DOCUMENT RESUME

ED 142 615 UD 017 098

AUTHOR	Horst, Donald P.
TITLE	Analysis of School Projects for the Development of Project Information Packages (PIPs).
PUB DATE	4 Apr 77
NOTE	27p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, N.Y., April 4-8, 1977) ; Tables may be marginally legible due to print size
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Compensatory Education Programs; Educational Assessment; Elementary Education; *Instructional Programs; *Models; Reading Achievement; School Personnel; *Summative Evaluation; Teacher Influence; Teacher Morale
IDENTIFIERS	*Project Information Packages

ABSTRACT

Six compensatory education projects, all of which were effective, are analyzed. The purpose of the Project Information Packages (PIPs) mandated by the United States Office of Education (USOE) was to determine whether compensatory projects could be packaged with sufficient clarity and in sufficient detail to encourage and enable their replication at sites where educational needs had not been adequately met by existing programs of instruction. The three stages of replication: selection/adoption, start-up, and operation are examined and analyzed. The major problems in describing the replication mechanism are explored. These concerned the differences between effective projects in the areas of: management and instructional features of the projects and in the role of the individual teacher in determining the instruction of the student. (JP)

ANALYSIS OF SCHOOL PROJECTS FOR THE DEVELOPMENT OF PROJECT INFORMATION PACKAGES (PIPs)

Donald P. Horst

RMC Research Corporation
Mountain View, California

Paper presented at the Annual Meeting of the American Educational Research Association, April 4-8, 1977, New York City.

ANALYSIS OF SCHOOL PROJECTS FOR THE DEVELOPMENT OF PROJECT INFORMATION PACKAGES (PIPs)

The analysis of projects for the development of Project Information Packages contrasts sharply with the analyses described by the previous speakers. Where the other studies have analyzed a range of programs looking for features which distinguish effective programs from ineffective ones, RMC analyzed a small number of projects, all of which were effective, looking for the unique features which led to the success of each project. Since no ineffective projects were included in the study, no dimensions could be defined which would distinguish projects in terms of effectiveness. In any case, the projects in question shared few common features at the level of specificity required for diffusion to new sites.

Certainly, a statistical search for common features would not have helped in this effort to package projects for dissemination. Instead, a form of systems analysis was used to identify project elements which contributed to project effectiveness, and this analysis is described below. Before turning to the analysis, however, I would like briefly to describe the history of the packaging studies in which RMC has been involved. I would also like to consider two issues which are central to current efforts to isolate and communicate essential project features: (a) Just how effective are exemplary projects in compensatory education? and (b) Where should we be looking if we want to find the things that produce apparent achievement gains?

History of the Packaging Studies

The USOE has long felt that the effectiveness of federally funded programs--particularly those intended to provide compensatory services to disadvantaged children--could be greatly enhanced through the diffusion of education projects and practices proven to be effective. It became clear recently, however, that the simple dissemination of information about such projects and practices was not a sufficient mechanism for bringing about widespread replication. For this reason, a decision was made in 1973 to attempt a more complete packaging of several exemplary projects. The purpose of this undertaking was to determine whether projects could be packaged with sufficient clarity and in sufficient detail to encourage and enable their replication at sites where educational needs had not been adequately met by existing programs of instruction.

In June 1973 a twelve-month research and develpment contract was awarded to the RMC Research Corporation by USOE to identify up to eight compensatory education projects and develop replication packages for them. Six exemplary projects were identified, and packages were delivered to sites for field testing late in the summer of 1974. A description of the entire contract effort is presented in a report by Tallmadge (1974).

A two-year evaluation of the PIP field test was the second step in answering the question of PIP viability. The field test consisted of attempts by two to five sites to replicate each of the six projects, a total of 19 replications in all. The evaluation of the field test was conducted by SRI with RMC as a subcontractor. The first-year evaluation activities included five visits to each site to observe progress and collect student test data. RMC participated in these site visits and provided general support but was primarily responsible for developing recommendations for PIP revisions (Horst, Piestrup, Foat, & Binkley, 1975).

The general conclusion of the field test was that the PIPs were fairly successful in enabling accurate replication of the exemplary projects (Stearns, 1977). However, the field test also provided many new insights into project analysis and PIP design. Based on these insights, the original projects were visited again and re-analyzed, the

PIPs were redesigned, and completely new packages were produced. These second-generation PIPs are currently being disseminated on a nationwide basis. This diffusion effort is being evaluated by the American Institutes for Research, with RMC as subcontractor.

How Effective Are Exemplary Projects?

