INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
EVALUATING THE IMPLEMENTATION OF SITE-MANAGED EDUCATIONAL PROGRAMS: THE RELATIONSHIPS OF PROGRAM CHARACTERISTICS AND IMPLEMENTATION CONTEXT WITH LEVEL OF IMPLEMENTATION

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

EDUCATIONAL PSYCHOLOGY

AUGUST 1995

By

Zhigang Wang

Dissertation Committee:

Paul R. Brandon, Chairperson
Mary E. Brandt
Ronald H. Heck
Kiyoshi Ikeda
Morris K. Lai
# TABLE OF CONTENTS

**CHAPTER 1. INTRODUCTION AND BACKGROUND** ........................................... 1  
The Importance of Evaluating Program Implementation and the  
Significance of This Study ................................................................. 2  
Implementation Theory for Program Evaluation .................................... 4  
Research Questions .................................................................................. 5  

**CHAPTER 2. VARIABLES RELEVANT TO EVALUATING THE**  
**IMPLEMENTATION OF SITE-MANAGED EDUCATIONAL PROGRAMS** ............. 6  
Variables About the Program ................................................................. 7  
Variables About Program Implementation Levels .................................. 9  
Variables About Program Implementation Context ............................... 10  
Variables About Program Implementors ............................................... 12  
The Relationships Among the Variables .............................................. 13  

**CHAPTER 3. METHOD** ............................................................................ 15  
Variable Selection .................................................................................. 15  
Instrument ............................................................................................. 19  
Data Collection ....................................................................................... 21  
Preparing Data for Secondary Analyses .............................................. 22  
Structural Equation Modeling for Data Analysis .................................. 27  
A Model of Implementation .................................................................... 34  
Analyses of the Structural Model ......................................................... 35  

**CHAPTER 4. RESULTS** ........................................................................ 40  
Distribution of the Data and the Input Matrixes for LISREL Analyses .... 40  
The Relationships Between Observed and Latent Variables ................. 41  
Model Fit ................................................................................................. 45  
Summary .................................................................................................. 50  

**CHAPTER 5. DISCUSSION, IMPLICATIONS, AND CONCLUSIONS** ............ 52  
The Construct Validity of the Model ...................................................... 53  
Implications of the Model for the Practice of Program Evaluation  
and for Research on Program Evaluation ............................................ 63  
Conclusions .............................................................................................. 63  

**APPENDIX A. SAMPLE QUESTIONNAIRE** ........................................... 65  

**APPENDIX B. VARIABLES FOR STUDYING PROGRAM IMPLEMENTATION** ......................................................................................................................... 68  

**APPENDIX C. CORRELATIONS BETWEEN OBSERVED VARIABLES** .... 75  

**REFERENCES** ....................................................................................... 76
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Explanations and Examples of the Three Types of Activities in the Local Side-Managed Program for This Study</td>
<td>25</td>
</tr>
<tr>
<td>Table 2. Distribution Statistics on the Interval Variables in the Model</td>
<td>40</td>
</tr>
<tr>
<td>Table 3. Results of Confirmatory Factor Analysis: Initial Loadings of the Observed Variables on the Latent Variables</td>
<td>41</td>
</tr>
<tr>
<td>Table 4. Results of Confirmatory Factor Analysis: Loadings of Final Set of Observed Variables on the Latent Variables</td>
<td>43</td>
</tr>
<tr>
<td>Table 5. Goodness of Fit Statistics of the CFA Model</td>
<td>43</td>
</tr>
<tr>
<td>Table 6. Goodness of Fit Statistics for the Structural Model</td>
<td>47</td>
</tr>
<tr>
<td>Table 7. Total, Direct, and Indirect Effects in the Model</td>
<td>47</td>
</tr>
<tr>
<td>Table B1. Variables About Programs, Program Outcomes, Program Implementation Context, and Program Participants and Their Inclusion in the Model of This Study</td>
<td>68</td>
</tr>
<tr>
<td>Table C1. Listwise Polychoric and Polyserial Correlations Between Observed Variables</td>
<td>75</td>
</tr>
</tbody>
</table>


## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>Proposed Model of Program Implementation at Site-Managed Educational Programs.</td>
<td>14</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>A Theoretical Model of Program Implementation</td>
<td>36</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>Final Structural Model of Implementation of Site-Managed Educational Programs</td>
<td>48</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION AND BACKGROUND

In the recent educational reform movement in the United States, many programs have been implemented at the school level. In some of these programs, funds have supported school-initiated improvement efforts at multiple sites. These programs are known as site-managed educational programs.

This study was conducted to propose and test a theoretical model of program implementation for evaluators of site-managed educational programs. The model is useful for (a) expanding our understanding of the variables in the implementation of site-managed educational programs, (b) guiding program evaluators in their selection and investigation of variables for studies of program implementation, (c) helping program evaluators interpret findings of these studies, and (d) providing formative evaluation information for program improvement.

The dissertation is organized in five chapters. In the present chapter, I provide the theoretical background for the study. In it, I discuss the importance of evaluating program implementation, the significance of the study, and the literature on program implementation theory. In Chapter 2, I review the critical program-implementation variables discussed in the literature on school improvement, program implementation, and program evaluation. In Chapter 3, I describe the methods and procedures for building and testing the theoretical model of program implementation. In Chapter 4, I present the model-testing
results. Finally, in Chapter 5, I discuss the findings, describe implications of the findings for evaluation theory and practice, and present conclusions.

The Importance of Evaluating Program Implementation and the Significance of This Study

Many program evaluators have focused their evaluation studies primarily on program outcomes, presuming a simple treatment-outcome model for program implementation and evaluation (Trochim & Cook, 1992). The disadvantage of this approach is that these studies do not provide sufficient meaningful and useful information for interpreting the evaluation findings and for improving the programs (Scheirer, 1987).

For program evaluations to generate adequate information for interpreting program outcomes, program evaluators need to look into the “black box” (Chen & Rossi, 1984), that is, the process of program implementation. By examining this “box,” evaluators can see how programs have been implemented and discover the features of the implementation process that might have helped or hindered implementation.

In the literature, the term “implementation” has been used with a wide range of connotations. What constitutes implementation depends on what is being evaluated (Brekke, 1987). For example, if the evaluand is a curriculum, the implementation may involve using curriculum materials and procedures with the appropriate target group. If the evaluand is a policy, the implementation may
entail following the specifications of the policy. Furthermore, if the evaluand is a site-managed educational program, the implementation is to carry out the “enabling activities” (Brekke, 1987, p. 16; see also Scheirer, 1987) that are designed by schools according to the guidelines for accomplishing the mandated program mission. Because schools in such programs are encouraged to design and conduct activities appropriate to their needs, program implementation varies from site to site.

In the past two decades, increasing numbers of researchers and evaluators (e.g., Brandon, 1992; Brandon, 1993; Chen, 1990; Fullan, 1985; King, Morris, & Fitz-Gibbon, 1987; Scheirer, 1987) have emphasized the importance and usefulness of studying program implementation. Studies on program implementation have shown that program implementation is extremely complicated and difficult within the human service areas. The difficulties and complications involved in implementation provide a serious challenge for evaluators to develop their own strategies and conceptual frameworks with which to deal with these problems (Chen, 1990, p. 31).

One important contribution of the recent theoretical and empirical studies about program implementation is that they have identified variables in program
characteristics, and in their implementation context, that affect program implementation levels (Pinto & Prescott, 1990). However, these studies often identify "too many variables" related to program implementation (Goggin, 1986, p. 335). Given the time and funding constraints in most program evaluations, it is impossible to consider all the variables in any evaluation. As some researchers have said (e.g., Goggin, 1986; Roberts-Gray, 1985), there is a need to weigh the importance of these variables and identify those that are crucial to program implementation in various settings. This study is an attempt to address this need.

Implementation Theory for Program Evaluation

Scheirer (1987) has defined implementation theory as a theory that describes the variables in the program implementation context and procedures that should be examined to explain the extent of program implementation. Although a few other researchers have discussed the variables in the implementation context and procedures and their effects on program implementation levels (for example, Bickman, 1987; Chen, 1990), only Scheirer has distinguished program implementation theory from program theory. She argues that program and implementation require different theories:

Implementation is a complex but understudied problem that can easily undermine well-conceived social programs. Whatever the
theory behind an innovation, the change must be implemented in
the day-to-day routine of the target social agency in order for it to
produce measurable effects (p. 67)

Recent empirical and theoretical research on program implementation has
focused on using program theory to guide program evaluation (Chen & Rossi,
1992). The use of implementation theory has been either subsumed under
program theory or largely overlooked (Scheirer, 1987). This study addresses
this deficit by emphasizing implementation theory in program evaluation.

Research Questions

This study addressed two research questions: What can we conclude about
the construct validity of the model? What are the implications of the model for
program evaluation theory and practice? The first question was addressed by
examining (a) the reliability of the observed variables and the latent variables
that can be defined with the observed variables and (b) the extent the latent
variables fit the data and account for the variance and covariance in the data.
The second question was addressed by making recommendations for practice
of program evaluation and for research on program evaluation.
To collect information about variables relevant to evaluating program implementation at the school-level, I searched the bibliographic data base of the Educational Resource Information Center (ERIC) for documents produced from 1980 through 1995. The literature search was based on three key terms combined: school improvement, program implementation, and program evaluation. A total of 84 references were found. The majority of them were technical evaluation reports for small local programs. Only a few were theoretical discussions of variables for evaluating school-level program implementation. Using the reference lists of these publications, I found and reviewed additional theoretical discussions about variables for studying program implementation at school level. The variables discussed in the literature as critical to the implementation of educational programs are summarized in Table B1, Appendix B and are described in this chapter.

