DOCUMENT RESUME

ED 209 351 TM 810 910

AUTHOR Kennedy, Mary
TITLE Longitudinal Information Systems in Early Childhood Title I Programs.
INSTITUTION Department of Education, Washington, D.C.
PUB DATE Nov 80
NOTE 76p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Data Collection; Disadvantaged Youth; *Early Childhood Education; Evaluation Methods; *Information Systems; *Longitudinal Studies; Program Effectiveness; *Program Evaluation; *Student Characteristics
IDENTIFIERS *Elementary Secondary Education Act Title I; *Student Information System

ABSTRACT

The topics discussed in this resource book include not only how to assess the long-term impact of early childhood Title I programs, but also how to develop and use a longitudinal information system for a variety of routine operations. The book is designed to help evaluators decide what a particular information system should look like. For longitudinal data to be useful, they must be collected and maintained systematically. That means: (1) data collection is done at regular intervals; (2) the population on whom data are collected remains the same; and (3) the data are filed in such a way that they can be easily obtained when they are needed. This resource book contains four chapters. In the first, two ways in which longitudinal information systems can be used are described (assessing long-term program impact, and helping with routine operations). The two main types of such systems are discussed in the second chapter, together with their strengths and weaknesses. Chapter 3 describes how to determine information needs, and Chapter 4 discusses different design features of systems that must be anticipated before building them. (Author/GK)

LONGITUDINAL INFORMATION SYSTEMS IN EARLY CHILDHOOD TITLE I PROGRAMS

U.S. DEPARTMENT OF EDUCATION

LONGITUDINAL INFORMATION SYSTEMS

IN

EARLY CHILDHOOD TITLE I PROGRAMS

November 1980

Mary Kennedy
The Huron Institute
123 Mount Auburn Street
Cambridge, Massachusetts 02138

TABLE OF CONTENTS

FOREWORD

This booklet has been prepared as part of a United States Education Department (ED) sponsored project on evaluation in early childhood Title I (ECT-I) programs. It is one of a series of resource books developed to meet the need expressed by state and local personnel for information to help them evaluate and improve their early childhood Title I programs. The series describes the potential usefulness of various options in making local decisions about program practice. It focuses on the following questions:

- Who will use the evaluation results?
- What kinds of decision will the evaluation results be used for?
- Does the importance of such use justify the cost of evaluation?

Together, the resource books address a range of issues relevant to the evaluation of early childhood programs for educationally disadvantaged children. Separately, each book focuses on a single set of problems. The series include the following:

- Evaluating Title I Early Childhood Programs: An Overview
- Assessment in Early Childhood Education
- Short-Term Impact Evaluations of Early Childhood Title I Programs
- An Introduction to the Value-Added Model
- Evaluation for Improving Early Childhood Title I Programs
- Longitudinal Evaluation Systems for Early Childhood Title I Programs
- Evaluating Title I Education Programs That Involve Parents

The resource books are not comprehensive technical manuals. Their purpose is to help local school personnel clarify their information needs and decide what evaluation strategies are most appropriate to meet those needs. Additional information and technical assistance in using the various evaluation strategies are available in the more technical publications cited at the end of each volume and from the Technical Assistance Centers (TACs) in the ten national regions.

These resource books and the evaluation approaches they describe do not constitute a single evaluation system. Instead they represent a variety of approaches to evaluation aimed at serving a range of information needs at the federal, state, and local levels. This strategy has been adopted because our study of the feasibility of developing evaluation systems for ECT-I programs revealed that the variety of potential uses of information at these three levels is so great as to defy any simple, concise or unified "answer." No single evaluation system, no matter how complex, could possibly satisfy all possible information needs (Bryk, Apling, & Mathews, 1978, p. 22-27). Nevertheless, we hope that the variety of ECT-I evaluation resource booklets which we have produced will help people to connect different evaluation strategies to the varieties of information needed to improve and strengthen ECT-I programs.

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

This resource book is one of a series designed to help directors of early childhood Title I projects evaluate and improve their programs. Taken together, the series addresses a variety of evaluation problems that are unique to early childhood programs for disadvantaged children. This particular book, however, may be useful for other purposes as well. The topics we discuss in this resource book include not only how to assess the long-term impact of your programs, but also how to develop and use a longitudinal information system for a variety of routine operations. This dual purpose came about gradually, as the plan for the book evolved. While our original goal was to describe methods for assessing the long-term impact, or "sustained gains," of early childhood projects, we soon realized that a major part of such long-term evaluations involved developing and maintaining a longitudinal data base. Soon after that, we realized that such a data base could be useful for a variety of different purposes, both within and outside of early childhood Title I programs. Since the development and maintenance of such a system depends on your particular needs and your particular situation, we have designed the book to help you decide what your particular information system should look like. But once you have decided that, you may want to consult a management expert, a consultant from your regional Title I Technical Assistance Center (TAC), or other, more technical books, to help you get the system going.

What exactly is a longitudinal information system? The term "longitudinal" has two meanings. First, it can refer to the length of time records are kept on individual children; that is, you may maintain all

data on a given child from the time he or she enters school until he or she leaves sixth grade. Such a recording system would be called "longitudinal." The second meaning relates to historical information. You may, for example, maintain data on the characteristics of all kindergarten classes over a ten-year period. These data are also "longitudinal." The system you choose to develop could reflect either (or both) of these definitions. But the more important term is the word "system." For longitudinal data to be useful, they must be collected and maintained systematically. That means:

- Data collection is done at regular intervals
- The population on whom data are collected remains the same (i.e., "all kindergarteners," "all grade levels of a particular group," etc.)
- The data are filed in such a way that they can be easily obtained when they are needed.

Such a systematic longitudinal data base may not be possible in all situations. For example, if your population is highly transient, so that you wind up serving different children every year, you may feel that the effort involved in creating such a system will not pay off. Or, if you have a very tight budget, and may have to lay off staff in the near future, you may not be able to follow through on a system long enough for it to pay off. If you feel these kinds of problems will not deter you, read on.

This resource book contains four chapters. In the first chapter, we elaborate on the two ways in which longitudinal information systems can be used: for assessing long-term program impact, and for helping with routine operations. In Chapter II, the two main types of such systems are discussed, together with their strengths and weaknesses. Chapter III describes how to determine your information needs, and Chapter IV

discusses the different design features of information systems that must be anticipated before one can be built. You may want to combine this book with information from other sources, such as the Department of Education's Title I Policy Manual, or materials prepared by your regional Technical Assistance Center.

ASSESSING LONG-TERM PROGRAM IMPACT

Many educators believe that if children's long-term academic or social progress is to be truly influenced, early assistance is necessary. If this is the reason for establishing programs for young children, then it is as important to assess the long-term effects of the programs as the short-term. Figure 1 summarizes the evaluation questions that might be asked to assess these kinds of long-term impact.

A Foundation for Future Learning

Early childhood Title I programs (ECT-I) may enable children to acquire certain attitudes or skills that are needed in order to attain other skills. If so, the real effects of the program will be seen in the attainment of the later skills. For example, an early program may be designed to foster positive attitudes toward education, so that children will be more motivated throughout their school years to study and learn. While early evaluation may show how the program affects children's attitudes, only later evaluation can determine whether those attitudes lead to academic improvement. Bilingual programs also illustrate the need to assess achievements that cannot be measured until some time after the program ends. The eventual goal of many bilingual programs may be to help children speak, read, and write

Long-Term Goals	Sample Questions
Providing a Foundation for Future Learning	• Did the program have an immediate impact on attitudes? Did children whose attitudes improved make later achievement gains? • Did the program have an immediate impact on bilingual children's first language skills? Did children whose first-language skills improved later improve in English skills?
Prevention of Future Problems	• How many ECT-I children returned to Title I later on? • How many ECT-I children were retained in grade later on? • How many ECT-I children were referred to special education later on?
Sleeper Effects	• If the program demonstrated no noticeable effects at its close, could children's scores still be affected in later grades?
Sustained Gains	• To what extent are effects apparent in the spring still visible the following fall?* • Are effects still apparent at the close of the following school year? Through several school years? • Did children who had two years of the program sustain their gains longer than children who had one year?

* The 1978 Title I Amendments require that evaluations be carried out "over at least a twelve-month period in order to determine whether regular school year programs have sustained effects over the summer" (Sec. 124(g)(2)). These studies are to be carried out at least once every three years.

Figure 1: Four Assessments of Long-term Impact of Early Childhood Programs

easily in both their first and second languages; but if facility with the first language precedes instruction in the second language, a short-term impact study may only demonstrate your success in the first language. Again, to evaluate the full impact of bilingual ECT-I programs, later assessment is necessary.

Prevention of Future Problems

A frequent goal of early childhood programs is to prevent later problems in school--such as truancy, grade retention, or even droppong out of high school--or the need for later services, such as special education services or more Title I services. ECT-I programs with the goal of preventing later problems may not differ in content (such as attitudes toward learning or facility with another language) from later programs, but are designed to give children an early start so that they can keep up with other children. Assessing the extent to which later performance is free of problems is clearly far beyond the time frame of early childhood programs, and therefore requires longitudinal evaluation.

Sleeper Effects

Sometimes early childhood programs have sleeper effects; that is, children may appear not to have benefited from the program, but the program may in fact facilitate their academic success several years hence. An evaluation that stops at the end of the early program may thus lead to the erroneous conclusion that the program was ineffective. Why and when sleeper effects occur is conjectural, but clearly, if they exist, longitudinal evaluation is needed to find them.

