leithwood, K. A., Russell, H.   Science Objectives & Item Pool.   Development
   proposal submitted for funding to the Government of Ontario Grants-in-aid
   Pool, 1971.

Maguire, L. M.   Observations and analysis of the literature on change.
   Philadelphia: Research for Better Schools, Inc., 1970.

Moore, Evelyn   The way it is in curriculum development, Part I: A note on
   the need for curriculum development studies.   Curriculum Theory Network
   Monograph Supplement, 1971.   12-14.

Pillet, Roger A.   Boundaries of a Curriculum Network.   Curriculum Theory
   Network Monograph Supplement, 1971, 7-11.

Popham, W. James   Probing the validity of arguments against behavioral goals.
   In Kibler, Robert J., Barker, Larry L. and Miles, David T.   Behavioral
   Objectives and Instruction.   Boston: Allyn and Bacon, 1970.

Popham, W. James   The instructional objectives exchange: new support for
   criterion-referenced instruction.   Evaluation and measurement newsletter
   OISE, 1971, No. 10.

Scriven, M.   The methodology of evaluation. Perspectives of Curriculum
   Evaluation, 1967, 1, 39-83.

Stake, R. E.   Toward a technology for the evaluation of educational programs.
   Perspectives of Curriculum Evaluation, AERA Monograph Series on Curriculum
   Evaluation.   Chicago: Rand McNally, 1967.

Whitehead, A. N.   The aims of education.   New York: MacMillan, 1956.

REFERENCE (cont.)

Denison, E. F. The sources of economic growth in the United States and the
alternatives before us. Supplementary Paper No. 13, Committee for
Economic Development, 1962.

Denison, E. F. "Appendix to Edward F. Denison's reply." In the Residual
Factor and Economic Growth. Organization for economic cooperation and
development (Paris), 1964.

Dewey, John Democracy and education. New York: The Free Press, 1966

Galbraith, K. The Affluent Society. New American Library, 1956.

Olivier, W. M. and Churchill, S. "Computer-Assisted Instruction at the
Ontario Institute for Studies in Education." In AEDS Monitor
February 1971

Whitehead, A. N. The aims of education. New York: MacMillan, 1956.