While the nominal purpose of this series of studies has been to explore the potential of packaging for the diffusion of exemplary projects, the ultimate goal has been to help solve the reading and math problems of educationally disadvantaged students. In particular we have been asked to identify and package the "essential elements" of the exemplary projects; that is, those elements which will, if replicated in other sites, produce the same achievement gains obtained at the original project sites. Unfortunately there are three interrelated problems which limit the ability to diffuse project effectiveness. Each of these problems is discussed briefly below.

The basic model is wrong. The fundamental problem has been in concentrating on the impact of the "project" while treating other sources of impact rather lightly. To belabor the point, we have acted as if

$$\text{Gain} = f(\text{Project}).$$

In fact of course, if the project is replicated in very different surroundings, with very different students and teachers, gains can be expected to differ too. In addition, we have found that the way in which achievement is measured has a major effect on apparent gains. In other words, if we want to examine project success we must start with a more realistic equation:

$$\text{Apparent Gain} = f(\text{Project, Context, Measurement}).$$

I will not attempt to develop this equation in any serious way, so precise definitions for the three independent variables are not too important. In general, the term "project" is intended to include only those elements of the project which can be packaged for diffusion, such as instructional techniques, management procedures, materials, and so on. The "context" is a catch-all concept including the students and their environment, school personnel with their skills and attitudes, and all other project features that cannot be exported. "Measurement" includes the tests which are used, the evaluation design, and the testing procedures. It might also include "test wiseness" and similar sources of test score variance which involve student learning but which we would not want to include under the heading of the "project." No one would doubt that all three classes of variates affect apparent gain, yet as we shall see, their relative importance is often overlooked.

Project effects are small in relation to our expectations. The purpose of compensatory projects is to take students who are not acquiring adequate skills and teach them those skills. Of course, we all profess to be realistic in our expectations as to what is possible, but a review of some basic components of test score variance can be quite sobering. Figure 1 illustrates the relative magnitudes of three possible components of reading scores as measured with a standardized test. The first component is the individual differences among students. These differences, as we all know, are very large. The second component is the normal expected skill increase over the course of a year for a given student. It includes the effects of (a) the regular school program, (b) maturation, and (c) all out-of-school learning. At grade 2, the sum of these three effects is small when compared to the differences among students. At grade 6 it is smaller still. The third component is the difference that an effective project might make. For the purpose of illustration, a value of one-third of a standard deviation with respect to the national norm group is shown in Figure 1. While this value is rather arbitrary, in practice we have been able to find very few compensatory projects with convincing evidence of even this amount of impact.