Sustained Gains

While some early childhood programs show strong effects at the end of the program, these may fade out over time. Eventually, the children may appear to be no better off for the program, and may even fall behind again as if they had never taken part in it. Finding out whether the apparent early benefits fade over time is as important as finding out whether sleeper effects occur. In fact, the Congress recognized the importance of investigating whether gains were sustained when they revised the law in 1978. The regulations now require that, at least once every three years, each project measure children's progress over a twelve-month period. If the early benefits fade, changes need to be made. Longitudinal evaluation may help determine whether early effects last long enough to justify their costs.

ROUTINE OPERATIONS

The above discussion points out the value of longitudinal information systems for assessing the long-term impact of the program. We now turn to the ways in which good information can contribute to a wide range of decisions routinely made in schools.

Longitudinal information, by definition, takes time, but it may still be useful for this year's decisions, and for future decisions, including some that cannot be anticipated now. Because the potential uses are difficult to define in advance, and because many of them may be idiosyncratic to particular school systems, we confine our discussion to four major kinds of use, rather than specifying particular uses. Figure 2 summarizes these categories and gives examples of the kinds of questions that might be asked in each. Let us consider now how longitudinal information might meet some of these needs.

Routine Uses of Longitudinal Information	Sample Questions
Programming Decisions	• Which beginning readers prepare children best for their later reading texts? • Do in-class services or pull-out programs have more "staying power"? • How does early ability grouping affect children later on?
Policy Development	• At what age should early education programs be offered? • What eligibility criteria have the best long-term predictive power? • Should Title I programs be offered at all ages or are they more effective at one age than another?
Establishing Program Accountability	• Do children who leave Title I re-enter the program later on? Do they enter enter special education? • What is the academic progress of Title I children after they leave the program? • How often are former Title I students retained in grade? How often do they skip grades? • Do former Title I students graduate from high school? • Are former Title I students involved in extra-curricular activities?
Decisions About Individual Children	• Should this child be referred to special education? • Should this child be taught by a different method? • Should we involve the child's parents more than we have?

Figure 2: Routine Uses of Longitudinal Information

Programming Decisions

This category refers to decisions about the content or organization of instruction, which for the most part affect what happens inside the classrooms. They include decisions about curricula, teaching methods, and grouping strategies.

If early childhood programs are to be improved, you must know what they are like currently. Information about the nature and intensity of Title I services provided each year, for example, offers a starting point for analysis. Since class sizes and Title I budgets may change from year to year, the intensity of services offered to children may also change. Even within one year, services vary from classroom to classroom or building to building. Accurate data on what is available can contribute to discussions about more equitable or more efficient distribution of resources and to decisions about changing the general structure of the program.

One school district, for example, discovered that because staffing patterns had remained relatively stable for several years while the population of needy families had relocated, the case loads for social workers were quite uneven. Whereas one school had three social workers for 100 children, another had only one. These inequities appeared only when an information system was developed. Similar discrepancies may be found in the availability of teacher aides, particularly of unpaid parent volunteers.

Comparisons across buildings may help pinpoint other imbalances--say variations in parent involvement. While such information does not suggest a remedy, it at least identifies an area for program improvement efforts. Further investigations may suggest ways to increase parent involvement: changing the times when the Parent Advisory Council normally meets, to avoid conflict with a popular neighborhood activity; ensuring that

form-letter invitations in a non-English-speaking neighborhood are not sent in English, and the like.

When information on services is combined with information on children's academic progress, comparisons across buildings may contribute to understanding the relative benefits of different service delivery systems. Such comparisons must be made with care, however, since a variety of other events also contribute to children's achievement.

Information on services that is tied to financial or staffing data can give a fairly complete description of the program, and thus suggest areas where costs could be reduced without influencing the services provided, or where resources could be more fairly distributed. Of course, if children in different classrooms or buildings have different needs for services, this must be taken into account.

Finally, full descriptions of the services provided at each grade will show the extent to which successive educational experiences build on earlier ones, rather than repeating lessons, skipping lessons, or haphazardly changing formats. The data can indicate, for example, that some children received the same instruction in first grade as in kindergarten, simply because the first-grade teacher did not know what was offered in the Title I kindergarten program. Or they may suggest that some children have moved from highly structured programs to open classrooms (or the reverse), simply because of anomalies in the placement process. Knowledge of continuity between Title I and other programs may be especially useful in designing Title I programs to prepare children for their later educational experiences.

There is a variety of ways, then, in which full and accurate information can assist in program planning. And longitudinal information adds the dimension of time to the possibilities for analysis. Variations in services offered over the years may, for example, reveal gradual shifts in program emphasis or in parent participation that otherwise might not be noticed. Good ideas put into practice by someone who has since left the staff might be forgotten if not recorded. Longitudinal data provide an "institutional memory" that extends beyond the memories of staff members who come and go.

Policy Development

Whereas we have used the term programming decisions for decisions about how best to offer services to children, we use the term policy development to refer to all other decisions about who should receive the services, what ages should they be served, what should be the eligibility criteria, and so on. Longitudinal information would not only provide useful descriptions of the student body, but might also demonstrate the relative success or lack of success of prior policies. Indeed, an institutional memory may be even more useful in the policy sphere than in program planning. Policies must be continually revised to meet changing social values, changing bureaucratic or budgetary constraints, and changing community needs. At the same time, the ways in which they affect children must be understood if policies are to be improved. Without an institutional memory, a policy once adopted and rejected as inappropriate may be adopted and rejected again, for the same reasons. Suppose, for example, that a certain test is used to determine children's eligibility for services, because it seems to be especially appropriate for young children and is easy to

administer; but that it is later discovered that the test missed several needy children. Four years later, a new program director might try the same test again if he or she has no access to information regarding its earlier use.

Institutional memory apart, information on the later academic success of children who took part in early childhood Title I programs is of great importance for judging program success. Consider, for example, the staff in one school district who thought their early childhood program quite successful until they learned that their third graders were having serious reading problems. Because the district had a relatively transient population, and because there were no longitudinal records on the children who had been in the early childhood program, the staff did not know which third graders had participated in the program and which had not. Longitudinal records of children's progress, then, could guide policy makers who must decide such matters as whether funds would be better spent in early childhood or in later programs.

Since differences among groups of children could stem from many causes other than early childhood programs, information about these differences can be useful for policy development. A technique often used by insurance companies is that of developing actuarial tables. No causal inferences are overtly drawn from the tables, but they are nevertheless used to establish rate policies. In the case of automobile insurance, for example, accident rates are maintained separately for males and females, persons under 25 and over 25, and so on. In a similar vein, actuarial tables may be useful to educators. For example, if truancy rates vary by sex or neighborhood, that fact may help in interpreting the apparent success or lack of success of programs. Similarly, tables of average reading achievement

at each grade, for children with different language or family-income backgrounds, may suggest what policy makers can reasonably expect as "baseline" performance. These tables function essentially like local norm groups, but as many different categories of students can be developed as are needed to fit your local population. Longitudinal information can be especially useful for infrequent events such as grade skips or retentions. The system allows data to accumulate over a long time, until enough different children have skipped grades or been retained in grades that patterns can be seen.

Program Accountability

We use the term accountability to refer to situations in which data may not lead to particular decisions, but are nonetheless useful in justifying the program to parents, the community at large, or the funding agency. Both immediate and long-term program effects are an important part of accountability. Longitudinal information suggesting that program participants maintain their academic progress relative to their non-disadvantaged peers provides strong evidence of program success. And an information system can contain much more than test scores: it can also show how often children enter and leave Title I programs throughout their academic career, how many of them are retained in grade or referred to special education programs at some point, and how many have histories of truancy or of exceptional academic progress.

When the program goal is to prevent the need for later services, these data are valuable evidence of program effects.* If the evidence suggests that success has not been significant, the data may be used to

* Provided that the criteria for receiving those services have not changed over time.

revise policies regarding when and to whom services should be provided. When goals are framed in terms of achievement, longitudinal information can indicate variations in test scores across grade levels, across subjects, or across children who have entered school at different ages.

The data can be summarized in various ways, depending on the audience. Parents, for example, may be more interested in truancy and academic progress, while funding agencies may want to know how frequently children who have been in the early childhood program return to Title I later on. Annual progress may be summarized for children with and without the early childhood program, children with one year or two years of the program, Spanish-speaking versus English-speaking children, and the like.

Decisions about Individual Children

Teachers, guidance counselors, and other staff members make daily decisions about individual children: to try different instructional strategies or seating arrangements, to retain children in grade, or to refer them to special education. Often such decisions are based primarily on the staff member's own immediate knowledge of the child, with only sketchy information on the child's previous educational experiences. Knowledge of which programs have been tried, whether the child has exhibited academic, behavioral, or truancy problems in the past, or what the academic success of his or her siblings has been may assist in these decisions.

Decisions about individual children are not confined to teachers. One school district recently set up descriptive files of all children receiving special education and discovered that several of them had IQ scores in the 90s or even over 100! Since the scores were from tests given several years earlier, the placement teams had not known about them

when they placed the children. The data allowed the program director to identify children who needed to be re-evaluated.

PROBLEMS IN INTERPRETING LONGITUDINAL INFORMATION

The value of early childhood Title I educational experiences lies in their connection to later experiences, so much so that it would be difficult to separate the effects of one from the other. How, then, can one estimate the contribution of the early experiences? Each child's education consists of continuous experiences, and his or her success at any given time reflects the influence of the whole series. This fact makes it difficult to interpret children's test scores.