Some researchers (for example, Fullan, 1993; Pinto & Prescott, 1990; Scheirer, 1987) have presented conceptual frameworks that illustrate the relationships among program implementation variables. For example, Fullan (1991) presented a framework of implementation that included three types of variables: program characteristics, the external governing system, and the
school characteristics. As seen in Table B1, I have categorized variables into four types: (a) variables about program characteristics, (b) variables about implementation levels, (c) variables about program implementation context, and (d) variables about program implementors. A definition for each of the variables is given in the second column of the table. From the four sets of variables identified in the literature, a theoretical model was developed and tested.

Variables About the Program

Variables about the program are useful for describing the characteristics of the program, such as the origin and history of the program, the clarity and appropriateness of program objectives, the size and characteristics of the target group, program implementation steps and schedule, and the importance and feasibility of the program.

The variables about the origin and history of the program are examined to describe where the program has been developed and how long the program has been implemented at the school. It has been noted in the literature that programs that have been initiated by a school, or implemented previously at a school, are likely to experience smooth implementation and achieve high implementation levels (Goggin, 1986; Firestone & Corbett, 1988).
The variables about *clarity and appropriateness of program objectives* are useful for describing the extent the objectives of the program are clearly defined and appropriate for schools in which the programs are to be implemented. Pinto and Prescott (1990) described findings showing that clearly defined objectives that were appropriate to schools improved the implementation process and implementation levels of programs.

The variable, *size and characteristics of the target group*, is examined to describe the school population to which the program is intended to provide services. Many researchers (e.g., Fullan, 1982; Pinto & Prescott, 1990; Roberts-Gray, 1985; Goggin, 1986) have observed that the size and type of the target group often affect the implementation process and level of implementation. For example, it is easier to implement a program that provides services to a small number of students that need the services than to implement a program to larger group of students who do not have the need for the services.

The variable, *program implementation steps and schedule*, is used to describe the steps and the time allowed for implementing each of the stages. Pinto and Prescott (1990) found that programs with well-planned steps and schedules were more likely to be highly implemented.

The *importance and feasibility of the program* are two variables for describing the extent a program is important for addressing school needs and the extent
a program is feasible to implement, given a school’s capabilities. Many researchers (e.g., Fullan, 1985; Pinto & Prescott, 1990; Roberts-Gray, 1985; Goggin, 1986) have observed that important and feasible programs are more likely to be fully implemented than less important, unfeasible programs.

Variables About Program Implementation Levels

Program implementation levels may be measured with four variables, extent of program implementation, quality of program implementation, frequency of use, and mutual adaptation. Extent of program implementation is a variable useful for describing the extent to which a program has reached its target group and completed its planned activities during implementation. The variable, quality of program implementation, is useful for describing the effectiveness and efficiency of program implementation procedures (Scheirer, 1987; Pancer & Westhues, 1989). A well-implemented program should effectively and efficiently provide quality services to its intended target group. The variable frequency of use is used to describe how often the materials and techniques prescribed in a program (such as a curriculum) are used with the target group in program implementation. A high frequency of use usually indicates high level of program implementation (Fullan, 1989). Finally, the variable, mutual adaptation is used to describe the changes in both the program and the implementation sites that occur during program implementation. In a
highly implemented program, not only is a school improved because of program implementation, but the program is also modified to better fit the school's situation and meet the school needs (Berman & McLaughlin, 1978).

Variables About Program Implementation Context

Because all site-managed educational programs are implemented at schools, the school culture provides the context for the program implementation. The extent that the conditions in the context are favorable for program implementation helps determine the extent that the program will be implemented smoothly and fully.

Variables about program implementation context are used to describe the components of the school culture that affect program implementation results. Previous studies (e.g., Chen & Rossi, 1992; Firestone & Corbett, 1988; Pinto & Prescott, 1990) have identified a large number of variables that can be used to describe school culture. Among them, the following have been discussed as critical to successful program implementation: school leadership support, school administrator's emphasis on academics and student attitudes, communication, collaboration, teacher input in decision making, external pressure and support for program implementation, rewards, resources, training, and program monitoring and evaluation. Among these variables, school leadership support probably is the most important one. It is used to describe the extent
that school administrators provide proper leadership for program implementation (Scheirer, 1987; Firestone & Corbett, 1988). The *administrators' emphasis on student academic achievements* and *administrator's emphasis on student attitude improvement* are variables denoting the extent a school focuses on the most important aspects of schooling (Firestone & Corbett, 1988). *Communication* among school staff, parents and community is a variable used for describing the flow of information about the program and its implementation among the school stakeholders (Pinto & Prescott, 1990). *Collaboration* among school staff for program implementation is a variable useful for describing the extent school staff jointly implement programs (Pinto & Prescott, 1990; Fullan, 1989; Scheirer, 1987). The variable about *teacher input* is used to describe an aspect of the decision-making process at the school. With this variable, we can measure the extent teachers play active roles in school management. (Goggin, 1986; Hord & Hall, 1987; Firestone & Cobett, 1988; Fullan, 1989; Pinto & Prescott, 1990). *External pressure and support for program implementation* is a variable for describing (a) the requirement the funding source of the program places on a school and (b) the support it provides for program implementation (Fullan, 1989). The variable about *rewards* helps us describe the rewards and incentives given to school staff for their participation, risk taking, and accomplishments in implementing the program (Fullan, 1989).
Resources available for implementing the program is a variable for describing the adequacy of funding, the availability of time, and the adequacy of information for implementing the program (Firestone & Borbett, 1988; Pinto & Prescott, 1990). Training is a variable for describing the training and in-service provided to participants for program implementation (Goggin, 1986; Evans, 1986; Pinto & Prescott, 1990; Roberts-Gray, 1985). Monitoring and evaluation is a variable for describing the managerial procedures for implementing the program (Pinto & Prescott, 1990).

Each of the variables in the school context contribute to one aspect of program implementation at the school-level. Together they constitute the process of program implementation.

Variables About Program Implementors

In site-managed educational programs, the group of program implementors usually consists of the school administrators and teachers. Sometimes, it also involves other school staff (e.g., school librarian), parents, and students. A study of program implementation should examine program implementors because their background, qualification, attitudes and so forth have a considerable effect on implementation. For instance, motivated and skillful program staff are more likely to implement program activities easily with better results than those who are less willing or less skillful. Variables that are used
to describe the characteristics of program implementors are abundant but only a few have been discussed as critical for program implementation, including background of the implementors (age, education, years of teaching, and years of working at current school), their familiarity with the school improvement program, their attitudes towards the school improvement program, their qualifications for implementing the program, and their motivation for participating in implementing the program (Goggin, 1986; Evans, 1986; Pinto & Prescott, 1990; Roberts-Gray, 1985).

The Relationships Among the Variables

To sum up, 26 variables have been identified in the literature on school improvement, program implementation, and program evaluation as crucial to successful program implementation at schools. These variables fall into four groups: variables about the characteristics of the program, variables about the implementation levels, variables about the program implementation context at the schools, and variables about the program implementors.

The following model of program implementation shows the relationships among the four types of variables in the implementation of site-managed educational programs.
Figure 1. Proposed model of program implementation at site-managed educational programs.
CHAPTER 3
METHOD

In this chapter, I describe the methods of the study, including selecting variables to build the model of the study, preparing the data for analysis, the structural equation modeling, the LISREL model, and model testing steps.

Variable Selection

Although all the variables discussed in last chapter are important for program implementation, not all of them are useful for evaluating the implementation of site-managed educational programs. From the 26 variables discussed in last chapter, I selected 12 to operationalize the theoretical constructs for this study. The other variables, although important for program implementation, were excluded from this study for one or more of the following reasons: (a) the variable did not apply to site-managed programs examined in this study, (b) there was typically no variation among the program sites on the variable, and (c) it was not feasible for program evaluators to collect data on the variable under typical program evaluation time and funding constraints. In Column 3 of Table B1 in Appendix B, I show whether the variables were selected and, in Column 4, I show the reasons for the selection decisions. In this section, I describe how the variables were selected for this study.
Variables About Program Characteristics

Of the six variables in this category, two were selected: importance and feasibility of the program. In site-managed educational programs, the program activities implemented at the schools are of different importance and feasibility for implementation. The variation in these variables might result in different levels of program implementation at the schools.

The other four variables in this category were not included in this study: (a) the origin and history of the program, (b) clarity and appropriateness of program objectives, (c) size and nature of the target group, and (d) program implementation steps and schedule. The reason for excluding these four variables from the model is because typically there is no variation among site-managed schools in these variables. For instance, site-managed educational programs are usually developed by the schools according to general program guidelines that apply to all schools. Furthermore, the target group participating in the program is usually selected according to common standards and similar student needs.

Variables About Implementation Levels

Of the four variables in this category, two were included in this study: extent and quality of program implementation. The variable on frequency of use was not selected because it is more appropriate for studying the
implementation of other types of programs (such as a curriculum) in which special prescribed teaching materials or methods were used. The other variable about implementation level, mutual adaptation between school and program, was also excluded from this study because it was more concerned with the effects of the program than with the implementation process of the program.

Variables About Program Implementation Context

Of the eleven variables in this category, six were selected. These variables were: leadership support for program implementation, administrator's emphasis on student academic achievements, administrator's emphasis on improving student attitudes, communication, collaboration, and teacher input into school decision making.

The other five variables in this category were not selected for this study: (a) external pressure and support for program implementation, (b) resources available for implementing the program, (c) training available to participants for implementation, (d) rewards, and (e) monitoring and evaluation. The reason for excluding the first two variables is because they typically do not have variation across the program schools. For example, the schools in the same program usually have the same level of funding and have the same resources for implementation. The third variable is not included because it is not consistent across all school sites; not all enabling activities in the type of programs
examined in this study require training. The last two variables (i.e., reward and evaluation) were not included in this study because it was not feasible for the evaluators to collect questionnaire data under constraints in typical program evaluation situations (e.g., lack of time and funds).