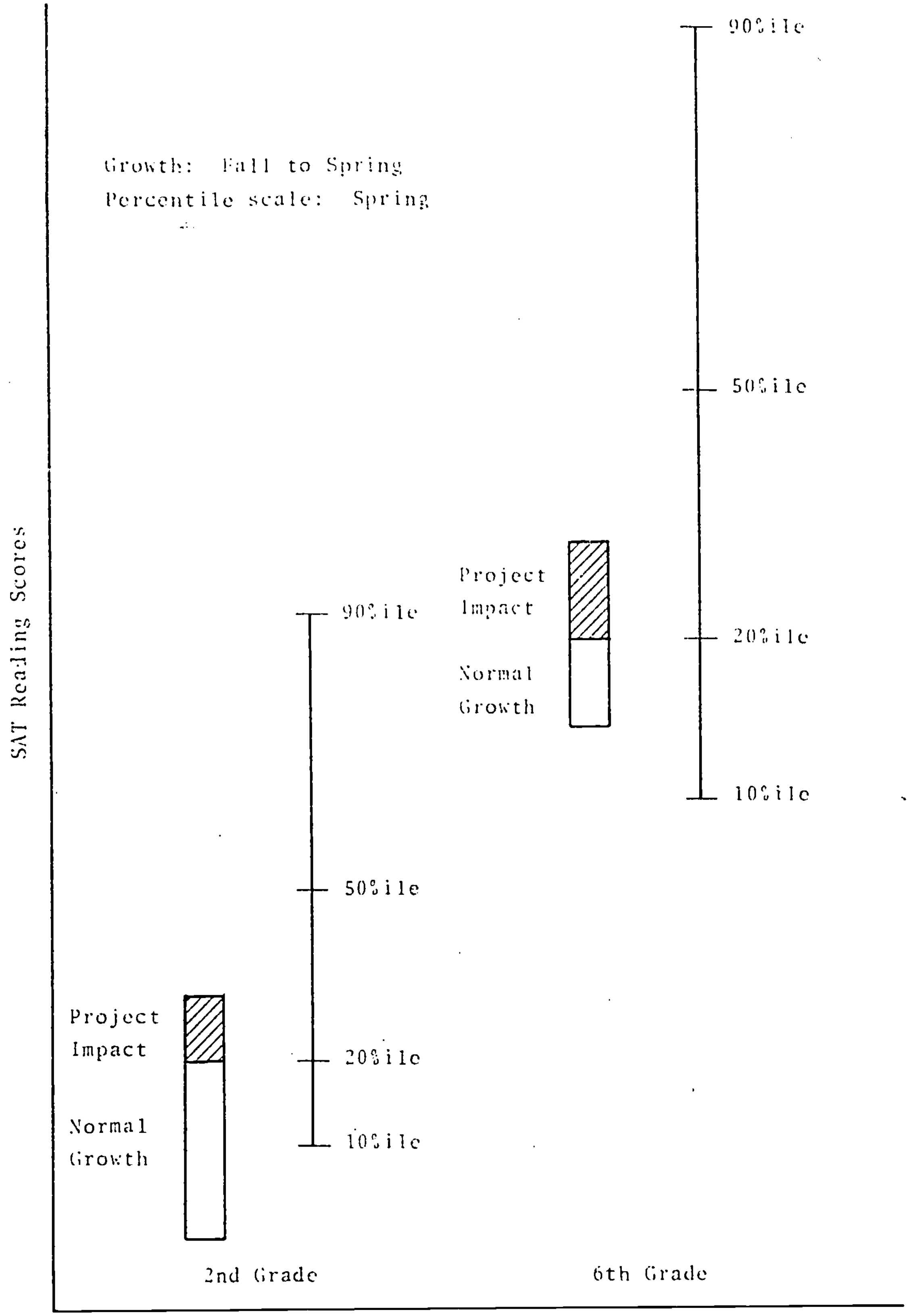

Figure 1. Comparing reading growth of 20%ile students against individual differences in reading.

Unfortunately, it seems that the scale of individual differences is often used as the frame of reference for judging project effectiveness. We know, for example, that the ninetieth percentile student, while not a genius, reads quite well. The tenth percentile student reads so poorly that special help is badly needed. We tend to use the fiftieth percentile student as our goal for compensatory projects, perhaps because fiftieth percentile reading skills seem minimally adequate, perhaps because we feel that no American should be below average in anything. The problem is, as Figure 1 suggests, that such a goal implies the need for truly dramatic changes in learning rates; far more dramatic than our experience tells us is possible with available instructional techniques. When faced with these facts, there is a tendency to become discouraged and to say in effect, if we can't have the moon, why even try? RMC does not take this position, but before considering what we would advocate, there is a third problem to discuss.

Project effects are small in relation to context and measurement variance. The problem is not simply that project effects are small in relation to our aspirations. They are also small in relation to the vast sea of confounding contextual and measurement effects from which we try to separate them. The effects of contextual variates are not well quantified, but I feel most educators would agree that a project which appears effective in middle-class suburbia may not have the same success in an inner-city setting. Similarly, a project which helps students whose main problem is limited skill in English may not have as great an impact on native speakers of English with more severe educational deficits, and a project which depends on exceptional teachers may not work as well with teachers having lesser skills. I would go further, and say that the context can make or break a project.

Measurement variance is somewhat better understood. We know, for example, that for a project which apparently meets our arbitrary standard (one-third of a standard deviation gain due to the project), half of that gain could easily be due to regression effect error (Anderson & Estes, 1976). The use of grade-equivalent scores may account for a comparable spurious gain (Tallmadge & Horst, 1976, p. 74). Dropouts from the low end of the range can inflate gains tremendously. For example, if the lower 40% of a group is dropped between the pretest and the posttest, a one-third standard deviation loss in mean score will appear to become one-third standard deviation gain (Roberts, 1977). Testing too early or too late can also make large apparent differences (Horst, Tallmadge, & Wood, 1975, p. 13). In addition, the test which is used can be a major factor, since a one-third standard deviation gain may represent as few as two or three raw score points. Even changing from one standardized test to another can make huge differences in scores. The effect of changing to other kinds of tests may be much larger. Clearly, then, the match between test items and instructional content can be critical (Bianchini, 1976). The general picture is one of many variates which might produce roughly equal impacts on apparent gain. The project is only one such variate, and not necessarily the dominant one.

Finding the Essential Elements

Given that project impacts as measured by standardized tests are small and elusive, one may ask whether there is any justification for pursuing the critical project elements which produce these illusive gains. I would say yes, because it is clear to any observer that the best projects are better than mediocre ones and far superior to no treatment at all. Giving all students access to the best projects available is still a worthwhile goal. While we cannot now support this assertion by pointing to impacts on student achievement, I believe that few educators would really disagree. However, some cautions are in order.