There are other sources of difficulty in interpretation. One of these relates to the tests themselves. While the desire for longitudinal information may stem from a theory of child development, or a recognition of the relationship between early and later experiences, available tests may not have been developed on that basis. Their content may change from grade to grade to accommodate naturally occurring curriculum changes across grades, but their score scales are rarely developed by actually testing the same group of children over time. Instead, the first-grade test is standardized on first-grade children, the second-grade test is simultaneously standardized on second graders, and so on. If, for some reason, fewer Title I children were included in the norming samples in one grade than in others, the scores of your Title I children relative to the "norm" may appear to change suddenly. Such changes would be due not to the actual educational experiences of your children, but to change in the group to whom they are being compared. Tests must therefore be selected not only to match the curriculum over the several years of the study, but also with an eye toward how changes in scores will be interpreted.

Interpretation of test scores may also be confusing because events have influenced children's attainment which were not part of the planned educational experiences, or because of historical changes in the programs themselves which are forgotten and therefore do not enter into the interpretation of children's later attainment. Longitudinal information is often interpreted under the assumption that the sequence of programs experienced by each wave of children is essentially the same. In fact, class sizes change with rises and declines in school enrollment; text books change as new editions come out; teaching styles change as teachers mature or change jobs; and the children themselves change over time, as their mothers do or do not work outside the home, and as other influences, such as television programming, change. Unless your longitudinal information system somehow takes historical events into account, the data may be erroneously interpreted.

Finally, since children may enter and leave the Title I program in the course of their education, and since many will leave the school system altogether while others transfer into your schools from other districts, information neglecting these changes may be difficult to interpret.

SUMMARY

This chapter has described in broad terms the two main uses of longitudinal information systems: assessing the long-term impact of early childhood programs, and contributing to routine decisions. In each case, the variety of suggested uses implies that such information systems would be valuable. The applications of longitudinal information are summarized in Figures 1 and 2. We have also seen that for several reasons the information may be difficult to interpret. The remainder of this resource book

will be devoted to more pragmatic concerns of developing and using an information system.

II. TWO KINDS OF LONGITUDINAL INFORMATION SYSTEM

There are two general strategies for developing longitudinal information systems. Each has its own strengths and weaknesses, and the two differ in their appropriateness to different situations. For convenience, we label these two strategies controlled studies and descriptive studies, to reflect one of the main differences in the two approaches: control. Whereas the first strategy entails careful control over which children receive which educational experiences, the second allows children to be assigned to programs in routine ways. There are other differences, too. "Control" means that children are assigned to different programming arrangements systematically, so that specific questions such as "what would happen if. . ." can be answered. Description, on the other hand, does not entail such controls, and so cannot provide specific answers; but it may provide useful insights into a number of different questions. And this leads to a third difference in the two strategies: the descriptive system, in its relatively broad scope and flexibility, is more appropriate for exploration, whereas the controlled study of groups permits strong inferences to be drawn from the data analyses. That is, the latter frees you from nagging questions about alternative explanations for observed differences among groups. With these differences clearly in mind, let us consider each approach in more detail.

CONTROLLED STUDIES

Controlled studies allow direct comparisons across groups of children who have had different educational experiences. For example, one group may follow a DISTAR program while another uses, say, Sullivan readers. Or one group may take part in preschool programs at ages three and four, another

only at age four, and a third not until kindergarten. The comparisons of interest must be specified ahead of time so that services and children can be arranged to allow comparison.

Figure 3 displays several aspects of an ideal controlled study of a one-year program. First, a population is identified--children eligible for an early childhood Title I program, for example. Next, a sample is randomly selected from this population and randomly assigned to a program group or a comparison group. These groups may participate in two different kinds of program; or if funds are scarce, one group may take part in the program while the comparison group does not. Each group can be tested before the program begins (time 1), but need not be. At the end of the program (time 2), all groups are tested, and testing continues for all groups over a period of time (time 3, 4, . . .)* as children continue through school. Finally, those children who leave the program early for any reason are followed over time and tested at the same specified intervals. This is crucial to later interpretation of the data. If the dropouts are random, they are not likely to cause any differences later between program and comparison groups. If they are somehow different from those who stay (for example, tend to come from poorer families), then comparisons of the remaining groups may be difficult to interpret.

If the program and comparison group members are compared at times 1, 2, 3, . . . , the immediate and long-term program effects can be

* It is highly desirable to have a series of measures for which results can be expressed in the same way, such as in percentile ranks. That is, the same or comparable tests or measurement techniques should be used throughout the study. These measures should be able to show growth across the years, and if possible, relative standing at any given time.

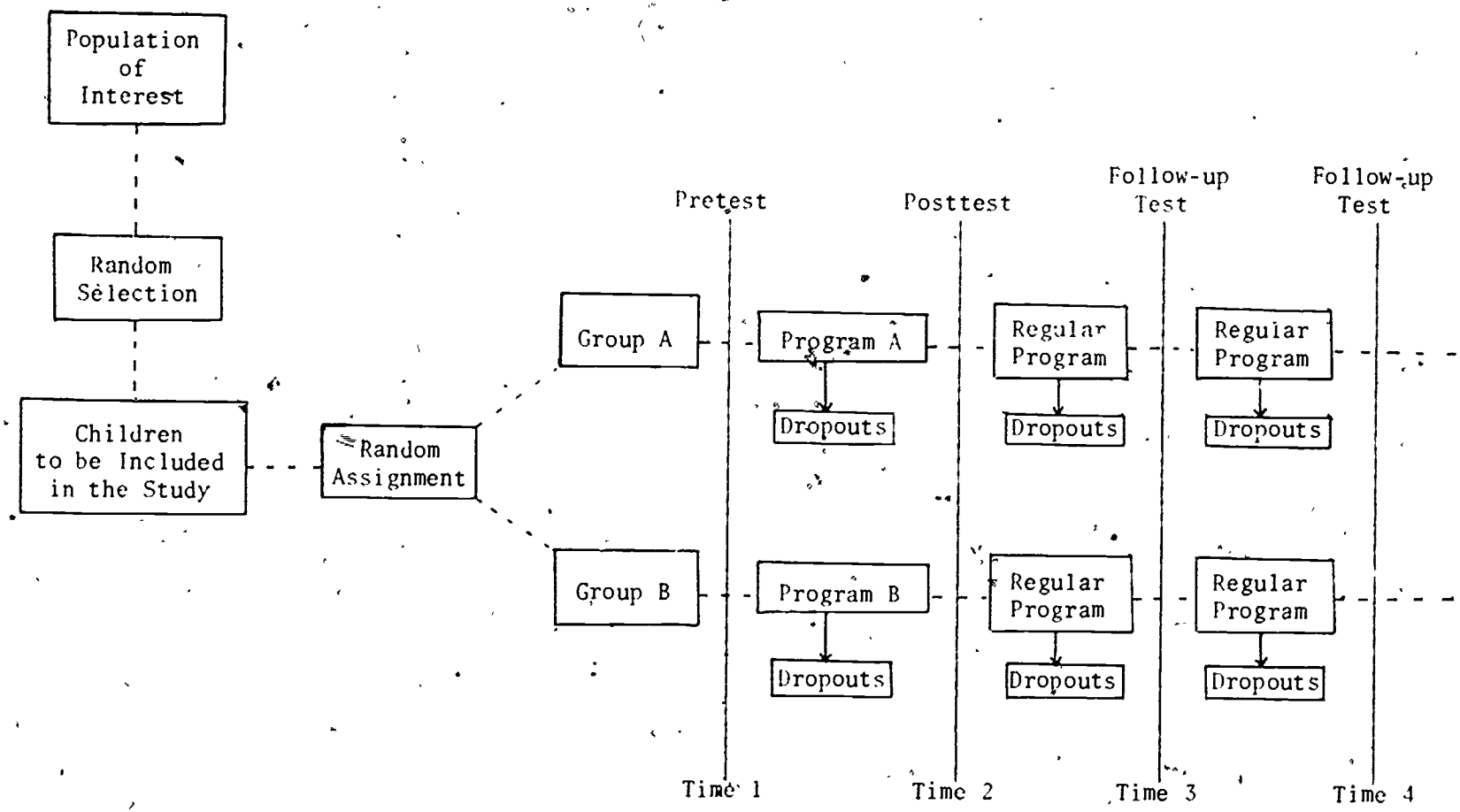

Figure 3: An Ideal Controlled Study of a One-Year Program

examined. In addition, comparing immediate and long-term effects will demonstrate whether short-term gains are sustained and whether long-term sleeper effects have occurred. The evidence is strong because of the control over the study; the program experience should be the only difference between final program group and comparison group. Thus, any gains of the program group vis-a-vis the comparison group must be due to the program.

An Example

Louise Miller and Jean Dyer (1975) compared four different Head Start programs as they affected disadvantaged children. Four-year-olds were randomly assigned to the different programs and attended them for one year. Miller and Dyer added a group of comparison children who had no Head Start experience. Head Start children were given a battery of tests when they entered the preschool programs, and they and the comparison children were regularly tested from entry to kindergarten through the end of the second grade.

In addition to randomly assigning children to programs, these evaluators took care to equate their groups on a number of other conditions, such as the kind of grade-school programs children went into after Head Start, and the general quality of the buildings in the schools the children attended. They also videotaped the classes, not only in Head Start but in kindergarten and first grade as well, to be sure that the programs were carried out as they had been planned. Clearly, this example of a controlled group comparison is one of considerable complexity, thoroughness, and cost. Just as clearly, the care paid off in complete and interpretable findings.