Variables About Program Implementors

Of the five variables in this category, only two about the background of program implementors are selected for this study. This variable is included because earlier exploratory analysis on the data (Brandon, Wang, Chan, & Peecook, 1993) indicated that the background of the respondents had potential effects on their perceptions of the program, the program implementation levels, and school implementation context. The other four variables in this category are not included because site-managed programs are not likely to have variation among program sites on these variables. For example, teachers at different schools should have the same or similar levels of familiarity and attitudes towards the activities implemented at their schools.

Summary

From the variables discussed in the literature, I selected 12 to build the model for this study: importance of the program, feasibility of the program, extent of program implementation, quality of program implementation, leadership support for program implementation, administrator's emphasis on
student academic achievements, administrator's emphasis on improving student attitudes, communication, collaboration, teacher input into school decision making, years since the respondents began working at schools, and years the respondents have worked at their current schools.

Instrument

Program implementation data for this study were collected with a questionnaire (shown in Appendix A) used for the evaluation of a local site-managed educational program in School Year (SY) 1992-93. In the evaluation, the evaluators paid careful attention to the development of the questionnaire. Portions of the questionnaire had first been developed in SY 1990-91 and improved in SY’s 1991-92 and 1992-93. All versions of the questionnaire were carefully pilot-tested at school sites (Brandon, Wang, Chan, & Peecook, 1993; Brandon, Wang, & Harman, 1992; Brandon, Wang, Harman, & Lindsey, 1991).

The 1992-93 version of the questionnaire consisted of four parts. In the first part, the evaluators explained the purpose of the questionnaire, reassured the respondents of the confidentiality of their responses, and explained the procedures for completing and returning the questionnaire. The second part of the questionnaire consisted a set of questions asking about the respondents’ background and status at the school. In the third part, four questions asked
about each activity's importance, feasibility, extent of implementation, and quality of implementation. This part of the questionnaire addressed the specific activities at each of the schools. Because the number of activities varied among schools, the number of times the four questions were asked was different from school to school. In the fourth part of the questionnaire, six questions asked about the characteristics of the program implementation context at the schools, including the administrator's emphasis on students' academic achievement and their attitudes toward school, school leadership support for program implementation, communication and collaboration among the school staff for program implementation, and teacher input into decisions about the program implementation.

Items in the questionnaire carefully operationalized the variables in the program implementation theory. The issue of potential halo effect in the responses to the questions was partially addressed by asking questions about specific activities instead of the entire program (although it was possible that there was a slight halo effect on responses to questions about school context). Other potential issues, such as normality of the data distribution and measurement errors, were addressed by the data analytical methods.

---

1A fifth question, not examined in this study, asked about the extent the activity achieved its objectives.
Data Collection

The data for this study were collected in School Year (SY) 1992-93 at 28 schools (22 elementary schools and 6 secondary schools) for the evaluation of the implementation of a local site-managed educational program. This program was designed by the Hawaiʻi Department of Education (DOE) to help selected schools address their “special needs characterized by high percentages of at-risk students and those who are chronically low achievers” (Hawaiʻi Department of Education, 1989, p. I). In this program, the DOE allocated funds to participating schools and mandated program objectives and guidelines for program planning and implementation. Based on the goals and guidelines, the schools designed and conducted activities to address their problems and needs.

In April 1993, CRDG administered the data collection questionnaire to all administrators and certificated staff (teachers, counselors, and librarians) at the 28 schools (N=1,304). The questionnaire return rate was 87.4% for all the schools combined (N=1,139). Individual school return rates ranged from 52% to 100%.

---

2 The objectives of the program included (a) improve rates of attendance, graduation and promotion, (b) improve student achievement, staff stability and retention, and school climate, and (c) decrease the number of disciplinary referrals and suspensions. Annual evaluations of the program have been conducted by the Evaluation Office of Curriculum Research & Development Group (CRDG), University of Hawaiʻi at Manoa. In SY 1992-93, the evaluation was in its third year and included an evaluation of program implementation.
Preparing Data for Secondary Analyses

Why is it Necessary to Select a Subset of Data for This Study?

The 1992-93 SNS evaluation data set included information about (a) the program, (b) program participants, and (c) program implementation context. For the variables about program participants and implementation context, the respondents were asked to provide information about themselves and the entire school, respectively. For variables about the program, the respondents were asked to rate each program activity instead of the entire program. The advantage of asking questions about the program activities was that, taken together, the ratings on activities could provide more accurate information about the program and its implementation than answers to overall questions about the program. The disadvantage, however, is that the information about the entire program is not easily summarized from the diversity of activities implemented at the schools.

To summarize the data on program implementation for this study, it was necessary to first combine the ratings about activities to generate data about the program and then analyze the data across program sites. However, the nature of site-managed educational programs poses some potential problems for combining the activity-level data into program-level data. As encouraged in such programs, the numbers and kinds of activities conducted in the program
varied considerably among the schools. Furthermore, the implementation of the activities involved varying numbers of participants; when responding to the evaluation questionnaire, the respondents provided information only about the activities in which they participated.

Given the diversity of the program activities and the responses on the data collection questionnaire, it may not be appropriate to combine and analyze all of the activities together. Instead, the combination and analyses must be based on a selected subset of activities (a) that are substantively and statistically similar for all schools and (b) for which a sufficiently large number of respondents answered questionnaire items at each school.

*How Were the Activities Selected?*

First, substantively, the selected activities should address the program's major objectives. For the SNS program, the major program objectives are to improve students' academic achievement and their attitudes towards school. The selected activities, therefore, should be those that aimed at (a) developing students' knowledge and skills in major academic areas, (b) providing support (such as materials, supplies, and equipment) for classroom instruction, and (c) enhancing students' personal development (in areas such as self-esteem, behavior, and social skills). These categories of activities are conceptually similar in that they all directly serve the students and were conducted at all
program schools. Explanations and examples of the three types of activities are given in Table 1.

To avoid any potential effects of school level (elementary schools vs. secondary schools) or location (urban schools vs. rural schools) on the data, only the elementary schools (N=8) in one urban district were included in the analyses for this study.

Second, statistically, the ratings on the selected activities should form a homogeneous, or internally consistent, set of measurements, indicated by high internal consistency coefficients among the selected activities for each school. If the reliability coefficient of the set of selected activities is low, the activities are not internally consistent and, therefore, it would not be appropriate to combine the activity ratings.

To examine the internal consistency of the selected data, I calculated Cronbach's alpha on the two variables of implementation outcomes (extent and quality of activity implementation) using the data on the substantively selected activities. Cronbach's alpha was repeatedly calculated on the selected activities for each school, each time deleting the activities with low item-total correlations until all were above 0.5 and the coefficient alpha for the remaining group of activities was greater than 0.80. A total number of 40 activities were
Table 1. Explanations and Examples of the Three Types of Activities in the Local Site-Managed Educational Program for This Study

<table>
<thead>
<tr>
<th>Type of activity</th>
<th>Explanation</th>
<th>Examples of school enabling activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing students' knowledge and skills in major academic areas</td>
<td>Activities designed to help student performance in language arts, mathematics, science, and social studies.</td>
<td>• Use mathematics manipulative to help students meet the new mathematics standards.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Operate learning and interest centers for each classroom in Grades 4 and 5.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Enhance the whole language program, develop students' global awareness through integrated social studies activities, and increase their knowledge of science through newspaper reading.</td>
</tr>
<tr>
<td>Providing support for classroom instruction</td>
<td>Activities designed to provide educational aides, materials, supplies, or equipment for classroom instruction.</td>
<td>• Provide each teacher with $478 to purchase books for language arts and cooperative learning activities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Provide teachers with classroom manipulative and hands-on supplies to be used with students (calculators for Grade 3, AIMS [science] booklets, and mathematics manipulatives for all other grades).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Employ additional part-time teachers to reduce student-teacher ratio at all grade levels.</td>
</tr>
<tr>
<td>Enhancing students' personal development</td>
<td>Activities designed to help students improve their self-esteem, behavior, and attitude towards school.</td>
<td>• Continue the administrators' luncheon for students-of-the-quarter.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Employ a part-time teacher to conduct a counseling program for the at-risk students identified from all classes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Improve Grades 3, 4, and 5 students' oral communication skills and self-esteem through Alliance for Drug Education.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Conduct leadership workshops for Grades 3-6 students to develop their leadership skills.</td>
</tr>
</tbody>
</table>

25
selected, with four or five activities for each school, to combine activity ratings into program-level data.

*Combining the Activity-Level Ratings into Program-Level Data*

Activity-level ratings were combined to generate program-level data for four variables: (a) importance of the program, (b) feasibility of the program, (c) extent of program implementation, and (d) quality of program implementation. For all respondents, the combined data were based on the selected activities for their schools.

The activity-level ratings of the four variables were combined in the following steps. First, for each activity in each school, the means for the four variables were calculated. Ratings on program importance illustrate how program ratings were developed for each respondent. A respondent's ratings on the importance for the selected activities were averaged, resulting in a mean importance rating for the respondent. Second, missing data were replaced with the appropriate activity mean. Third, raw activity-level mean ratings were standardized with means of 100 and standard deviations of 10. Finally, standardized mean activity ratings were calculated, resulting in one final mean rating on each of the four variables for each respondent.
The Final Data Set for Testing the Program Evaluation Model for This Study

The final data set for testing the program implementation model for this study included observations on 229 respondents, with 12 variables per observation. The variables were (a) the number of years since the respondent began working at schools, (b) the number of years the respondent had worked at the current school, (c) the administrator's emphasis on student academic achievement, (d) the administrator's emphasis on student personal development, (e) the school's leadership support for program implementation, (f) the communication among school staff during program implementation, (g) the collaboration among the school staff during program implementation, (h) teacher input into decisions about program implementation, (I) the importance of the program, (j) the feasibility of the program, (k) the extent of program implementation, and (l) the quality of program implementation.