Limitations of multivariate statistical techniques. The other three studies described in this symposium represent straightforward attempts to find the elements common to effective projects, or more precisely, those elements which distinguish effective projects from less effective ones. RMC could not apply such procedures for two reasons. First, we had only six projects to analyze in depth, and they were all effective. In fact, only effective projects were even nominated for our screening procedures. Thus we had no ineffective projects to provide contrast. Secondly, the six projects we examined shared little in common in terms of instructional approaches. Thus, concentrating on common features would have eliminated all instructional features from further consideration.

This is not to say that common features are unimportant. Our impressions, based on our intensive experience with these six projects and our informal experience with many others are completely consistent with the findings of the other studies. There certainly are general features which characterize effective projects, but I would classify most of them as management rather than instructional features. Furthermore, while they are probably among the most important features to communicate to a new adopter site, they do not tell the complete management story. In particular, they cannot include the unique approaches developed in exceptionally innovative projects. Statistical techniques have helped to identify the common management features of effective projects, but for purposes of diffusion, we must also isolate those features which are unique to each effective project.

My impression is that, while well implemented projects tend to share certain effective management features, any reputable instructional approach can work, if it is implemented well. However, it would also be wrong to conclude that specific instructional features are unimportant in a diffusion effort. It seems obvious that different instructional approaches may teach different skills, and that a given approach may be more effective in one context than another, even though our current measurement model is so crude that we cannot say exactly what the various approaches accomplish. It is also clear that the insructional features of an effective approach will be the first thing that a potential adopter will look for. The fact that <u>many</u> different approaches are comparably effective does not imply that <u>all</u> approaches will work, or that all approaches will be equally attractive to all educators. Therefore, if we intend to analyze projects for the purpose of diffusing them to other sites, our analysis process must identify the important instructional features as well as the management features.

Diffusing gains vs. diffusing projects. At the beginning of the PIP development effort, RMC was focusing on a rather narrowly defined question. The question was this, "Given that a particular effective project is to be diffused, to what extent can a package of printed and other materials replace the services of a technical consultant?" This is an important question, because diffusion via a technical consultant is expensive, and the consultant (usually the project director who developed the project) has a very limited amount of time. The method of project analysis used by RMC was developed in direct response to this question, and it focuses on identifying project goals and processes. As an aside, I believe that we can now answer the original question rather well. In fact, if we combine what has been learned in the PIP studies with the results of several concurrent diffusion studies, we can even give pretty good answers to the more general questions of how to diffuse projects effectively.

However, as I mentioned earlier, these were never the real PIP questions. The real question was whether we could package the large achievement gains reported by a small number of exemplary projects and diffuse those gains throughout the entire educational system. Analyzing projects with a view to replicating their gains involves much more than capturing their goals and processes. Specifically, it requires careful analyses of the project "context" and "measurement." Only where the impacts of these variates are fully understood can we be confident that any kind of "project" analysis will contribute to producing achievement gains in other school districts.

Diffusing measurement approaches and contexts. Since measurement and context seem to be major factors in the apparent effectiveness of projects, a brief comment on the possibility and desirability of diffusing these components is warranted. The issue is not a facetious one. Measurement effects are generally exportable and fall into two categories. The first category covers errors in evaluation methodology. It is difficult to prove, but my impression is that a substantial proportion of reported project impacts are the result of widely accepted but inappropriate evaluation practices. These practices are easy to package, but we probably do not wish to diffuse them.

However, the second category of evaluation practices is equally easy to diffuse and should probably be strongly encouraged. These practices include selecting relevant tests, and in general designing evaluations which capture the exact learning that the project has produced. It is important not to pretend that specific new skills are the same as a general improvement in reading or math, but it is equally important not to belittle real impacts on student skills simply because those skills are not accompanied by big jumps in standardized test scores.

The concept of diffusing project context may at first seem absurd, since context was defined as those features which cannot be packaged for export. However, while strictly speaking context cannot be exported, it is possible to be sure that there is a good match on critical context variates before a new site attempts to adopt a project. The concept that the context in a new site must be compatible with a given project has been basic to PIP development from the beginning and much of the PIP selection material focuses on critical context variates. Unfortunately, it has been difficult to get anyone to take context considerations seriously, and projects frequently end up in inappropriate settings with predictable results.

Analyzing a Project for Diffusion

Independent of the problems in finding effective projects, if a project contributes to student achievement gains, other districts may very well wish to adopt it. As I have noted, both management and instructional features may be included among the important components of project success. In fact, one of the major features of the original PIP projects has been their longevity. All six have survived for more than ten years. Obviously, a project which does not survive does not help students, and survival requires much more than demonstrating achievement gains. Over the past three years, RMC has evolved an analysis procedure which has proven very useful in sorting out the many goals and procedures which make up a successful project. Essentially, it is a systems analysis which organizes the various initial conditions, processes, and outcomes in the project. Since the intent is to let other sites replicate the original project, the project is described as a component of a replication system.