If controlled comparisons of groups can be done, they can provide strong and convincing evidence of both short- and long-term program effects. No doubt, however, you are already listing reasons why this strategy is not practical. In addition to the problem of costs, finding and maintaining a comparison group may be difficult if not impossible. Randomly assigning children eligible for Title I to program and non-program groups may not be legal, since the programs are mandated to serve the neediest children. Some program directors may be genuinely interested in the relative benefits of two alternative programs, however, and for them controlled comparisons may be possible, since each program can serve as a comparison for the other. But for most, the question of real interest is "early program versus no early program"--a comparison that is almost impossible to make by randomly assigning children to groups. Moreover, locating any non-program comparison group, especially for prekindergarten programs, can be difficult, because parents will not be eager to allow testing and observation of their children without receiving program benefits. Although this is not as severe a problem with kindergarten or first-grade programs, for which there is an available pool of children not in Title I programs, random assignment of those eligible for Title I may still not be possible. And children who were not eligible for Title I will provide no clear evidence of how those who were would have done over time in the absence of the Title I program. Finally, if the comparison group is in school, it cannot be said that it follows no program, for many of its members may be in other compensatory programs, some quite similar to Title I. There is a need, then, to define what kind of educational program can be considered a "no-program" experience that would be meaningful to compare to the Title I experience.

Attrition may also make controlled studies less feasible. If the attrition rate differs for the program and comparison groups, and if the difference cannot be attributed to chance, then observed differences in outcomes may be due to differences in the children rather than the programs. If loss is random--that is, if the final sample is not significantly different from the dropouts--attrition is only a technical problem.

DESCRIPTIVE STUDIES

In many cases, a simpler descriptive approach more readily fits the realities of early childhood Title I programs. Since ECT-I children are selected because of educational need, random assignment, as we have said, is generally impractical, if not illegal. Descriptive systems do not require random selection or assignment of children to programs; they just follow up whoever was selected. In addition, using the same or comparable measures over time may not be possible, given the range of program goals for different age groups. Descriptive systems do not require the same or comparable measures. Finally, the descriptive system can portray the actual educational programs that children take part in over time, and so provide a complete knowledge of both program experiences and academic progress.

Because systems can be designed in different ways to meet different information needs, there is no "ideal" descriptive system. Figure 4 indicates the groups on whom data could be collected if the system spanned several school years (grade k through 6). Even more groups could be defined, since each wave of children may contain several different subgroups: English- or Spanish-speaking; children with or without Head Start experience; children with or without early childhood Title I programs;

Wave	Year						
	1980	1981	1982	1983	1984	1985	1986
1	k	1	2	3	4	5	6
2		k	1	2	3	4	5
3			k	1	2	3	4
4				k	1	2	3
5					k	1	2
6						k	1
-							k

Figure 4: Different Groups of Children in a Longitudinal Descriptive Information System

children who did or did not drop out of those programs, and so forth. The data maintained could also include achievement test scores, teacher ratings, grades, attendance, assignments to special education, grade retentions, or re-entries to other Title I programs.

An Example

Lazar et al. (1977) analyzed the descriptive information available on a project in Vineland, New Jersey. Their analysis shows how a comparison group can easily be added to a longitudinal descriptive information system, and illustrates the usefulness of program outcome measures such as assignment to special education and grade retention. The participants in this program, much like ECT-I children, were the highest-risk children; thus no similar group could have been found for comparison. The researchers therefore decided to compare the children in the program to the general school population, represented by a random sample from the previous year's first-grade classes. This sample, of course, included middle-class as well as lower-class children. To compensate for the higher percentage of Spanish-surnamed children in the program group than in the general school population, 36 additional Spanish-surnamed children were selected randomly from the previous first-grade group. This, the researchers claimed, compensated for ethnic and social class differences.

Two comparisons were then made. First, the program children were compared to the random sample of the general school population, excluding the 36 Spanish-surnamed children. Then, the Spanish-surnamed program children were compared to the sample of Spanish-surnamed children in the general school population, including both those in the first sample and the 36 additional children (p. 70).

Tables 1 and 2 illustrate group differences in grade retention and in assignment to special education. Since statistical tests assume random assignment, they are not appropriate for these comparisons; but frequencies and percentages alone can be useful. For example, Spanish-surnamed children are retained in grades at a higher rate than other children; this seems to warrant further investigation. And program children are more likely to be assigned to special classes. This could mean that they need further help, or that there is bias in the special-education assignment system.

The descriptive longitudinal information system clearly differs from the controlled strategy described earlier, and therefore cannot be expected to provide the same amount and kinds of information. Most significantly, the descriptive study cannot be used to infer that observed differences in children's accomplishments were caused by the program, since too many other contributions are left uncontrolled. But it can, for example, describe the conditions of ECT-I graduates. Do these children remain in Title I programs? Do they leave Title I only to be cycled back in later? Do they end up in other compensatory programs? Or do they enter and continue through regular education programs? Descriptive information can also tell you how many of your second- or third-graders participated in the early childhood program, how many of them transferred into your school district, and even how often children's absences contribute to missing data.

But the descriptive study, like the controlled study, has its own difficulties and requires a substantial commitment of time and skills, as well as continuity of staff for maintenance. In addition, since its

Table 1

The Effects of the Micro-Social* Learning System on Grade Failure

Program Children vs. General School Population

	Not Retained	Retained	Total
Program N % of Program	132 (68.0%)	62 (32.0%)	194 (100%)
General Population N % of General Population	66 (65.3%)	35 (34.7%)	101 (100%)
Total	198	97	295

Spanish-Surnamed Program Children vs. Spanish-Surnamed School Population

	Not Retained	Retained	Total
Program N % of Program	80 (67.2%)	39 (32.8%)	119 (100%)
Population N % of Population	23 (37.1%)	39 (62.9%)	62 (100%)
Total	103	78	181

(Source: Lazar, 1977, p. 80)

* This was the name of the program developed for the children.

Table 2

The Effects of the Micro-Social*Learning System on Assignment to Special Education

Program Children vs. General School Population

	Not in Special Education	In Special Education	Total
Program N % of Program	169 (86.2%)	27 (13.8%)	196 (100%)
General Population N % of General Population	95 (94.1%)	6 (5.9%)	101 (100%)
Total	264	33	297

Spanish-Surnamed Program Children vs. Spanish-Surnamed School Population

	Not in Special Education	In Special Education	Total
Program N % of Program	104 (86.)%)	17 (14.0%)	121 (100%)
Population N % of Population	54 (87.1%)	8 (12.9%)	62 (100%)
Total	158	25	183

(Source: Lazar, 1977, p. 81)

*This was the name of the program developed for the children.

benefits derive from its ability to contribute to ongoing hypothesis testing and problem solving, it is only as good as the descriptive information it contains.

SUMMARY AND COMPARISON

We have outlined two different kinds of longitudinal information system, the controlled study and the descriptive study. Each has advantages and disadvantages. The controlled study requires great care in selecting the educational experiences that could meaningfully be compared with Title I experiences, and in manipulating children's placements so that the comparisons are valid while the Title I rules are honored. The descriptive study, on the other hand, requires maintenance of quite a bit more information, since you do not know in advance what kinds of comparison will be useful later on. It could eventually be more costly and more complex to manage.

Generally speaking, the controlled study is better suited to assessment of long-term benefits of early childhood programs, since it compares the achievement of children with and without the program; while the descriptive study is better suited for routine operations, since it can provide information not only on outcomes, but also on other aspects of children and on program characteristics as well. But this distinction is not absolute. In the controlled study, several kinds of information need to be retained, including data on the characteristics of children who remain in the program and of those who leave, so the information system developed may turn out to be useful for other purposes as well. And descriptive studies may enhance your understanding of long-term program benefits, provided conclusions are carefully drawn. The next chapter discusses in greater detail how you can determine your own information needs.

III. DEFINING YOUR INFORMATION NEEDS

Both types of longitudinal information system described in Chapter Two require a great deal of planning if they are to be useful for the purposes we have discussed. Because such systems can easily become too large to be manageable, one of the most important parts of planning is to decide exactly what information you will need. This chapter provides a number of exercises to help you do that. They are divided into three stages: getting started, expanding, and compromising. The two getting started exercises involve deciding who will use your information system and determining how they will use it. When you finish these exercises, you should have a set of "wish lists" of information different people would like to have. The next stage consists of expanding those lists by considering all sorts of other information that you might need to interpret the data on your original wish lists. By the time you finish this stage, your list of possible information should be quite long--too long, in fact. Then you need to start the third stage, which consists of compromising. To help you with your compromises, we provide a number of tables you can use to list and compare all your possible types of information, so that you can choose those that will be most feasible for you to collect.

GETTING STARTED

Since information can be used by many people and in many ways, the first step in developing your system is to determine who will be the primary users and what will be the primary uses.

Identifying Users

Different people in the district have different information needs and may require information at different times. Guidance counselors, for example, are more likely to want data on individual children than aggregates of data. Teachers may also want data on individual children, and occasionally on class averages. Principals may be more interested in data on the children in their buildings than on those in the whole district, while program directors and other district administrators may need aggregated data on all children.

It may not be possible to develop a single system that responds to all these needs. If data are stored longitudinally by child, aggregations across children may be more difficult; if they are stored by grade level, it will be difficult to reconstruct a file on an individual child.

The question of who will use the system influences more than just the way the information is organized. If more than one person has access to the files, safeguards for maintaining confidentiality must be built into the system, and each user must be trained in how the system is organized and how to find the information he or she needs.

The fact that different people may want different kinds of information need not mean that the system must exclude some possible users. Yet attempting to serve all potential users may not be the best solution, since a compromise may mean that no one can easily get the information needed. One reasonable alternative may be to maintain two sets of files--one with longitudinal data on individual children, and another with aggregates of data maintained for making comparisons or reporting general statistics to

the parents and the community. This is not such a monumental task as it might seem. The raw data (test booklets, teacher ratings, program enrollment data) are available anyway, and can easily be filed in one system, while the scores from these data can be summarized in another system.