Structural Equation Modeling for Data Analysis

Among recommended analytic methods, Messick (1989) illustrated how structural equation modeling (SEM) could be used for construct validation. In this study, I used the SEM method to test the construct validity of the program implementation model. In so doing, I (a) appraise the soundness of the selected variables and the underlying constructs in the model and (b) examine the relationships among the constructs. A familiarity with SEM is useful for
understanding the analytic procedures and results. Therefore, in this section, I review the literature on using SEM for building and testing social theories and describe LISREL, the computer program for SEM that I used for analyzing the data. Finally, I specify the model of program implementation with the variables selected earlier in this chapter.

**Structural Equation Modeling in Social Inquiry**

The growing interest in studying the relationships among social variables, the increasing availability of data from large-scale studies, and the development of computer programs have "led to greater utilization of sophisticated methods of data analysis, most particularly structural equation modeling" (Walberg & Reynolds, 1992, p. 210). Since the late 1980s, the literature about SEM has been accumulating, and computer programs for SEM are now widely available. In the field of program evaluation, SEM has been recommended as a method for testing program theory (Chen & Rossi, 1987; McClintock, 1987). This fast growing popularity of SEM in social science and program evaluation is largely due to its powerful data analysis approach that, "properly employed, offers great potential for theory development and construct validation in psychology and the social science" (Anderson & Gerbing, 1988, p. 422).

When properly used, SEM can be a powerful approach for proposing, developing, testing, and validating social theories in a number of ways. The
foremost, perhaps, is its heuristic feature that helps researchers reduce data and interpret findings. As many researchers (for example, Chen & Rossi, 1989; Costner, 1989; Heck, Marcoulides, & Glasman, 1989; Lipsey & Pollard, 1989; Wang & Walberg, 1983) have pointed out, SEM forces researchers to consider the expected causal relationships among observed variables and theoretical constructs in an explicit framework. To define the framework, the researcher must clarify and explain the theoretical reasons for allowing certain variables to interrelate. Therefore, SEM encourages the researcher to think clearly and in detail about the causal mechanisms underlying the correlations in the data.

Second, SEM is useful in theory construction because "it provides a convenient and efficient way to test theories and their inherently complex effects" (Walberg & Reynolds, 1992, p. 220). Theories of social process often involve many interrelated variables whose overall effects cannot be easily understood and explained. Many of these effects involve underlying hypothetical constructs. With the help of SEM, we can simultaneously estimate the effects of underlying variables and their relationships included in the model. By carefully and rigorously examining the relations among constructs, we can have greater precision about social phenomena, which is often not possible with simple correlation, regression, or qualitative data analysis. Additionally, because SEM can examine the indirect effects of
variables, it can aid the researcher in identifying more complex causal relations through investigating the variables that may have been discounted because they lacked direct effects.

Third, SEM is flexible in handling problems of measurement errors and omitted variables (Walberg & Reynolds, 1992). The statistical basis of SEM methods provides a means of inference regarding the plausibility of a model. It provides information on how well the selected variables serve as estimates of the latent constructs and how much of the variance in the model is explained by the included variables. In this manner, the researcher can assert the likelihood of important effects of included versus possible missing variables.

Despite its recent growing popularity in social science research, SEM, like other data analytic methods, can be misused. When inappropriately applied, SEM may result in seriously distorted models, misleading users of the results. Therefore, researchers should follow guidelines to improve the soundness of their models. First, they should use theories of the field and findings of previous studies to guide the selection of variables and construction of models (Walberg & Reynolds, 1992). Second, they must understand the principles of SEM and improve their skills for conducting SEM analysis (Muthen, 1992). Third, they must collect reliable and valid data and develop thoughtful models to explain the structure underlying the data (Cliff, 1992). Finally, they should
conduct exploratory data analysis before using rigorous modeling procedures (Muthen, 1992; Jöreskog & Sörbom, 1993).

The LISREL Approach to SEM

In this study, the model for evaluating program implementation was tested with the computer program LISREL VIII (Jöreskog & Sörbom, 1993). This program is popular because of its ability to test a wide variety of proposed theoretical models. In this section, I explain some technical terms of the LISREL program.

Observed variables and latent variables. Most structural models consist of two types of variables, observed variables and latent variables. An observed variable is measured with data collection instruments such as questionnaires. Latent variables are theoretical constructs that are defined by measuring the observed indicators of these variables because they cannot be directly observed. Many theoretical concepts in social science, such as self-esteem, motivation, and prejudice, are formulated with these kinds of hypothetical constructs (Long, 1983).

Measurement component and structural component. There are two components in the LISREL model: the measurement component and structural component. In the measurement component, the latent constructs are identified through confirmatory factor analysis (CFA) on the covariances among the observed
variables. Any proposed model must be well measured. This helps establish the construct validity of the latent variables. In the structural component, the relations among the identified constructs are estimated through (a) specifying the causal relations among the constructs, (b) determining the strengths of such relations, and (c) specifying the amount of variance in the data that is explained by the model.

**Errors in a model.** Two types of errors are measured in model testing: the random measurement error associated with each observed variable and the residual error in each structural equation representing the unexplained variance in each latent construct of the model. Measurement errors result from problems in the data collection procedures. The residual error of latent variables results from variables and relationships that are not included in the model. Both types of errors are indicators of the soundness of the model.

**Model testing.** To test a model is to examine how well the hypothetical model defined by the researcher fits the empirical data. A well-fit model is one that "not only fits the data well from a statistical point of view but also has the property that every parameter of the model can be given a substantively meaningful interpretation" (Jöreskog & Sörbom, 1993, p. 115).

**Output of LISREL analysis.** The output of an LISREL analysis shows (a) the latent constructs that can be defined from the observed variables, (b) the
reliability of the observed variables as estimates of the latent constructs, (c) the strength of the causal relationships among the latent constructs, (d) the amount of variance in the data that is left unexplained by the model, and (e) indexes of the model's fit to the empirical data. LISREL output also provides recommendations to the researcher on how to modify the model to improve its fit with the data. However, "only the researcher is capable of judging the balance between statistical and substantive model fit," and "model respecification must remain the decision of the researcher and not of the LISREL program" (Byrne, 1989, p. 57).

Path diagram. A LISREL model is conventionally depicted with a path diagram for ease of presentation and interpretation. In a LISREL path diagram, the observed variables are enclosed in rectangles and the latent constructs are enclosed in circles or ovals. Causal relationships among constructs in a model are indicated by straight lines with arrows, leading from the causal variables to the affected variables. The strength of the effect is written on the line. The unexplained portion of variance (or error) in a latent construct is shown in parentheses, with an arrow pointing at the corresponding latent construct.

In the next section, I specify the model for evaluating program implementation with the selected variables.
A Model of Implementation for Site-Managed Educational Programs

There are three latent constructs in the model presented in Figure 2. The first construct, *school program implementation context*, represents the program implementation context at schools. It represents the focus, support, and collaboration among the administrators, faculty, and staff implementing the program. It is measured with six observed variables indicating respondents’ opinions about administrators’ emphasis on student academic achievement, administrators’ emphasis on student attitude improvement, leadership support for program implementation, communication among faculty and administrators, collaboration among faculty and administrators, and teacher input into decisions about the program implementation. The second construct, representing *program characteristics*, is measured with two observed variables: the respondents’ opinions about the importance of the program activities and about the feasibility of the activities. The third construct, representing the *program implementation level*, is comprised of two observed indicators: respondents’ opinions about the extent of program activity implementation and about the quality of the implementation. Finally, the model includes two demographic variables: (a) the number of years since the respondents began working at schools and the number of years the respondents have worked at their current schools. These two demographic variables are considered primarily...
because they are known to affect teachers' opinions of schooling (e.g., see Brandon, Wang, & Heck, 1994), an issue of validity.

Figure 2 shows the hypothesized relationships among the observed and latent variables. First, the elements in the program implementation context are hypothesized to have significant direct and indirect (i.e., through program characteristics) effects upon program implementation level. Second, program characteristics also have a significant direct effect upon program implementation levels. Third, the variables for the number of years since the respondents began teaching and the number of years they have worked at their current schools may account to some extent for their opinions about the program implementation context, program characteristics, and program implementation levels. A relationship found between either of these two demographic variables and the variables for respondents opinions would suggest possible limitations in the application of the model.

Analyses of the Structural Model

The two goals of the data analysis in this study are to (a) estimate the strength of the independent variables in explaining the level of program implementation and (b) assess the amount of variance in the level of program implementation that can be accounted for by the variables included in the structural model. The data were analyzed in three stages: (a) using the PRELIS
Figure 2. A theoretical model of program implementation.
computer program to examine the distribution of the data and to generate input matrixes for the LISREL analysis, (b) using confirmatory factor analysis (CFA) to examine how well the latent variables are defined by the observed variables, and (c) using LISREL VIII to estimate the structural relations among latent variables in the model. In this section, I explain each of the steps.

Step 1: Using PRELIS for Examining the Distribution of the Data and to Generate Input Matrixes for the LISREL Program

Users of LISREL do not always have sufficient knowledge about the characteristics and problems of their raw data when using LISREL to test models. Without thorough understanding of the data, the user of LISREL might select an inappropriate modeling method for the data, resulting in biased and abnormal LISREL parameters (Jöreskog & Sörbom, 1993). Therefore, before using LISREL for more elaborate analysis of data, researchers should first find out about the distribution characteristics and quality of the data and create appropriate input matrixes for the LISREL program.

Following Jöreskog and Sörbom’s recommendation, the PRELIS 2 (Jöreskog & Sörbom, 1993) program was used for the first stage in data analysis. After checking the distribution characteristics of the data, two matrices were produced with PRELIS 2 as input matrixes for LISREL VIII, a polychoric and polyserial correlation matrix (Appropriate for ordinal and
interval data) and, to correct for any violations of normal distribution in the data (which is likely with ordinal data), an asymptotic variance-covariance matrix.