Stages of Replication.

The replication of a project is divided into three stages: (a) selection/adoption, (b) start-up, and (c) operation. Selection/adoption is the process whereby sites select the projects that suit their needs and apply through appropriate channels for funds. Start-up begins with the delivery of a PIP and associated funds to a suitable site, and ends when the project is ready for operation. In general, start-up is completed by the end of the summer vacation. Operation is the instructional and administrative activity of the school year, beginning with the first contact between teachers and students for instruction or testing.

Each of the three stages is further broken down into inputs, processes, and outcomes, and these substages are organized to reflect the basic project components of personnel, other resources, and students. In this model all three components are receivers of inputs and processes, not initiators. For example, in the personnel section, the project is described in terms of what is done to teachers, not by them. The resulting descriptive model is summarized in Figure 2.

The selection/adoption stage is included in Figure 2 for completeness, but it is incidental to the analysis of the project, and will not be discussed further here. However, several other features of the replication model should be discussed. The intended outcome of the selection/adoption stage is a site that meets the requirements for replicating the project chosen. This outcome is described in terms of the characteristics of personnel, other resources, and students at the site, and constitutes a major component of the inputs to the start-up stage. The remainder of

	Selection/Adoption				Start-Up				Operation			
	Inputs	Processes Activities	Time	Outcomes	(New) Inputs	Processes Activities	Time	Outcomes	(New) Inputs	Processes Activities	Time	Outcomes
	Sites which need projects Dissemination mechanism (ASK)	Sites select PIP projects		Sites suitable for project replication	PIP Techical assistance Money Constraints (Field try-out inputs)	Installation of operational project		Project ready for student selection/ instruction	PIP-specific new inputs (minimal)	Operation of project		Successful project
Descriptive Conventions												
Personnel Project Director Project Staff Non-Project Personnel				Roles Skills Attitudes	PIP: Information Self-training materials Job aids: PR materials Calendars Sample Forms	Establishing roles Selection Training		Roles Skills Attitudes		(Establishing roles) (Selection) Training		Roles Skills Attitudes
Other Resources Materials Facilities				Availability Adequacy Acceptability Related systems	PIP: Information Sample materials	Selection/ Ordering Distributing/ Allocating		(Availability) Adequacy Acceptability		(Selection/ Ordering) Distributing/ Allocating		(Availability) Adequacy Acceptability
Students				Skills Attitudes	PIP: Information	(Selection)		Skills Attitudes		Selection Training		Skills Attitudes

Figure 2. Descriptive model of the PIP replication mechanism.

inputs to start-up include the PIP, money, technical assistance, and other inputs that are supplied to the site to help in replicating the desired project. In other words, assistance supplied to the site is a subset of the inputs to the start-up stage. Inputs to the start-up stage also include the site itself with its personnel, other resources, and students.

Similarly, the outcomes of start-up are inputs to the operation stage. However, unlike start-up, there are very few additional inputs during operation. This is a consequence of a basic PIP replication concept: the concept that a package of information (and associated funds) can be sufficient to produce the replication of a project if delivered to an appropriate site. To the extent that this concept proved unworkable in the PIP field test, new inputs in the forms of monitoring, technical assistance, and the like have been identified for the operation stage.

The process column under each stage includes the activities carried out at that stage and the timing considerations for each activity. Timing includes the amount of time allotted for each activity, the sequence of activities, and the period in the school year during which the activity is expected to take place.

Descriptive Conventions in the Model

The lower section of Figure 2 displays the specific descriptive conventions (see Horst et al., 1975, Appendix D, for definitions) adopted by RMC in categorizing field test data. Personnel are divided into (a) the project director, (b) other project staff, and (c) nonproject personnel. These categories were adopted because of the qualitatively different ways in which the PIP mechanism interacts with them. The project directors are central to the project replication. They are intended to be selected by the time of the site proposal or grant request, and most of the PIP materials are directed to them. While the PIPs provide them with some job aids in the form of calendars, sample forms, and the like, it is assumed that they possess most of the required characteristics at the time they are selected, and much of their PIP material simply describes what they should accomplish, with minimal suggestions on how to proceed.

Personnel inputs and personnel outcomes are described in terms of the titles of the personnel required, and their required characteristics. Characteristics of personnel, or outcomes, are categorized under skills, attitudes, and quantity. In addition, for the purposes of the replication model there are critical roles, interrelationships, job positions, and the like, which are not considered to be either skills or attitudes. Thus, for example, the amount of authority given to a project director by the district administration may be one of the most important factors to describe. A third category of roles is provided to encompass such personnel descriptors.