Identifying Uses

Earlier, we described four broad categories of use (policy development, accountability, program decisions, and decisions about individual children) for descriptive information systems, and suggested various kinds of comparison for which the systems are suitable. Since different uses imply slight differences in system organization and information to be collected, it might be useful to rank-order the comparisons in terms of the likelihood that you would want to make them, or in terms of the relative value they would have for you.

One way to sort out uses is to review instances in which people wished they had had more information. The following situations, for example, may have led to that wish.

- A visit by a Title I program review team
- Preparation of the last Title I grant proposal
- A school board budget-making meeting
- Negotiations with the teachers' union
- A visit from the state education agency
- A conflict between the Title I director and the director of special education programs
- A local newspaper report on the school district
- A PAC debate
- A decision on staff or textbook changes.

It may be helpful to ask several people who were involved in these events to list information they wish they had had, or, if they can, to outline tables showing how they would like to have seen it displayed.

Once you have determined who will be the primary users of the information system, how they would use it, and what information they needed during recent critical events, you will already have developed a good-sized list of potentially useful information. But to make that information as useful as possible, you may need to expand it further. The next section discusses some ways to anticipate the need for ancillary information.

EXPANDING

Upon reviewing different "wish lists," you may discover that information is often wanted for comparisons such as those made when assessing long-term program impact. That is, people may want to know whether this eligibility criterion picks up more needy children than that one, whether Spanish-speaking children benefit from the program as much as English-speaking children, or what the additional benefits of a summer program are relative to the regular school-year program.

But unless these comparisons are based on controlled studies, they cannot lead to sound conclusions about <u>why</u> things look the way they do. For example, since many things can influence children's academic progress, differences between children who have participated in the early childhood program and those who have not cannot be attributed solely to the program. they may have been there earlier, or be due to post-program experiences.

For any comparison that can be made, there are several plausible explanations for observed differences. If you can anticipate comparisons you will want to make, you can design your information system accordingly--to control children's assignments, or to include descriptive information that may help you interpret uncontrolled comparisons. Let's consider several kinds of comparison that could be made, to see how planning can make them more useful.

Program vs. No Program

Suppose children who were in early childhood Title I programs are tested on reading comprehension at the end of third grade, and that the scores are lower than those of children who have never had Title I services. Can we infer that the lower scores are due to the program, that children are better off without it? Certainly not, for the program children differed in several ways from the non-program children:

- They were behind academically when they first started school
- Many of their families may be poorer, and may not be able to provide as many stimulating experiences to their children as other families can
- They may be more likely to have single parents and to be cared for during the day by their grandparents.

If you know that you will want to make such a comparison, you can do one or both of two things in advance to help you interpret it: you can control assignments to the program, so that the two groups are more similar; and you can collect and maintain information on the characteristics of the two groups, so that you will know exactly how the groups differ. Since Title I services

must be provided to the neediest children, you cannot randomly assign children to program and no-program groups. But you may be able to identify other groups of children who are similar: poor children who are less educationally disadvantaged, or higher-income children in other schools who are equally educationally disadvantaged. Neither group offers a perfect comparison, of course, so information about their characteristics, as well as about those of the Title I children, should be maintained. The important point here is to identify the comparison group in advance, and to collect information about it in advance.

Program vs. No Program: Title I Eligibility

As data are being collected, it may become apparent that some of the children were eligible for the program, but for various administrative reasons were never identified or served. If you are operating a preschool program, for example, your screening procedures may not have found all eligible children. Although this would normally occur only rarely, the existence of even a handful of such children may be enormously useful for comparison purposes, since their backgrounds are probably more like those of children in the program than are the backgrounds of children who are not. Such children may invite frequent comparisons and ready interpretation of observed differences, even though they are not really the same as those who took part in the program. One must ask why these children were overlooked initially. Are their families more out of touch with the school? More leery of educators? Do they live in different neighborhoods? Speak different languages? Though it may be hard

to determine how served and eligible unserved children differ, these differences must be sought out, for they could influence the educational process as well. This information could also become a part of the longitudinal information system, and may be useful not only for interpreting comparisons between children in and outside the program, but also for improving the process of identifying eligible children.

Programs in Different School Buildings

If the program is offered in more than one building, program results from one building to the next may be of interest. But children in different buildings may vary in many other ways. Each building may serve a different neighborhood, and each neighborhood may have its own cultural values and attitudes toward education.

Since children in all schools for which outcomes are to be compared are receiving Title I services, they can, in principle at least, be randomly assigned to schools. But random assignment may not be practical, since it would entail busing children away from their neighborhood schools. If no special assignment procedures are used, comparisons of program effects in different schools should include the same information on background characteristics as do comparisons of program vs. no-program.

Furthermore, if the comparison is of interest because of known differences in programs, then you should also keep information on the nature of the different programs. Was the program in one building changed so that it would better meet the needs of children in that neighborhood? How so? Or were new materials simply placed there to try them out? If so, why

in this building and not another? Is there a teacher there who advocated these materials? If so, the differences in outcomes may reflect the teacher's zeal more than the effect of the materials. Finally, the buildings themselves may differ--one has larger rooms, one is newer, one is designed for open classrooms while another has seats bolted to the floor, one has a library, and so forth.

These slight differences between neighborhoods, or between program characteristics from building to building, may be difficult to document and maintain in an information system, but they will prove to be enormously useful to those who wish to interpret comparisons of child outcomes across school buildings.

Waves of Children

If the program changes over time, or if it was only recently initiated, it may be useful to compare the academic progress of different waves of children, since they will have had different program experiences. But just as other comparisons can be interpreted in several ways, so can comparisons of waves. Unless the district has introduced a new desegregation program, one wouldn't expect the population to change too radically from year to year, but other events could occur that would affect these children:

- Local employment rates may rise or fall rapidly if a large company obtains or loses a major contract
- A new public television station may bring educational programs such as Sesame Street and The Electric Company into range
- A lengthy teacher strike may deprive different waves of different educational experiences
- An historical event such as a rocket launching or an election may alter curricula so that some waves are taught more current events and less standard curriculum than others.

Once again, these various influences may be difficult to anticipate or to document; but including such information in the system may be a great help in interpreting cross-wave comparisons.

Age or Grade Level

A major advantage of longitudinal information systems is that they allow children's academic progress to be tracked over time. Do the early program effects last? Do scores rise temporarily, only to fall later on? Of course, all children's scores go up over time, so progress is usually measured relative to that of some other group of children--children in your district, or children in the norming group on which the test scores were developed. Longitudinal questions, then, tend to be phrased as, "Were the differences between Title I children and the total school population larger at the end of first grade than at the end of fourth grade?" Interpretation of these comparisons requires just as much knowledge of the comparison groups of children as any other comparison does. If program children are being compared on first-grade versus fourth-grade standard scores in reading comprehension, then the characteristics of the norming sample should be compared with the characteristics of your program children. What is their family income? How many had only one parent? What wave were they in, and how might that wave differ from that of your program children?

In addition to differences between groups, differences in tests used at different times may also affect these comparisons. For example, the first- and fourth-grade reading tests may differ on any of the following:

- Emphasis on word recognition vs. sentence comprehension
- Test reliability

- Time needed to take the test
- Oral vs. written test-taking instruction
- Match between content of test items and curriculum
- Qualifications or experience of the test administrator.

Since all of these differences among tests could affect children's scores, they should be taken into account in interpreting cross-age or cross-grade comparisons. This will be easier if the peculiarities of each test and each test administrator are documented and placed in the information system as they are met, so that they are not forgotten later when the comparisons are made.

Attrition

If the community population is highly transient, you may need to know the nature and extent of attrition in your programs, as well as the characteristics of children who leave compared with those who stay. That information can be used for program planning, but is also especially helpful when other comparisons are interpreted. Suppose you are looking at fifth-grade achievement scores in schools with and without early childhood Title I programs. How many of the children in Title I schools actually had the early childhood program? In a highly transient community, there may be few who did, and that is worth knowing.

Summary

Since we cannot list all the ancillary information that you might need for your population and your evaluation questions, we have tried instead to provide examples of how such information may be used to interpret different comparisons. You may easily think of other kinds of information that you might need to interpret comparisons of particular interest to your particular Title I early childhood program. But you may also feel like you can't

possibly collect all this information. The next section suggests ways to select those pieces of information that will be of most help to you.

COMPROMISING

Once you have completed your starting and expanding exercises, you may find that your list of potentially useful information is quite long. Tables 3 and 4 list a variety of kinds of information that might be useful, but only as examples. Your exercises may lead you to several others. Once you have your ideal "wish list," you need to find ways to pare it down to a list of things on which you can really collect data.

The information people would like to have is often not easy to obtain. For example, one may wish to know how frequently children have trouble with school, whether that trouble is in their academic progress or in their attitudes toward education. The concept of "having trouble" is a broad one, but several indicators of trouble may be available: tardiness, truancy, referral to the principal's or counselor's office, grade retention, and referrals to special education, are all indicators of trouble, just as are test scores, grades, and behavior rating scales or checklists. Similarly, "success" may be measured by test scores, participation in extra-curricular activities, winning prizes, or being on the honor roll.