Step 2: Using CFA for Examining the Relationships Between Observed and Latent Variables

The objective of CFA is to examine how well the observed variables measure the hypothesized latent variables and construct and test the fit of a measurement model for further analysis with LISREL VIII. If the observed variables load highly (e.g., between .7 and 1.00) on the latent variables, we can conclude that the latent variables are valid underlying constructs for the observed variables and that the observed variables, in turn, are reliable measures of the latent variables (Kim & Mueller, 1990). The statistic significance of the loadings were tested through t-tests (the ratio of the loading to its standard error). T-values are significant at $p<.05$. The loadings, together with other model fit indexes, also provide the researcher with the information about whether the CFA model needs modification. CFA could be run several times, each time with the CFA model slightly modified, until all the loadings of the observed variables on their desired latent variables and CFA model fit indexes become acceptable.
Step 3: Estimating the Fit of the Structural Model

After achieving an adequate fit of the CFA model to the data, the final step of data analysis was to test the fit of a structural model. The weighted least squares (WLS) method of LISREL was used to analyze the data. This fitting method is appropriate when the model includes both interval and categorical variables and when there was a potential violation of normality of distribution in the data (Jöreskog & Sörbom, 1993)
CHAPTER 4
RESULTS

Distribution of the Data and the Input
Matrixes for LISREL Analyses

Table 2 shows the statistics on the distribution characteristics of the data.

Table 2. Distribution Statistics on the Interval Variables in the Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of program implementation</td>
<td>-1.31</td>
<td>2.68</td>
</tr>
<tr>
<td>Quality of program implementation</td>
<td>-1.24</td>
<td>2.48</td>
</tr>
<tr>
<td>Importance of the program</td>
<td>-1.69</td>
<td>4.03</td>
</tr>
<tr>
<td>Feasibility of the program</td>
<td>-1.07</td>
<td>1.53</td>
</tr>
<tr>
<td>Years since the respondent began working at schools</td>
<td>-0.59</td>
<td>-0.73</td>
</tr>
<tr>
<td>Years the respondent has been at current school</td>
<td>0.41</td>
<td>-1.37</td>
</tr>
</tbody>
</table>

As shown in Table 2, all except one variable are negatively skewed and the kurtosis for the variable of importance of the program is somewhat too high, suggesting that violation of normality may exist in the data. Therefore, weighted least squares (WLS) method of SEM should be used to estimate the model fit.

The polychoric and polyserial zero-order correlations among all the variables in the model are shown in Table C1, Appendix C for those who are interested in replicating this study. These correlations are produced from a variety of dichotomous, ordinal, and interval variables and result in more
accurate correlations than Pearson product moment correlations for these type of data (Jöreskog & Sörbom, 1993)

The Relationships Between Observed and Latent Variables

Table 3 shows the loadings of the variables on their corresponding latent variables in the first run of CFA.

Table 3. Results of Confirmatory Factor Analysis: Initial Loadings of the Observed Variables on the Latent Variables

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Observed variable</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of program implementation</td>
<td>Extent of program implementation</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>Quality of program implementation</td>
<td>.97</td>
</tr>
<tr>
<td>Characteristics of the program</td>
<td>Importance of the program</td>
<td>.84</td>
</tr>
<tr>
<td></td>
<td>Feasibility of the program</td>
<td>.93</td>
</tr>
<tr>
<td>School program implementation context</td>
<td>Administrator's emphasis on student academics</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Administrator's emphasis on student attitudes</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>School leader's support for program implementation</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>Communication among faculty and administrators</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Collaboration among faculty and administrators</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Teacher input into program decision making</td>
<td>.99</td>
</tr>
</tbody>
</table>

The initial results of CFA shown in Table 3 indicate that the factor loadings on five variables are greater than 1.0. These abnormal loadings suggest that some the variables are multicolinear and need to be either deleted or
combined to improve the loadings. I elected to combine some of the variables because, as Marsh (1994) has pointed out, the combinations of closely related variables "tend to be more reliable and valid indicators of the latent construct" (p. 4). In the following runs of CFA, two pairs of variables, (a) administrator's emphasis on student academics and on student attitudes and (b) communication among faculty and administrators and collaboration among faculty and administrators, were combined into two new variables. The first pair of variables were averaged into a new variable called *proper school focus* while the second pair of variables were averaged into a new variable called *communication and collaboration among faculty and administrators*.

In Table 4, the loadings of the final set of observed variables on their corresponding latent variables are shown. From the results shown in Table 4, we can see that the variable loadings have improved after the combination of variables (i.e., none of the loadings are greater than 1.0). The variable loadings ranged from .79 to 1.00, showing that the latent variables are well measured by the observed variables (with t-ratios greater than 2). The statistical indexes that describe the fit of the CFA model to the data are shown in Table 5.

The goodness-of-fit index (GFI) and the adjusted goodness-of-fit index (AGFI) can be considered measures of the relative amounts of variance and covariance in the data accounted for by the proposed model. It is generally
Table 4. Results of Confirmatory Factor Analysis: Loadings of Final Set of Observed Variables on the Latent Variables

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Observed variable</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of program implementation</td>
<td>Extent of program implementation</td>
<td>.91*</td>
</tr>
<tr>
<td></td>
<td>Quality of program implementation</td>
<td>1.00*</td>
</tr>
<tr>
<td>Characteristics of the program</td>
<td>Importance of the program</td>
<td>.79*</td>
</tr>
<tr>
<td></td>
<td>Feasibility of the program</td>
<td>1.00*</td>
</tr>
<tr>
<td>School program implementation context</td>
<td>Proper focus of the school efforts</td>
<td>1.00*</td>
</tr>
<tr>
<td></td>
<td>School leader support for program</td>
<td>.99*</td>
</tr>
<tr>
<td></td>
<td>Communication and collaboration among staff</td>
<td>.99*</td>
</tr>
<tr>
<td></td>
<td>Teacher input into school decision making</td>
<td>1.00*</td>
</tr>
</tbody>
</table>

*p<.05

Table 5. Goodness of Fit Statistics of the CFA Model

<table>
<thead>
<tr>
<th>Index name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>.95</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>.92</td>
</tr>
<tr>
<td>Root mean square residual (RMR)</td>
<td>.09</td>
</tr>
<tr>
<td>Normed fit index (NFI)</td>
<td>1.0</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>1.0</td>
</tr>
<tr>
<td>Chi-square (df)</td>
<td>12.88 (23) p&gt;.90</td>
</tr>
<tr>
<td>Chi-square/df ratio</td>
<td>.56/1</td>
</tr>
</tbody>
</table>
accepted that GFI and AGFI values close to or above .90 indicate a good model fit (the maximum possible value for the index is 1.0). For this model, the GFI is .95 and the AGFI is .92, indicating a good model fit. The root-mean-square residual (RMR), in contrast, is a measure of the variance and covariance that are unexplained in the model. For a good model fit this index should be close to 0. In this model, the RMR is .09, suggesting that little of the variance and covariance was left unaccounted for by the hypothesized model. This finding further confirms the soundness of model fit. Other indexes used to determine the assessment of fit are the normed fit index (NFI), comparative fit index (CFI), chi-square, and the ratio of chi-square to the degree of freedom. Generally speaking, a reasonable fit is considered to have NFI and CFI close or above .90 and a chi-square/df ratio of 2:1. For this model, the NFI and CFI were found to be 1.00\(^3\) and the chi-square/df ratio was .56:1, all indicating a well fit model. The significance value \((p)\) of the chi-square is greater than .90, suggesting that the model cannot be rejected and, therefore, can be considered a plausible representative of the data. Furthermore, the modification indexes for the final model indicated that no further parameters could be freed to significantly improve model fit. The conclusion, therefore, could be drawn that the proposed CFA model accurately accounts for the variance and covariance

---

\(^3\)The fit of 1.00 is the result of rounding.
in the data, and the model satisfactorily explains the relationships between the observed variables and their corresponding latent constructs. Having confirmed that the observed variables account accurately for the latent constructs, it is appropriate now to examine the fit of the structural model to the data, that is, how the latent constructs are related.

Model Fit

The Fit of the Overall Structural Model to the Data

The statistical indexes that describe the fit of the structural model to the data are shown in Table 6. The GFI and the AGFI for this model are 1.00, indicating a strong model fit. The RMR is .08, further confirming the soundness of model fit. The NFI and CFI for the model were found to be 1.00 and the chi-square/df ratio was .83:1, all indicating a well-fit model. The significance value \( (p) \) is greater than .70, suggesting the model cannot be rejected. Furthermore, the modification indexes for the final model indicated that no further parameters could be freed to significantly improve the model fit.

The conclusion, therefore, could be drawn that the proposed model accurately accounts for the variance and covariance in the data and the model satisfactorily explains the relationships among the variables (observed and latent) in the data.

---

4The fit of 1.00 is the result of rounding.
The Significance of Parameter Estimates in the Model

The parameter estimates in a model represent the simultaneous contribution of observed and latent variables in the overall model (Stage, 1990). The significance of the parameter estimates are tested with t tests (i.e., the ratio of the estimate to its standard error). In Figure 3, only the statistically significant relationships ($p < .05$) are shown. The direct and indirect effects among the constructs and the two demographic variables are summarized in Table 7.