Processes related to personnel are broken down into selection and training. Training is taken in the broad sense to include any activites designed to change skills or attitudes. In addition to conventional skill training, this definition encompasses all orientation and instructional activities. Procedures designed to create roles or authority relationships may not fall under either selection or training (e.g., giving a project director the task of hiring teachers in order to establish his or her authority over them). Such procedures are included under a third category, establishing roles.

The remaining column is labeled "(new) inputs." During start-up, these inputs include the PIP, other technical assistance, money, constraints, and, for the original six PIPs, field-test events. Most of these inputs should be generally self-explanatory and no special descriptive conventions were adopted. The exception is the content of the PIP, for which the following conventions are used:

Information includes descriptions of tasks and activities in terms of intended outcomes and their sequences. The significance of this category for the development of PIPs is the assumption that project directors and other staff will know how to accomplish the particular tasks and activities described in this brief manner.

Self-training materials are intended to help personnel acquire new skills, and attitudes, and range from informal tips and suggestions to, in one PIP, a programmed tape/slide training sequence. While it is not always possible to categorize materials unambiguously under information or training, the distinction is helpful in developing a package because it makes explicit the assumptions on the extent to which various personnel will come to the project with the necessary skills or will require PIP assistance in acquiring them. By and large, the PIPs were designed as information packages with only limited self-training materials. The major project training activity, training of project staff, was generally assigned to the new project directors, and it was assumed that they will have most of the basic skills needed to conduct such training.

The major remaining component of a project is its students. Students are described under the same conventions as personnel. Their characteristics (outcomes) include skills, attitudes, and quantity; selection and training are the processes in which they are involved.

Most of the description of students and the processes that affect them are included under operation processes and outcomes. While student outcomes in terms of skills and attitudes constitute the ultimate goals of any educational project, the emphasis of the PIP replication mechanism is on the replication of selected instructional processes. These processes are included under the operation process of student training. It should be noted that these instructional processes can be defined, in large part, in terms of teacher behaviors and interactions with materials. Thus, student training will encompass the operation outcomes for personnel and other resources.

Applying the Model

Conceptually, the model in Figure 2 provides a simple framework for isolating the critical elements of the PIP replication mechanism. In practice, however, there are major problems in describing the replication mechanism in terms of this or any other model. In particular, the number of specific items that could be listed under each heading is virtually infinite, so there is no practicable way of describing the mechanism exhaustively. For example, consider the problem of describing the skills and attitudes of the project director which are assumed to exist at the outcome of the selection/adoption stage. It is clear that the project director must have appropriate technical and managerial skills and must have a generally positive attitude toward the project. However, enumerating all of these skills and attitudes at a detailed level is clearly impractical, as it would be necessary in the process to exclude all the possible human characteritics the project director must not have. Attempting to list all the processes intended to modify personnel characteristics presents a similar problem.

Early attempts by RMC to produce appropriate lists led to considerable frustration. The lists quickly became long and unwieldy with details that were obvious or trivial, yet when used as a basis of comparison for the field-test sites, the lists never seemed to have a place for the critical site-specific problems. What was needed was a list of categories to systematize the description; this would provide an overall perspective on the mechanism in question and would highlight those parts of the mechanism requiring revision. Clearly, a systematic procedure for generating such a description was needed.

The procedure which was adopted took as its primary input the field-test data on problems encountered by sites. The rationale for this procedure was simply that any aspect of the replication process that did not go as intended at one or more sites indicated a possible defect in the intended replication mechanism and a potential point for a recommended revision.

The data used to develop the model came from a variety of sources. The reports of site observations by SRI, RMC, and USOE site visitors were the major source of data. While it was not practicable to list every sense in which sites were replicating successfully, it was quite possible to list the major ways in which sites are deviating from intended practices. Included were examples of observed failures to replicate, problems observed, user comments on reasons for problems, reported ambiguities in the PIPs, and lack of information in the PIPs. Observers obtained verbal reports, reviewed marginal notes written in the PIPs by project staff, and looked for specific evidence of use or nonuse of PIP components. In addition, all requests by project staff for assistance from outside sources were monitored via contact report forms.

The first step in the analysis procedure was to screen all of the data described above. All reports and other data sources were systematically reviewed, each problem or deviation was entered on a separate card, and the cards were sorted into the various categories shown in Table 1. Initially, cards from each site were processed separately so that 19 different sets of cards resulted.

The resulting picture of the problems at a given site was incomplete. While the outcomes of start-up were generally portrayed fairly systematically, there were gaps in the preceding columns. The practical problems faced by the site observer virtually preclude recording every possible aspect of each process and every possible characteristic of each person involved in the project and, in many cases, the critical things to look for became apparent only after an associated outcome went awry. To identify and fill in the gaps, RMC used the following procedure: the start-up outcome problems for a given site were organized into manageable units. For example, one unit might consist of required teacher skills which were not present at the end of start-up. The intended mechanisms which were designed to produce the skills were then systematically reviewed beginning with other start-up outcomes, then working backward through the related start-up processes, start-up inputs, and selection/adoption outcomes.