Whereas a review of recent situations in which information was needed may yield examples of ideal information, it is equally important to determine the best indicators of each ideal. This is true even of more readily agreed upon concepts such as achievement, which could be measured by criterion-referenced tests, teacher-made tests, grades, pages completed in workbooks, writing samples, and number of objectives mastered. Thus, two steps are needed: to define the ideal information,

Table 3

Kinds of Student Information That Could Be Catalogued In a Longitudinal Information System

Student Status
• Age
• Age at entry to school
• Address
• Parent income and education
• Race*
• Second language
• Health impairment (e.g., asthma, diabetes)
• Number of adults in the home
• Number of siblings; their ages, grades
• Title I eligibility, by year
• Presence in Title I programs, by year*
• Whether student dropped out of Title I mid-year
• Test scores or grades, by year*
• Teacher ratings, by year
• Attendance
• Grade retentions or skips
• Presence in special education programs, by year
• Extra-curricular activities, by year
• Honors or awards

Test Score Data
• Content
• Match between content and curriculum
• Test-taking skills required
• Norm group average income
• Norm group cohort
• Average family size of norm group

* Already required for Title I reports

Table 4

Kinds of Program Information That Could Be Catalogued In a Longitudinal Information System

Title I Services

- Classroom
- Curriculum, by year*
- Individual/small group/large group instruction
- Use of aides
- Number of days/year, by year
- Number of hours/week, by year
- Class size
- Mean class ability
- Parent participation in class*
- Parent attendance in PAC*
- Parent attendance at teacher-parent meetings
- Other special services available
- Ratio of service providers to children*
- Unusual curriculum changes that reflect current events (e.g., a rocket launch, an assassination)
- Unusual events that modified service delivery (e.g., a flood, a teacher strike)

Finances

- Cost of teachers
- Cost of workbooks
- Cost of consultants, itinerant teachers, etc.
- Cost of special equipment
- Cost of miscellaneous materials
- Cost of overhead
- Cost of administration

Staff

- Teacher age, years experience, education
- Aide age, years experience, education
- Itinerant teacher age, years experience, education

*Already required for Title I reports

and to define the indicators that may be used to provide insight into the concepts of interest. Let us review some methods for selecting among the kinds of information listed in Tables 3 and 4 that you can afford to collect and maintain.

Choosing Influences on Outcomes

Earlier, we pointed out that descriptive studies of groups yield less certain evidence of program effects than controlled studies, and that uncertainty makes it necessary to maintain data on various non-program factors that could influence the children's progress. Since many of these potential influences may be difficult to measure, it may be useful to construct a table summarizing them and their contributions to comparisons you will want to make. Figure 5 illustrates such a table. Since the possible influences on children's academic progress are infinite, some rules are needed for inclusion in the table. We have chosen four kinds of influence:

- Those that school districts already tend to control
- Those that school districts could control, but tend not to
- Those related to the quality of the indicators
- Those that school districts probably could not control.

Whether these influences are or are not controllable by the school district should not affect the decision to maintain data on them, since the purpose of such files is to aid in interpreting comparisons. If the children served by different schools come from significantly different family backgrounds and there is reason to believe that this influences academic progress, then the backgrounds are relevant to interpreting academic progress, whether or not the school can control them.

Your own knowledge of the children and the community you serve may lead you to hypothesize other influences on children's academic or social progress in school. You should list these as well. Once you have listed all

Influence	Strength of Influence on Outcomes	Availability of Data
Controlled by the School • Class Size • Curriculum • Eligibility Policies • Teacher Education • Presence of Teacher Aide • Teacher Race or Language		
Could be Controlled by the School • Family Attitudes Toward Education • Teacher Strikes • Distance From Home to School • Parent Participation in School Decisions • Parent Tutoring		
Related to the Indicators • Test Anxiety • Reliability of Tests • Match Between Test Content and Instruction • Non-substantive Requirements (Child Patience, Docility)		
Not Controllable by the School • Parent Income • Parent Education • First Language • Neighborhood Value Placed on Education • Presence of Older Siblings • Number of Adults in the Home • Pretest Scores		

Figure 5. Influences on Outcomes: Interpretation of Comparisons

the influences, you need to identify those that you consider the most important and that must therefore be recorded in your information system. A simple rating scale, where, say, a score of 1 means highly influential and a score of 10 means not very influential at all, may suffice for choosing what you should record. You may also want to score these influences on a second scale that roughly indicates how easily accessible the information would be.

Choosing the Indicators

No indicator will perfectly measure a concept of interest. Just as the Gross National Product imperfectly indicates the general economic health of the nation, so do reading scores imperfectly indicate the general academic progress of students. But for any given concept, many indicators are possible--more than can easily be maintained in a descriptive information system. Thus, you need some means of selecting the most appropriate indicators. Various criteria can be used to select the indicators most feasible and useful to maintain. Consider, for example, the following:

- Availability: Is the information already available and just needs to be pulled together, or must new data be collected?
- Relevance: Of the different possible indicators of a concept of interest, which most closely reflects the concept?
- Precision: How accurate is the indicator?
- Cost: What is involved in collecting data on the indicator? How much staff time? How much paperwork? If it is a purchased test or survey instrument, how much do the copies cost, and how much does scoring cost?
- Missing Data: How likely is it that you can obtain the information on all children or all families? What percentage of the data is likely to be missing?
- Face Validity: If audiences outside the school system--e.g., funding agencies or the community--will review your data, which indicators will seem most valid to them? Which will seem less relevant?

Figures 6 and 7 indicate methods for summarizing different criteria for indicators to show their relative advantages and disadvantages at a glance.

Indicator	Characteristics of the Indicator						
	Availability[1]	Relevance to General Concept	Precision	Estimated Cost of Data Collection		Likely Percent of Missing Data	Face Validity
				Personnel	Non-Personnel		
Frequency of Attendance at PAC							
Frequency of Attendance at Parent-Teacher Conferences							
Visits to School							
Volunteer Home Tutoring							
Volunteer Teacher Aide Work							
Active Participation at PAC							

[1] AA = Already available in someone's file; just needs to be pulled together
NA = Not available but could be easily recorded; person and mechanism are in place
AD = Available with difficulty; new procedures would have to be established to get the data

Figure 6: Indicators of Parent Involvement

Indicator	Characteristics of the Indicator						
	Availability[1]	Relevance to General Concept	Precision	Estimated Cost of Data Collection		Likely Percent of Missing Data	Face Validity
				Personnel	Non-Personnel		
Standard Test Score							
Criterion-Referenced Test Score							
Teacher-Assigned Grades							
Teacher-Assigned Grades, Writing							
Pages in Workbook							
Objectives Accomplished							
Tardiness							
Truancy							
Grade Retention							
Referral to Special Education							

[1] AA = Already available in someone's file; just needs to be pulled together
NA = Not available but could be easily recorded; person and mechanism are in place
AD = Available with difficulty; new procedures would have to be established to get the data

Figure 7: Indicators of Achievement

Of course, that glance does not necessarily make selection easier, but it does provide the information needed for selection.

STARTING SMALL

The exercises that we have described here should help identify the data that are most likely to be useful to you. However, they all rely on your guesses about ease of data collection, likelihood of use, and so on. And because they are based on conjecture, it would be wise not to initiate a large and complex longitudinal information system at once. Instead, begin with a small system, perhaps developing a file on one or two waves of children, or even one or two classrooms. After collecting data for a year or so, you will no doubt want to modify your system. For example, the first time you create summary tables, you may wish you had gathered other information. In fact, if you show your tables to those who asked for this information, they may well want still further information to interpret the tables. The data on your first small group should then be revised as needed and maintained for at least two years. This mini-longitudinal system will enable you to discover other categories of children and measures of progress that you had not anticipated, which will lead to still further modifications in system design.

SUMMARY

This chapter has outlined exercises designed to help you determine your information needs: obtaining "wish lists" of information from those whom the longitudinal information system is designed to assist; refining those lists by anticipating comparisons that might be made; and reducing the many kinds of potentially useful information to those that are affordable, practical, and most useful. Chapter 4 is devoted to characteristics of the information system itself that must also be determined in advance.

IV. CONSIDERATIONS IN DESIGNING A DESCRIPTIVE INFORMATION SYSTEM

Aside from difficult decisions about the information to be collected and maintained in a descriptive information system, the nature of the system itself must be decided. To some extent, system characteristics depend on the information chosen; but since the system will cost money, the information you choose may also depend somewhat on cost limitations. This chapter therefore reviews the potential limitations of different information systems.

An information system could be merely a file drawer of folders on individual children, or it could be a large computer with records stored on cards, tapes, or disks. Whether information is stored in file drawers or in computers, four characteristics define the system: size, organization, how it is gathered and entered, and how information is retrieved. These characteristics must be decided upon before the system is developed, and each depends partly on factors such as cost and feasibility, and partly on who will use the information and how.

SIZE

The size of the system means simply the number of pieces of information -- data points -- contained in the files. We saw earlier that a system describing children from kindergarten through sixth grade could contain records on 28 groups/grades of children. If each group contained 30 children, and one test score was maintained for each child at each grade, the system would contain 840 (30 x 28) test scores. Now suppose that each test has six subtests, and that the system records those scores as well: it would then contain 840 x 6, or 5,040, test scores!

The size of the system just described is defined by the number of child-

ren (30 per year), the number of waves of children (seven), the frequency of data collection (once a year for seven years), and the number of indicators (six subtest scores per child per year). But this system already has a limitation. It does not retain data on children once they graduate from the sixth grade; that is, it never contains a full seven years of data for more than one group of children. If all data on all classes were maintained, the system would eventually contain 8,820 data points (30 children x 7 classes x 7 testing times x 6 subtest scores).

This system is limited in other ways, for it contains no information on children's educational experiences (which programs they were in, what textbooks they used, when or how often they received Title I services), on their personal characteristics (whether they are bilingual, their parents' income, their health), or on outcomes other than test scores (grade retention, tardiness or absence, social development). Yet, as we have seen, such information may be of interest for various comparisons that might be desired later on.