As shown in Figure 3, 68 percent of the variance in program implementation level is accounted for by program characteristics and program implementation context and teacher demographics, leaving 32 percent of the variance due to variables outside the model (shown by the residual error estimated for that latent construct). As shown in the model, three variables account for 20 percent of the variance in program characteristics: program implementation context, years since began working at schools, and years at current schools. In the model, the number of years the respondents have worked at their current schools is the only variable that has statistically significant effects on the program implementation context, accounting for 2
Table 6. Goodness of Fit Statistics for the Structural Model

<table>
<thead>
<tr>
<th>Index name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>1.00</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>1.00</td>
</tr>
<tr>
<td>Root mean square residual (RMR)</td>
<td>.08</td>
</tr>
<tr>
<td>Normed fit index (NFI)</td>
<td>1.00</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>1.00</td>
</tr>
<tr>
<td>Chi-square (df)</td>
<td>23.43(28) p&gt;.70</td>
</tr>
<tr>
<td>Chi-square/df ratio</td>
<td>.83/1</td>
</tr>
</tbody>
</table>

Table 7. Total, Direct, and Indirect Effects in the Model

<table>
<thead>
<tr>
<th>Construct</th>
<th>Years since working at schools</th>
<th>Years at current school</th>
<th>Program implementation context</th>
<th>Program characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program implementation context</td>
<td>0.0</td>
<td>-.20</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Program characteristics</td>
<td>-.24</td>
<td>.17</td>
<td>.41</td>
<td>--</td>
</tr>
<tr>
<td>Program implementation level</td>
<td>0.0</td>
<td>0.0</td>
<td>.74</td>
<td>.69</td>
</tr>
</tbody>
</table>

Note: Total effects (in italics) include direct effects and indirect effects that result from correlations among exogenous variables and reciprocal effects; indirect effects (i.e., effects through combined paths) are in parentheses.
Figure 3. Final structural model of the implementation of site-managed educational programs. (Only the statistically significant \( p<.05 \) unstandardized path coefficients are shown.)
percent of its variance, with 98 percent of the variance due to factors outside the model or error in the model. The rest of the findings are summarized for the hypotheses introduced in page 35.

**H1: Program context has direct and indirect effects on program implementation.**
First, it was hypothesized that the elements in the program implementation context have statistically significant direct and indirect effects on program implementation. As shown in Figure 3 and Table 8, both the direct and indirect effects of program implementation context on program implementation level are statistically significant (.25 and .49, respectively). The results show that program implementation context has the strongest total effects (defined as the total of the indirect and direct effects, which is .74) on program implementation level.

**H2: Program characteristics have direct effects on program implementation.**
Second, program characteristics were hypothesized to have statistically significant direct effects on program implementation levels. This hypothesis is confirmed and the findings show that program characteristics have the strongest direct effects on program implementation level,

**H3: Teachers’ years of experience affect their impressions of the process.** Third, it was hypothesized that years since the respondents first began working at schools and the number of years they have worked at their current schools
might have effects on the respondents' opinions about the program implementation context, program characteristics, and program implementation levels. As shown in Figure 3 and Table 13, this hypothesis is only partially confirmed. More specifically, only the variable of years at current school has statistically significant negative effects on opinions about program context (-.20). Both variables have statistically significant effects on opinions about program characteristics (i.e., -.24 for years since working at schools, and .17 for years at current school). Neither variable has any statistically significant effects on opinions about program implementation levels.

Summary

The results of data analysis have shown that the observed variables are reliable measures of the three latent variables and that the theoretical model fits the empirical data satisfactorily. The fit of the model helps support its construct validity as a means of representing program implementation. Although not much of the variance in the program implementation context and program characteristics is accounted for by the model (2 percent and 19 percent respectively), the variables in the model account for 68 percent of the variance in program implementation levels. While program characteristics (i.e., importance and feasibility of the program) have the strongest direct effects on program implementation levels, program implementation context (e.g., school
leadership support, teacher input, and communication and collaboration among school staff) has the strongest total impact on the implementation levels. It may be concluded that the opinions about implementation levels of site-managed educational programs are very well accounted for by the observed and latent variables included in the model. The model tested in this study has therefore verified the importance of several variables for studying program implementation at the school level. Moreover, these observed indicators represent valid theoretical constructs for studying program implementation.

Having tested the model fit and examined the significance of the model parameters, I now turn to the implications of the findings for evaluating the implementation of site-managed educational programs.
CHAPTER 5
DISCUSSION, IMPLICATIONS, AND CONCLUSIONS

This study was built on current discussion of program implementation (see Chen & Rossi, 1992). It was conducted to verify a model of program implementation that program evaluators could use to study the implementation of site-managed educational programs. The model was developed from previous studies on program implementation, program evaluation, and school improvement. From the variables discussed in the literature, I identified those that are useful for studying site-managed educational programs. In an original contributions to the research on program evaluation, I examined the relationships among the variables and, using structural equation modeling, I tested the construct validity of a program implementation model.

The findings of my study have shed some light on the “black box” of program implementation and contributed to the theory and practice of program evaluation as well as program implementation. In this chapter, I discuss the findings and their implications for program evaluators by addressing the two research questions: (a) What can we conclude about the construct validity of the model? and (b) What are the implications of the model for program evaluation and practice? Finally, in this chapter, I present conclusions and make suggestions for future studies.
The Construct Validity of the Model

The construct validity of a statistical model is reflected in (a) the reliability of the observed variables for estimating the latent variables in the model and (b) the extent that the relationships among the latent variables fit the data and account for the variance and covariance in the data.

Reliability of the Observed Variables for Estimating the Latent Variables

The model of program implementation consists of three hypothetical latent variables: program implementation context, program characteristics, and program implementation levels. The model also consists of 10 observed variables that have been described in previous studies as being crucial to successful school program implementation. Of the 10 observed variables, four load on the latent variable of program implementation context, including, school leadership support for implementation, proper program focus, communication and collaboration among school administrators and staff, and staff input into decisions about school program implementation. Of the 10 observed variables, two others, importance of the program and feasibility of the program, load on the latent variable of program characteristics. Another two variables, quality of implementation and extent of implementation, load on the latent variable of program implementation levels. The other two variables about the background of the program implementors, which are number of years since
the respondents began working at schools and number of years they have worked in their current schools, are not designated to load on any of the latent variables because they are demographic variables. The statistical and substantive results of model testing showed that the observed variables load highly on their corresponding latent variables, indicating that they are reliable measures of the latent variables. This is also confirmed by the high goodness-of-fit indexes. Therefore, I conclude that the model has included reliable variables.

The implementation of social programs is a complicated process which may involve many variables. Among all the variables related to program implementation discussed in the literature, “to decide which ones to try and which to safely ignore requires experience, judgment, and serious knowledge of the subject matter” (Achen, 1992, p. 51). The selection of variables for building a model of site-managed program implementation useful for program evaluators should be based on at least two criteria. First, it should be based on whether a variable is theoretically sound and provides useful information for studying the implementation of programs across sites. Second, it should be based on whether it is feasible for evaluators to collect useful data on these variables under the constraints of limited time and funds, the typical conditions of program evaluation. Accordingly, to define a useful model for program
evaluators, variables that might be important for program implementation but not feasible for evaluators to collect useful data within the limits of typical program evaluation conditions should not be recommended to program evaluators.

Furthermore, variables sometimes may not be used exactly as they are described in the literature. Some variables are so closely related that in actuality they comprise one variable. In this case it is necessary to combine the closely related variables so that the data become more concise and accurate for analysis. For example, in this study, the results showed that administrators' emphasis on improving students' academic performance and on improving students' attitudes towards schools were so closely correlated that they could be combined into one variable representing the proper focus of the program. Another example is about communication and collaboration about program implementation among the school staff and school administrators. Again, these two variables could be combined as one variable representing communication and collaboration at the school. In my study, combining these two pairs of variables improved the loadings of the observed variables on latent variables because combining closely related variables "substantially reduces the number of parameters to be estimated and improves the ratio of the number of estimated parameters to the number of subjects" (Marsh, 1994, p. 7).
The Extent That the Relationships Among the Latent Variables Fit the Data and Account for the Variance and Covariance in the Data.

In light of the strong theory guiding the model's construction, the model-testing results (i.e., the findings on the model fitting indexes) show that the three-construct model satisfactorily fits the data, indicating that the model is a valid description of the implementation of site-managed educational programs. Moreover, the two independent latent variables (i.e., program characteristics and program implementation context) in the model account for 68% of the variance and covariance in the dependent latent variable (i.e., program implementation levels). The relatively high variance and covariance accounted for by the model further confirms the construct validity of the model representing the implementation of site-managed educational programs. Therefore, I conclude that this study has described a valid model of program implementation for site-managed educational programs.

The structural model (i.e., the relationships among the latent variables) of the study shows that, in school staff's opinions, program characteristics and program implementation context strongly affect the level of program implementation of site-managed educational programs. It indicates that, in school staff's opinions, high levels of program implementation are closely associated with program feasibility and importance and with a favorable implementation context. These empirical findings about the relationships
among the three constructs represent a significant contribution to the program implementation and program evaluation literature.

The model shows that program characteristics have the strongest statistically significant direct effects on the level of program implementation. This indicates that the feasibility and importance of the program as perceived by the school staff have the strongest direct effects on the extent and quality of program implementation. This provides support for Fullan's (1991) view that characteristics of the proposed change affect its likely implementation. Although the context of the program implementation is also thought to have statistically significant direct effects on levels of program implementation, these effects are only about one third as strong as that of the program characteristics.

However, program implementation context also has statistically significant direct effects on program characteristics and, through program characteristics, statistically significant indirect effects on the level of program implementation. Together, the direct and indirect effects of program implementation context make the elements in the program implementation context the strongest force influencing the level of program implementation. Since variables about program implementation are related to program implementation theory (Scheirer, 1987), this finding shows the importance of applying program implementation theory when conducting program evaluations.
Finding direct effects of program implementation context on program characteristics is not surprising because, for site-managed educational programs, the school context for program implementation is also the context for program development. This indicates that at schools where the program implementation context in program participants' opinions is favorable for program implementation, the school-developed programs are likely to be perceived as important and feasible. This suggests that elements in the school context, such as teacher input, leadership support, communication and collaboration, and focus of the program are relevant for studying characteristics of school-managed programs. This supports Fullan's (1991) contention that the school is another critical component of reform implementation.