To be specific, when looking for the reasons why teachers lacked certain desired attitudes, the other start-up outcomes to consider would include, at least, the expected attitudes of the project director toward the project and the availability and adequacy of the materials and facilities. Next, start-up processes would be reviewed to see where breakdowns

Table 1

FACTORS LEADING TO LACK OF TEACHER ENTHUSIASM
IN A PROJECT R-3 SITE

	Selection/Adoption Outcomes	(New) Input	Process: Activity	Process: Time	Outcome
Personnel	• Project director uncertain if R-3 can be replicated this year • Project director is hired as "helping teacher" instead of "Project director" • Project director has little administrative experience • Project director was not involved in or identified in grant application • Project director did not have district job appropriate to PIP role	• Inadequate funds allocated to pay two additional teachers (total of 12 is needed, project has 10) • PIP is not clear on how many planning periods teachers need per day • PIP does not discuss alternative staff configurations for more than or less than 250 students	• Principal rather than project director hires staff • Principal assigns existing seventh grade teachers to R-3 • Self-orientation for project director is limited • Project director has difficulty describing gaming/simulation to staff • Project director selects overqualified aides (80% have their B.A. or M.A. degrees) • Staff orientation lacks enthusiasm	• Only one month instead of anticipated four months for start-up • Orientation occurs after schools open in September	----------- Teachers are not enthusiastic about the R-3 project ----------- Project director expresses lack of confidence in being able to replicate R-3 for the current school year
Other Resources (Materials, facilities)	• Inadequate lead time provided for obtaining other resources on PIP schedule* • Adequate copies of R-3 materials (from exemplary site) not available • R-3 materials viewed as relating to a different socioethnic group • District/school channels for approval of orders in conflict with PIP • Appropriate space for project director's and secretary's office not available • Appropriate classroom space for teachers not available	• Inadequate funds for 345 students • PIP is not clear on which resources are essential to purchase • PIP does not state rational for purchasing certain facilities (e.g., carpeting and hexagonal tables) • PIP does not contain an adequate number of brochures on the core materials • Information on which gaming/simulation materials to order is unclear • Some gaming/simulation materials are illegible	• Facilities are not ordered (no carpeting and hexagonal tables ordered) • Classroom space for each teacher is not allocated • Materials and facilities are not distributed to each classroom as specified	• Ordering begins in August instead of May	Classrooms are not ready for instruction Adequate R-3 materials not available
Students	• Number of students to be served does not match PIP configuration (345 as opposed to 250)		• Students not grouped in groups of 20-22 (students grouped alphabetically in groups of 16-30)	• Project director has less than two weeks for grouping and scheduling as opposed to an entire month	Inappropriate pupil-teacher ratios

*Technically, time falls under selection/adoption processes.

occurred in training the teachers or ordering materials. Then start-up inputs would be examined to determine where the PIP appeared deficient, where the outcomes of selection/adoption were not as expected, and where money, USOE constraints, local policies, field-test inputs, and the like, affected training or ordering processes.

As ea category of the mechanism was considered, there were three judgments to be made: (a) the category does not bear on the problem under consideration (e.g., lack of desired attitudes), either directly or through any relation to an intervening category that relates to the problem; (b) the category is relevant to the problem, but cards are already prepared for the deviatons at this site that fall in this category and affect the problem; (c) there appear to be relevant deviations in the category that were either overlooked in screening the data or not recorded by site visitors. Where a deviation was simply overlooked, a card was prepared and included in the appropriate category. Where the information was missing (e.g., where there was no information as to why specific materials were unavailable), a note was made to obtain the information by phone or during a subsequent site visit. An example of the resulting site description is shown in Table 1, abridged somewhat for illustrative purposes.

Once this backward (from "outcome" to "input") review process was complete for a given problem at a given site, it was repeated for the next start-up outcome and so on until all of the unintended outcomes were covered. As may be observed from Table 1, beginning with teacher attitude problems immediately brings in a variety of other outcome problems. Thus, successive outcome problems can be processed quickly, since many of their causes will have been previously noted.

Before the process was considered finished for the site, a "forward" review was also completed. This involved taking each deviation in the selection/adoption outcomes and looking for expected problems at successive stages. This review was carried out in a manner analogous to that of the backward review, with each category of each stage given individual consideration. The result of this entire procedure was a set of cards for each site, categorized according to the model in Figure 2, and detailing for each site the critical steps in the replication mechanism where problems occurred. An example of the content for one site (start-up personnel processes) is presented in Horst, et al. (1975a).