There are some mechanical ways of reducing the physical size of the files. For example, instead of retaining test booklets, scores could be transferred to individual summary sheets or note cards. This may seem a trivial modification, but 8,000 note cards take up far less room than 8,000 test booklets. Or only class averages could be kept, so that 30 individual files are reduced to one class average; but that may greatly reduce the potential for later comparison. The 30 children who were in the same early childhood program may have very different experiences later on, which cannot be recorded without individual records.

ORGANIZATION

The way in which information is organized may also help save space, and

will certainly affect ease of access. For example, if a given record contains all the data on one child, longitudinal assessment of individual children's progress will be easy, whereas comparisons of group averages at any one time will require pulling those scores out of their longitudinal sequence and averaging them for each group of interest.

Textbooks on information systems usually use the terms subsystem and integration in discussing organization. Subsystems are certain sections of the files, and integration is the way in which they are linked, or cross-referenced. There are three ways of integrating subsystems of information: horizontal integration, vertical integration, and longitudinal integration (see Figure 8). Since each is suitable for a different kind of analysis, it pays to think about what analyses you will want to do before establishing the system.

Horizontal integration links subsystems that contain different kinds of information -- for example, one containing financial data with another containing achievement data. The best method of horizontal integration depends on how the data will be used. If, for example, all the questions asked will be about the relationship of these two domains, you may want to store the two sets of information together, say on a note card containing both cost and achievement data on a given child. But if important questions may relate only to cost (or only to achievement), that storage method will make data retrieval cumbersome; a better method would be to store the cost data separately and to code the achievement data by the relevant cost categories, so that it can be connected as needed.

Vertical integration is the summing up of data. If, for example, you want to average achievement across classrooms, school buildings, or the entire district, data should be stored so as to facilitate aggregation; in this

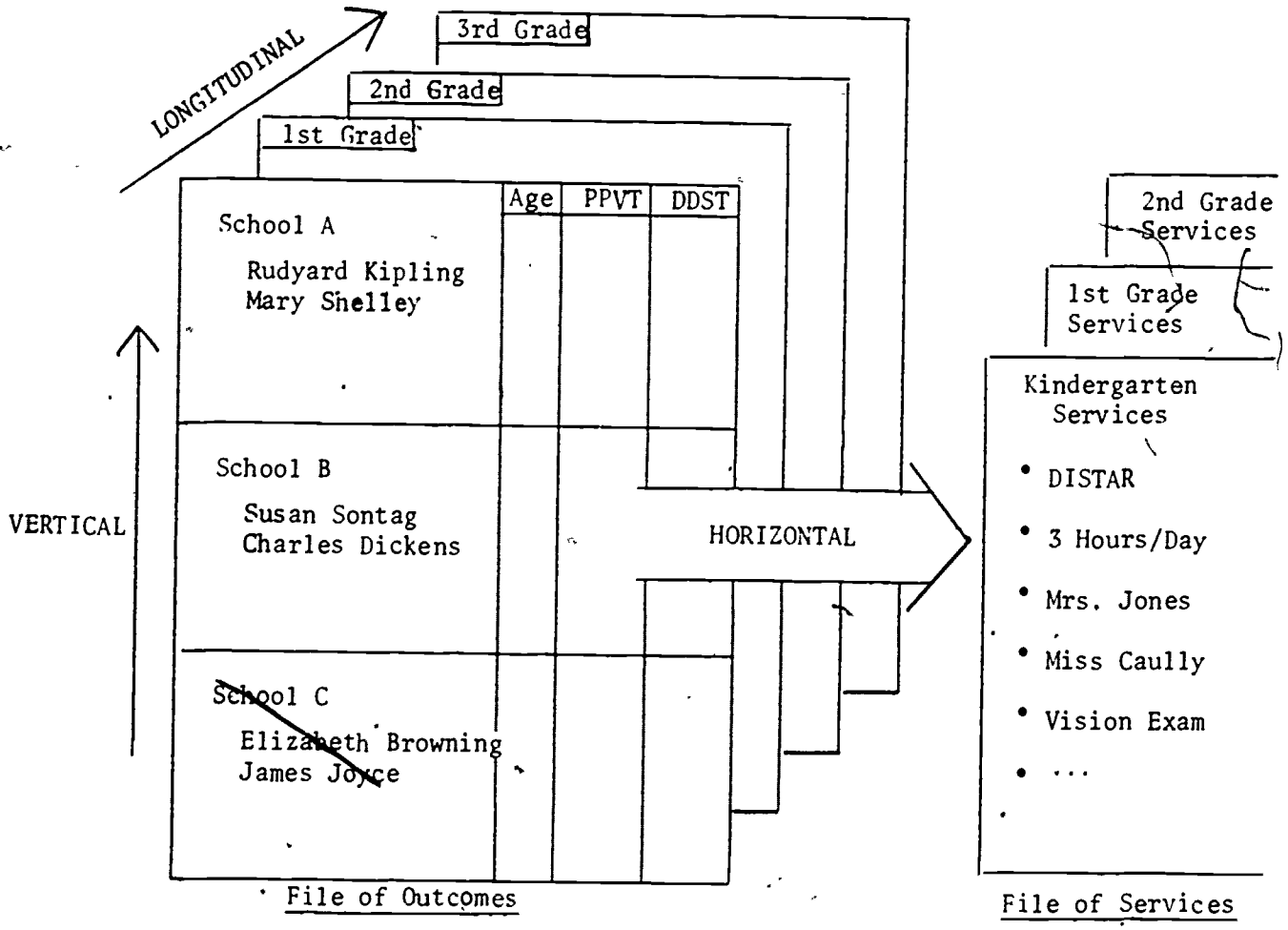

Figure 8: Three Ways of Integrating Subsystems

case, each unit (classroom, school) may be a subsystem. Aggregation across units also means that each subsystem should flag the data that should not be aggregated. For example, if some children were pretested and then moved to another school district, you may want to exclude their scores when averaging. Or if a flood caused one school building to close for a month, you may not want to average the scores of the children in that building with those of children in others. The individual records in each subsystem, then, should indicate any unusual circumstances that might affect the averages and render them uninterpretable.

Longitudinal integration links information over time. For example, you may want to compare the achievement at third and at sixth grade of children who attended school in your district for their entire elementary school years. Since you do not want to compare all available scores, the grade-specific subsystems must be integrated in a way that allows you to identify the children you are comparing. To preserve confidentiality, you might want to use some identification code rather than children's names.

The kinds of integration method used will depend on the content of the subsystem. For example, one subsystem might contain data on services, and another data on achievement; or one might contain first-grade and another second-grade achievement data. Subsystems are merely filed; but since a longitudinal data system may contain information on (a) different kinds of children, (b) different kinds or quantities of services, (c) different kinds of outcomes, and (d) different years, some thought should be given to how the subsystems should be defined, and how linked.

One way to do this is to think about information that will not be summed or averaged. For example, if you think you will never (or rarely)

want to pool data on English-speaking and Spanish-speaking children, then you may want to develop separate subsystems for these groups. Or if you will never want to pool data on children who did and did not have early childhood programs, then again the two sets of data might be stored separately. If you know you will not average first-grade test scores with second-grade test scores, then these data should be separate subsystems, coded so that they can be linked when comparisons are desired.

Conversely, you might think about information that you will want to link. For example, if you know you will often want to tie data on services to data on academic progress, you may want to store both together in one subsystem. Test score averages can then be computed for each kind of service delivered.

To define categories of service, you may need to develop some table outlines for the kinds of comparison that might be made. Figures 9 and 10 illustrate two such tables for defining subsystems. Each column or row (or cell) can represent one subsystem in the total descriptive information system.

HOW INFORMATION IS ENTERED INTO THE SYSTEM

You have determined the size of your system and how it should be organized; now you must consider how information will be collected and stored. It is one thing to decide to transfer scores onto note cards; quite another to find someone with the time to do it. If separate subsystems are to be developed for each building, perhaps someone in each building could record those data; this is especially useful if the data are to be used mainly by the staffs in the different buildings. If, however, the data are intended for use by a central office, it might make more sense if someone at that office took charge of all records. In either case, there are costs: in collecting the data, in transmitting them from the classroom to the person

Early Childhood Experiences	Later-Grade Title I Services						
	None	2nd Grade Only	3rd Grade Only	4th Grade Only	Grades 2nd and 3rd	Grades 2nd, 3rd and 4th	Grades 3rd and 4th
Preschool Only							
Kindergarten Only							
1st Grade Only							
Preschool and Kindergarten							
Kindergarten and 1st Grade							
Preschool, Kinder-garten and 1st Grade							

(Combined data for classes entering in 1974, 1975, 1976, and 1977).

Figure 9: Number of Children Receiving Different Amounts of Title I Services

Eligibility Determined by	Primary Parent Language			
	Spanish	Italian	Vietnamese	English
Denver Developmental Scale				
Gates-MacGinitie				
Peabody				

(Averaged across children entering kindergarten in 1976, 1977, and 1978)

Figure 10: Average Second-Grade Achievement for Children Admitted by Different Criteria

who maintains the system, and in recording, storing, and updating them as needed.

HOW INFORMATION IS RETRIEVED

The fourth decision about the design of your system is how the information will be retrieved for use. For example, will you want regular reports on certain things, or would you rather simply get information as you need it? Most people probably want both. For example, regular reports on enrollment and budgets may be useful for annual budget-planning meetings, and regular reports of achievement may be useful for accounting to the public. On the other hand, many important decisions that cannot be anticipated may require special analyses. For these occasions, an ideal system would include someone who could respond to special requests by pulling appropriate data, constructing whatever tables are needed, and calculating whatever averages or percents might be required.