Another interesting finding of the study is the statistically significant, although small, effects of both (a) the number of years the respondents had worked at schools and (b) the length of their employment at their current schools on their perception of the program characteristics, the school context, and the program implementation level. The findings suggest that the number of years school staff have worked at their current schools negatively affects their opinions about program context but positively affects their opinions about program characteristics. The findings also suggest that the number of years since the respondents first began working at schools negatively affects
their opinions about program characteristics. However, respondents' number of years of working at schools has no statistically significant effects on their opinions about program implementation context, and neither variable has any statistically significant effects on staff's opinions about program implementation levels. These findings are similar to the findings of Brandon, Wang, & Heck (1994), who found that the background of teachers had effects on the teachers' perception of the school context and their participation in school decision making, but not on their agreement with the decisions made. The effects of teacher background on their opinions about program implementation context should be examined further using ethnographic or other non-survey methods.

The effect of respondents' background on their opinions is a validity issue. That is, findings based on data collapsed across groups of respondents with different backgrounds might be somewhat invalid. Cautions for program evaluators based on these findings are evident: Evaluators should be aware of the effects of respondent background that may exist in their data.

The effect of respondents' background is also a substantive issue. Even though the length of the respondents' teaching experience did not affect their opinions about program implementation levels, it did affect their opinions of the program, suggesting that long-time teachers might be skeptical about new
programs (Weiss, Cambone, & Wyeth, 1992). Furthermore, the effects also suggest that the longer the teachers work at their current schools, the more negatively they think about their school context and the more positively they think about school improvement programs.

Implications of the Model for the Practice of Program Evaluation and the Research on Program Evaluation

The findings of the study have implications for both the practice of program evaluation and research on program evaluation. In this section, I discuss these implications with suggestions for future studies.

Implications for Evaluation Practice

The findings show that the model tested in the study is a valid representation of program implementation of site-managed educational programs. Therefore, the first implication for evaluation practice is that evaluations of program implementation levels should address two constructs: program characteristics and program implementation context. The examination of the program and the implementation context will help evaluators interpret the findings about implementation levels and provide formative recommendations for program improvement.

The second implication is that evaluators of the implementation of site-managed educational programs should strive to study the variables included in the model tested in the study. The ten observed variables in the model are
reliable indicators of the latent theoretical constructs. These variables are critical for the success of program implementation and feasible for evaluators to collect data under typical program evaluation constraints on time and funds.

Third, evaluators should not necessarily use variables exactly as they are described in the literature. As shown in this study, some of the variables described in the literature are highly correlated. Evaluators should identify the closely related variables in their data set and combine the data on the variables. By combining the data on the closely correlated variables, they can capture as much information as possible on the combined variables. In addition, evaluators can improve their data collection instruments (for example, reducing the number of items in questionnaires) and save valuable time and funds for other evaluation tasks.

Fourth, the relationship between program implementation context and program characteristics suggests that evaluators should examine the program implementation context for better understanding of the program characteristics. The findings of this study show that program implementation context at the schools can have direct effects on the characteristics of the program. By examining the components in the program implementation context, such as whether the teachers have been given opportunities to
participate in school decision making, evaluators can gain knowledge about the characteristics of the site-managed program.

Finally, the background of respondents might have an effect on their perceptions about programs and the school context. These effects can be harmful to the validity of the findings. Therefore, when designing data-collection instruments (e.g., a questionnaire), evaluators should include items about respondents' background. When analyzing the data that show the effects of respondent background, program evaluators should consider treating the respondents of different background groups separately. They must be careful about collapsing data across groups of respondents with different backgrounds and comparing findings among sites with program implementors of different background.

**Implications for Research on Program Evaluation**

First, researchers on program evaluation should go beyond the 10 variables included in the model and experiment with other critical variables. They can refine the model by adding more variables and defining additional constructs or redefining the three constructs found in the model. This is suggested because the model tested in the study does not completely account for the variance and covariance in the data. The unaccounted variance and covariance may be due
to variables that are important for program implementation but are not included in the model (see Table B1 in Appendix B).

Second, researchers can replicate the model testing method (i.e., the SEM approach) with data collected from rural schools, secondary schools, or other types of programs to see if the model found in this study still explains program implementation. Furthermore, researchers can separate the respondents according to their background and test the model again to find the evidence of the effects of respondent background on the relations among the observed and latent variables. In this manner, it should be possible to develop more refined knowledge of program implementation.

Conclusions

It is now widely realized that, for program evaluations to produce meaningful and useful information for the users of evaluation results, program evaluators need to examine the process in which a program has been implemented. The study reported here confirms the importance of applying program implementation theory when evaluating site-managed educational programs. By studying how programs have been implemented, program evaluators can provide information about the program’s intervening mechanisms and suggest ways to improve the program.
In this study I have demonstrated using SEM is an effective way to identify variables for program evaluation and study the relationships among the variables. This study, therefore, has opened new opportunities for future studies about program evaluation. This study has identified a theoretical model that satisfactorily accounts for the variance and covariance among the selected variables. It is useful for program evaluators to study the implementation of site-managed educational programs. Because the findings have shown some effects of teacher background on teacher perceptions about the program and the school program implementation context, background variables should also be included as statistical controls on teachers' perceptions of these processes.

Because the findings reported here are based on data collected from urban elementary schools, the generalizability and the application of the model may somewhat be limited. Obviously, the model tested in the study does not completely account for the variance and covariance in program implementation level. However, any model that helps reduce the complexity of organizational life to a finite series of measured variables is necessarily incomplete (Heck & Marcoulides, 1992). If, in our persistent efforts, we identify a few important variables at a time that are reliable and useful for program evaluators, eventually we will be approaching the complete understanding of the complicated process of program implementation.
APPENDIX A
SAMPLE QUESTIONNAIRE

Special Needs Schools Program Evaluation
Questionnaire for Fern Elementary School

Aloha.

As part of our evaluation of the Special Needs Schools Program, please complete this questionnaire about the program's implementation at your school in 1992-93. Your responses are confidential and will help improve the program.

If you complete this questionnaire during a faculty meeting, please seal it in the attached envelope when you are finished and give it to your principal before leaving the meeting. Otherwise, complete the questionnaire, seal it in the envelope, and return it to your principal's office by May 12, 1993.

Thank you for helping us evaluate your Special Needs Schools Program!

Questions About Your Background

1. What is your highest degree? (Check one box only.)
   - Bachelor's
   - Professional diploma
   - Master's
   - Doctorate

   For Questions 2-5, please put your answer in the space provided.

2. In what year did you begin your teaching career? __________

3. How many years have you taught in Hawaii public schools? __________

4. How many years have you been working at your current school? __________

5. How many years have you been in your current position or department or at your current grade level? __________

Questions About Program Activities

On the following pages, questions are asked about the Special Needs Schools Program activities conducted at your school in 1992-93. For each activity, please carefully read the activity statement and then answer the questions. Circle one number or check the box (for "don't know") for each question. When answering questions about activities conducted in classrooms, please consider how each activity was conducted throughout the school, not just in one classroom. If you do not recognize an activity, check "don't know" for each question about the activity.

-Please turn the page over-
### Activity No. 107.1: Conduct a school-wide speech festival for all students to develop their oral communication skills and prepare them for the district speech festival. (Awards to students were provided for this activity.)

<table>
<thead>
<tr>
<th>Item</th>
<th>Don't know</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How important was this activity for improving students' education?</td>
<td>not important</td>
<td>1 2 3 4 5 6 very important</td>
</tr>
<tr>
<td>2. How feasible was it to implement this activity as planned?</td>
<td>infeasible</td>
<td>1 2 3 4 5 6 quite feasible</td>
</tr>
<tr>
<td>3. To what extent were all steps of this activity implemented?</td>
<td>not at all</td>
<td>1 2 3 4 5 6 fully</td>
</tr>
<tr>
<td>4. How high was the quality of the implementation process (e.g., timeliness, attention to details, following procedures, etc.) of this activity?</td>
<td>very low</td>
<td>1 2 3 4 5 6 very high</td>
</tr>
<tr>
<td>5. To what extent has this activity achieved its objectives to date?</td>
<td>not at all</td>
<td>1 2 3 4 5 6 fully</td>
</tr>
</tbody>
</table>

### Activity No. 107.2: Expand and use classroom literature collections. (Funds were provided to all teachers to purchase books to be used by all students.)

<table>
<thead>
<tr>
<th>Item</th>
<th>Don't know</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How important was this activity for improving students' education?</td>
<td>not important</td>
<td>1 2 3 4 5 6 very important</td>
</tr>
<tr>
<td>2. How feasible was it to implement this activity as planned?</td>
<td>infeasible</td>
<td>1 2 3 4 5 6 quite feasible</td>
</tr>
<tr>
<td>3. To what extent were all steps of this activity implemented?</td>
<td>not at all</td>
<td>1 2 3 4 5 6 fully</td>
</tr>
<tr>
<td>4. How high was the quality of the implementation process (e.g., timeliness, attention to details, following procedures, etc.) of this activity?</td>
<td>very low</td>
<td>1 2 3 4 5 6 very high</td>
</tr>
<tr>
<td>5. To what extent has this activity achieved its objectives to date?</td>
<td>not at all</td>
<td>1 2 3 4 5 6 fully</td>
</tr>
</tbody>
</table>
Questions About the Entire Program

The questions on this page ask for your overall impression about the activities shown on the previous pages of this questionnaire. For Questions 1-6, please check one box only. For Question 7, write your comments (if any) about the implementation of the program at your school.