It should be reemphasized that the description of individual sites was in terms of the actual problems they encountered. The next step in processing the field-test data was to combine the data from all sites that used a given PIP and to restate the problems positively in terms of the processes and characteristics that make up the intended replication mechanism. The products of this step were six descriptions, one for each PIP, of the project replication mechanism. These descriptions are, in essence, similar to those developed by SRI/RMC from the PIPs as the first step in the study. The two major differences are that the stages of the replication mechanism are explicitly modeled, and the steps which have proven critical in the field test are systematically included in the descriptions. These descriptions provided the required basis for specifying required PIP revisions.

For the purposes of describing the overall findings relevant to PIP revision and ensuring that formats of PIP descriptions are consistent across PIPs, one final processing step was undertaken. The descriptions for the six individual PIPs were combined and summarized into a generalized model of the PIP replication mechanism. Basically this model was an elaboration of the model shown in Figure 2, but, like the individual PIP descriptions, the model reflects the experience of the field test.

Generalizations from the Analysis

While the analysis described above did not produce a set of features common to effective projects, some generalizations about the differences between effective projects can be made. One, which has been mentioned above, relates to the distinction between management and instructional features of a project. Certain management features (e.g., maintaining good relations with non-project personnel) are shared by many effective projects as the other speakers have noted. Others vary widely (for example, the degree and techniques of control exercised by the project director). Many of the management techniques are focused on project survival and have little direct influence on the instruction received by the students.

A second dimension on which effective projects differ is the role of the individual teacher in determining the instruction of the student. At one extreme might be projects using computer-controlled terminals in which the only "teacher" is a monitor who maintains order. At the other extreme, some projects simply hire excellent teachers and provide them with the resources they need to teach as they choose. In between fall a complete range of projects including programed tutoring projects in which paraprofessionals follow explicitly prescribed tutoring sequences, and teacher-centered projects in which average teachers receive frequent training and monitoring in project-specified techniques.

A final difference of critical importance to a potential adopter site concerns the interaction between management and instructional approaches. For example, it is obvious that a laissez faire, teacher-centered management style cannot work with paraprofessional staff; it is equally true that a project based on tight control of highly skilled specialists may fail to work. In short, when attempting to analyze an effective project for the purpose of diffusion, it must be considered as a unique entity, and the complex interrelationships among many project elements must be described in terms of their contribution to the achievement of the ultimate project goals.

SUMMARY

1. The impact of "projects" is small in relation to our aspirations for these projects.

2. The impact of "projects" is small in relation to the many sources of confounding variance.

3. The best compensatory projects are much better than average projects, and far superior to no treatment at all. However in order to measure their superiority we need better criteria and measures of project success.

4. The diffusion of apparent achievement gains would require careful analysis of project measurement techniques and context, as well as analysis of the management and instructional features. We can expect disappointing results from such an effort.

5. Multivariate techniques have provided information which the project packager should know, but quite different analyses are needed to identify the features that contribute to the success of an individual project.

REFERENCES

Anderson, J. I., & Estes, G. D. Estimating the effect of regression to the mean. Paper presented at the Rocky Mountain Educational Research Association Meeting, Flagstaff, Arizona, 1976.

Bianchini, J. C. Achievement tests and differentiated norms. Paper presented at the U.S. Office of Education Conference on Achievement Testing of Disadvantaged and Minority Students for Educational Program Evaluation, Washington, D.C., May 1976.

Horst, D. P., Piestrup, A. M., Foat, C. M., & Binkley, J. L. Evaluation of the field test of Project Information Packages: Vol. II--Recommendations for revisions. Los Altos, CA: RMC Research Corporation, December 1975. (a) (RMC Report No. UR-293)

Horst, D. P., Tallmadge, G. K., & Wood, C. T. A practical guide to measuring project impact on student achievement. Washington, D.C.: U.S. Government Printing Office, 1975. (b) (017-080-01460) Also published as Measuring achievment gains in educational projects. Los Altos, CA: RMC Research Corporation, October 1974. (RMC Report No. UR-243)

Roberts, A. O. H. Personal communication, February 1977.

Stearns, M. S. Evaluation of the field test of Project Information Packages: Vol. I--Summary Report. Menlo Park, CA: Stanford Research Institute, in press.

Tallmadge, G. K. The development of Project Information Packages for effective approaches in compensatory education. Los Altos, CA: RMC Research Corporation, October 1974. (RMC Report No. UR-254)

Tallmadge, G. K., & Horst, D. P. A procedural guide for validating achievement gains in educational projects. Washington, D.C.: U.S. Government Printing Office, 1976. (017-080-01516) Also published in Los Altos, CA: RMC Research Corporation, Second revision, December 1975. (RMC Report No. UR-240)