By this time it should be clear that not all spontaneous requests for data analysis will be equally easy to meet. The size and organization of the system will necessarily favor some kinds of analysis over others. The way the data are stored may make it extremely time-consuming and tedious to retrieve and organize information for uses that were not anticipated.

PRIVACY OF RECORDS

Longitudinal information systems require that children can be identified in some way so that their scores over time can be linked. In the past, when school records were considered the property of the schools, this was no problem. Today, however, children's right to privacy is a controversial issue. Schools can no longer completely control the disposition of school records; parents, and in some cases children, also have a say about the use

of grades, test scores, and teachers' comments written in students' files. The chief guarantee of parents' and children's rights in this matter is the Family Educational Rights and Privacy Act, often known as the Buckley Amendment, which became law in 1974. This Act, which applies to any institution receiving program funds through the Department of Education, has two main provisions:

- To guarantee parents or guardians (and students over the age of 18 or who attend a post-secondary school) access to their own records
- To restrict other people's access to these records by requiring parental consent before the data can be released.

The second provision is particularly significant for longitudinal information systems, which require that the identity of students be safeguarded. This poses no problem in studies you conduct, since you have a legitimate educational interest and are therefore not required to secure parental consent before reviewing a child's record. The use of records outside the district, however, requires written permission--even if students' are replaced by numbers or other codes, since identities still could be traceable. These rules should be taken into account in the design of the system, so that they can be observed while your own information needs are met.

COMPUTER VS. MANUAL SYSTEMS

While almost all other prices are rising, the price of computers is declining. And small computers that can perform a variety of tasks are becoming increasingly available. This trend suggests that it may be worth while for you to take time to weigh the costs and benfits of a small computer relative to those of manual information systems.

Computers have many advantages. For example, if you know that you will want certain tables produced each spring, you can program your computer to

produce them automatically; you can also program it to respond to new requests--say, the average reading scores of children who have been retained in second grade vs. those of children there for the first time. Computers can have built-in safeguards, so that only certain personnel have access to the information. And of course their calculations are not subject to human error.

But computers have disadvantages, too. Though many of the new ones use languages very close to English, you will still need to train someone in using your new computer. And because they are mechanical devices, they can break down, so some amount of repair and maintenance will be needed.

Finding out about computers is as simple as looking in the yellow pages under Data Processing Systems. Your investigation should include consideration of the following.

- Capability to Meet Your Needs: You should be fairly sure of your system needs (size, organization, etc.) so that you can assess the specific capabilities of different computers to meet those needs.

- Small Computer vs. Terminal: If you get a small computer, you have the whole thing right in your building. If you get a terminal, you have telephone access to a much larger computer that resides elsewhere. Usually you buy computers, but you rent terminals. Terminals vary in the extent to which the files of information can be stored in the terminal itself as opposed to in the main computer.

- Maintenance: You should seek a system that includes a maintenance service agreement as part of the purchase or lease.

- Training: Check to see how complicated the computer is to program and use, and how much staff training will be needed. Also check to see how much training will be offered as part of the sale or lease.

- Characteristics of the Files: The information you keep on file can be stored in the machine iself or on auxiliary equipment such as cards, tapes, or "floppy disks." It is not safe to keep it only in the machine, for a power failure could mean the loss of the entire file. Most people today prefer floppy disks to other storage methods. Be sure to investigate

the advantages and disadvantages of each for (a) entering the information, (b) revising or updating it, and (c) retrieving it.

- Technical Assistance: We suggested earlier that you should start small, and increase the size and scope of your system only after you are sure of your needs. Check which computer companies offer assistance in determining your needs, or programming the computer to meet them. Don't buy until you are sure that you know what you want and that this system is best suited to the uses you have defined.

WHERE TO GO FROM HERE

We have described several characteristics of information systems that must be determined before the system is designed: size, organization, how information is entered into the system, how it is retrieved from the system, privacy of records, and computer vs. manual systems. These characteristics are interrelated. System size and organization will depend on the staff or resources available, on the form in which information will be retrieved, and on the intended use. Since few school districts can afford a system containing all possible useful information, system characteristics are partly a function of how large and sophisticated a system the district can afford. Thus, the costs and benefits of different system designs must be compared, relative to the value of the information they will provide for you.

What do you do once you have read this book? Start with a lot of thinking. Make some of the wish lists suggested here to help you estimate your needs. Talk to people. Consult with your Title I Technical Assistance Center staff. Then start a small trial information system. See how you use it. See what questions people ask of it. Gradually, you will refine and revise your system as you enlarge it, so that eventually--perhaps as much as five years from now--you will have a useful information system.

NOTES ON SOURCES OF FURTHER INFORMATION

Writing Computer Programs in BASIC. Wellesley Hills, MA: Applied Decision Systems, 1970.

This book is a manual on how to use a particular computer program system. The BASIC program is designed especially for use on computer terminals, and the manual provides guidance on how to use a terminal, and how to link it with a computer, as well as on how to use the BASIC systems on your terminal. The total book is just over 50 pages.

Baltes, Paul B., Reese, H.W., and Nesselroade, J.R. Life-Span Developmental Psychology: Introduction to Research Methods. Monterey, CA: Brooks/Cole, 1977.

This convenient little paperback book provides a readable introduction to the longitudinal concepts and the various methodological issues involved in studying and analyzing developmental change. Not only do these authors discuss research design and control issues, but they also have chapters on measurement problems in longitudinal data-collection efforts and data analysis and interpretation. The book is divided into five major parts, and one of them deals exclusively with descriptive systems.

Campbell, Donald T. Temporal changes in treatment-effect correlations: A quasi-experimental model for institutional records and longitudinal studies. Published in the Proceedings of the 1970 Invitational Conferance on Testing Problems. Princeton, NJ: Educational Testing Service, 1970 (pp. 93-110).

This 30-page article discusses the problems of determining the effects of special programs when the ideal conditions of an "experiment" cannot be realized. Campbell demonstrates the use of correlation coefficients as tools to assist in the analysis of data, and provides some very well thought-out

discussions about how to interpret different longitudinal patterns in the data.

Diederich, Paul B. Pitfalls in the measurement of gains in achievement. School Review, 1956 (Vol. 64), pp. 59-63.

Although several articles have been written on this topic since Diederich's, this short and straightforward article offers a good statement of some of the most prevalent problems in interpreting data on children's gains.

Grossman, Alvin, and Howe, R.L. Data Processing in Education. Chicago: Educational Methods, Inc., 1965.

Although much of the information about particular computers is now out of date, this book offers a variety of helpful advice on how to make the initial decisions, such as selecting a computer, estimating what it will be able to do for you, and managing the transition from a manual filing system to an automated system. The book also discusses some of the economic and personnel considerations in using computers. Most of these topics are discussed in the first half of the book. The second half describes various applications of automated record-keeping in education, such as test scoring and reporting, attendance accounting and bookkeeping.

Nie, Norman H., Hull, C. Hadlai, Jenkins, J.G., Steinbrenner, K., and Bent, Dale H. SPSS: Statistical Package for the Social Sciences. New York: McGraw-Hill, 1975.

This book is actually a manual on how to use a particular package of computer programs, the SPSS. The SPSS computer programs are most widely used by researchers, since they include a wide variety of complex statistical analyses. The bulk of this very large book describes the specific

procedures for using these programs to conduct specific analyses, but the first two chapters (a total of 27 pages out of over 600) describe general information about coding data so that these SPSS programs can be used to analyze it. These first two chapters are particularly worth reading if you think you may want to conduct very complex analyses and may therefore want to use the SPSS programs on your computer.

Tuckman, Bruce W. Conducting Educational Research. New York: Harcourt Brace Jovanovich, 1972.

This textbook is an excellent introduction to a variety of different aspects of educational research. Three chapters may be particularly useful to those who are planning to develop longitudinal information systems: Chapters 5 and 6 describe some of the considerations involved in making comparisons among groups, and Chapter 10 gives very explicit guidance on how to code data for computer data processing. It also shows examples of hand-written summaries of data, key-punched cards, and printouts. (Chapter 10 is about 30 pages.)

REFERENCES

Lazar, I., Hubbell, V.R., Murray, H., Rosche, M., & Royce, J. The persistence of preschool effects: A long-term follow-up of fourteen infant and preschool experiments. Washington, D.C.: Department of Health, Education and Welfare, 1977.

Miller, L., & Dyer, J. Four preschool programs: Their dimensions and effects. Monographs of the Society for Research in Child Development, October 1975.

EXAMPLES OF LONGITUDINAL STUDIES OF EARLY CHILDHOOD PROGRAMS

Becker, W. C. Teaching reading and language to the disadvantaged -- What we have learned from field research. Harvard Educational Review, 1977, 44 (November), 518-543.

Bronfenbrenner, U. Is early intervention effective? In M. Guttentag & E.L. Struening (Eds.), Handbook of evaluation research (Vol. 2). Beverly Hills, CA: Sage, 1975.

Brown, B. (Ed.). Found: Long-term gains from early interventions. Boulder, CO: Western Press, 1976.

Pelavin, S.H. & David, J. L. Evaluating long-term achievement: An analysis of longitudinal data from compensatory programs. (Research Report EPRC 4537-15). Menlo Park, CA: SRI International, 1977.

Spodek, B. & Walberg, H.J. (Eds.). Early childhood education: Issues and Insights. Berkeley, CA: McCutchan, 1977.