1. To what extent does your school administration emphasize the importance of improving students' academic achievement in the Special Needs Schools Program?
   □ Not at all  □ A little  □ Somewhat  □ Very much

2. To what extent does your school administration emphasize the importance of improving students' attitudes toward school and improving their self concept in the Special Needs Schools Program?
   □ Not at all  □ A little  □ Somewhat  □ Very much

3. Overall, how well did the school administration support the implementation of the activities shown on the previous pages of this questionnaire?
   □ Not at all  □ A little  □ Somewhat  □ Very much

4. Overall, how well did the school's administration, faculty, and staff communicate during the implementation of the activities?
   □ Not at all  □ A little  □ Somewhat  □ Very much

5. Overall, how well did the school's administration, faculty, and staff collaborate on implementing the activities?
   □ Not at all  □ A little  □ Somewhat  □ Very much

6. Overall, to what extent did you have input into decisions about the activities?
   □ I had no input  □ I had little input  □ I had some input  □ I had a lot of input

7. What comments or suggestions, if any, do you have about the Special Needs Schools Program at your school?

-Thank you again for your time and effort!-
APPENDIX B. VARIABLES FOR STUDYING PROGRAM IMPLEMENTATION

Table B1. Variables About Programs, Program Outcomes, Program Implementation Context, and Program Participants and Their Inclusion in the Model of This Study

<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>The origin and history of the program (Goggin, 1986; Firestone &amp; Corbett, 1988)</td>
<td>The organization (the school or other organizations) that developed the program; how long or how often the program has been implemented.</td>
<td>No</td>
<td>Site-managed programs usually have little variation in origin and history across the program sites.</td>
</tr>
<tr>
<td>Clarity and appropriateness of program mission, goals, and objectives (Pinto &amp; Prescott, 1990)</td>
<td>The extent the mission, goals, and objectives of the program are clearly defined and appropriate for the schools in which the program will be implemented.</td>
<td>No</td>
<td>Site-managed programs are not likely to have variation among program sites on this variable (the program mission and objectives are provided by the funding sources and the schools are selected according to common standards).</td>
</tr>
<tr>
<td>Size and nature of the target group (Fullan, 1982; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985; Goggin, 1986;)</td>
<td>The number and type of school population the program is intended to serve.</td>
<td>No</td>
<td>Site-managed programs have little variation among program sites on this variable.</td>
</tr>
<tr>
<td>Variable name and references</td>
<td>Definition</td>
<td>Included in model?</td>
<td>Why or why not included</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------</td>
<td>--------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Program implementation schedule (Pinto &amp; Prescott, 1990)</td>
<td>The duration and the planned implementation stages of the program.</td>
<td>No</td>
<td>Site-managed programs have little variation among program sites on this variable (the duration and implementation stages are usually specified by the funding sources).</td>
</tr>
</tbody>
</table>

Importance, complexity, and feasibility of the program (Fullan, 1985; Pinto & Prescott, 1990; Roberts-Gray, 1985; Goggin, 1986) | The importance, complexity, and feasibility of the program for the school to implement. | Yes | The importance, complexity, feasibility of the program are different for the school sites and may result in different levels of program implementation. |

**Variables about the implementation outcome**

<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of program implementation (Sheirer, 1987; Pancer &amp; Westhues, 1989)</td>
<td>The extent to which the implementation has reached its target group and completed its planned activities.</td>
<td>Yes</td>
<td>This variable must be examined to see how much the program has been implemented.</td>
</tr>
<tr>
<td>Quality of program implementation (Sheirer, 1987; Pancer &amp; Westhues, 1989)</td>
<td>The effectiveness and efficiency of the implementation procedures.</td>
<td>Yes</td>
<td>This variable must be examined to see how well the program has been implemented.</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of use (Fullan, 1989)</td>
<td>The frequency the prescribed materials, techniques, etc. are used</td>
<td>No</td>
<td>This variable is not useful for studying the implementation of site-managed programs, in which the implementation is carried out by conducting a variety of “enabling activities.”</td>
</tr>
<tr>
<td>Mutual adaptation between the program and the implementation site (Berman &amp; McLaughlin, 1978)</td>
<td>The modification of the program to fit the school's situation and the changes in the school because of the implementation of the program.</td>
<td>No</td>
<td>This variable is not useful for studying the implementation process of a program; the changes of the schools observed during the implementation may not be stable, nor final.</td>
</tr>
</tbody>
</table>

### Variables about the program implementation context

<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication among school staff, parents and community (Pinto &amp; Prescott, 1990)</td>
<td>The flow of information about the program and its implementation among the school stakeholders.</td>
<td>Yes</td>
<td>Communication among program participants is important for successful program implementation. The differences among the schools on this variable may result in different levels of program implementation.</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaboration among school staff for program implementation (Pinto &amp; Prescott, 1990; Fullan, 1989; Scheirer, 1987)</td>
<td>The joint effort of the school staff for implementing the program.</td>
<td>Yes</td>
<td>Collaboration among program participants is important for successful program implementation. The differences among the schools on this variable may result in different levels of program implementation.</td>
</tr>
<tr>
<td>Leadership support for program implementation (Scheirer, 1987; Firestone &amp; Corbett, 1988)</td>
<td>School administrators' proper leadership for implementing the program.</td>
<td>Yes</td>
<td>School leadership support is important for successful program implementation. The differences among the schools on this variable may result in different level of program implementation.</td>
</tr>
<tr>
<td>Decision-making process (Goggin, 1986; Hord &amp; Hall, 1987)</td>
<td>Democratic decision making procedures at the school for making decisions for program implementation.</td>
<td>Yes</td>
<td>How program implementation decisions are made at the school is important for successful program implementation. The differences among the schools on this variable may result in different level of program implementation.</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal opportunities for teacher and other stakeholder participation (Firestone &amp; Cobett, 1988, Fullan, 1989; Pinto &amp; Prescott, 1990)</td>
<td>The active role the stakeholders play and their formal responsibilities in implementing the program.</td>
<td>Yes</td>
<td>Whether school staff are given opportunities to participate is important for successful program implementation. The differences among the schools on this variable may result in different level of program implementation.</td>
</tr>
<tr>
<td>External pressure and support for program implementation (Fullan, 1989)</td>
<td>Requirement and support of an outside organization given to the school for implementing the program.</td>
<td>No</td>
<td>No variation among the schools on this variable; the schools are under the same external pressure and support for program implementation.</td>
</tr>
<tr>
<td>Rewards (Fullan, 1989)</td>
<td>Rewards and incentives given to school staff for their participation, risk taking, and accomplishments in implementing the program.</td>
<td>No</td>
<td>It is not easy to collect accurate objective data on this variable.</td>
</tr>
<tr>
<td>Resources available for implementing the program (Firestone &amp; Borbett, 1988; Pinto &amp; Prescott, 1990)</td>
<td>Skillful personnel, adequate funding, available time, and up-to-date information for implementing the program.</td>
<td>No</td>
<td>No variation among the schools on this variable.</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training available to participants for implementation (Goggin, 1986; Evans, 1986; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985)</td>
<td>Training provided to participants for implementing the program</td>
<td>No</td>
<td>Not all of the “enabling activities” in the program requires training. Therefore, no systematic data could be collected from the schools on this variable.</td>
</tr>
<tr>
<td>Monitoring, troubleshooting, and evaluation (Pinto &amp; Prescott, 1990)</td>
<td>Managerial procedures for implementing the program.</td>
<td>No</td>
<td>It is not easy to collect accurate objective data on this variable.</td>
</tr>
</tbody>
</table>

**Variables about program participants**

<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background of the participants (Goggin, 1986; Evans, 1986; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985)</td>
<td>Age, education, years of teaching, and years of working at current school.</td>
<td>Yes</td>
<td>This variable reflect the teachers’ experience for implementing educational programs and their familiarity with the program being implemented at their schools</td>
</tr>
<tr>
<td>Familiarity with the school improvement program (Goggin, 1986; Evans, 1986; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985)</td>
<td>Participants’ knowledge of the program.</td>
<td>No</td>
<td>No variation among the schools on this variable and it is not easy to collect accurate objective data on this variable.</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Variable name and references</th>
<th>Definition</th>
<th>Included in model?</th>
<th>Why or why not included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes towards the school improvement program (Goggin, 1986; Evans, 1986; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985)</td>
<td>Participants' attitudes towards the program.</td>
<td>No</td>
<td>No variation among the schools on this variable and it is not easy to collect accurate objective data on this variable.</td>
</tr>
<tr>
<td>Qualifications for implementing the school improvement program (Goggin, 1986; Evans, 1986; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985)</td>
<td>Participants' skills and qualification for implementing the program.</td>
<td>No</td>
<td>No variation among the schools on this variable and it is not easy to collect accurate objective data on this variable.</td>
</tr>
<tr>
<td>Motivation to participate in the school improvement program (Goggin, 1986; Evans, 1986; Pinto &amp; Prescott, 1990; Roberts-Gray, 1985)</td>
<td>Participants' willingness to participate in implementing the program.</td>
<td>No</td>
<td>No variation among the schools on this variable and it is not easy to collect accurate objective data on this variable.</td>
</tr>
</tbody>
</table>
APPENDIX C. CORRELATIONS BETWEEN OBSERVED VARIABLES

<table>
<thead>
<tr>
<th>Observed variables</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
<th>Y5</th>
<th>Y6</th>
<th>Y7</th>
<th>Y8</th>
<th>Y9</th>
<th>Y10</th>
<th>X1</th>
<th>X2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrator's emphasis on academics (Y1)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrator's emphasis on attitudes (Y2)</td>
<td>.94</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School leaders support implementation (Y3)</td>
<td>.84</td>
<td>.82</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication among faculty and administrators (Y4)</td>
<td>.69</td>
<td>.70</td>
<td>.75</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaboration among faculty and administrators (Y5)</td>
<td>.78</td>
<td>.76</td>
<td>.79</td>
<td>.92</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher input into decision making (Y6)</td>
<td>.59</td>
<td>.64</td>
<td>.64</td>
<td>.62</td>
<td>.72</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance of the program (Y7)</td>
<td>.42</td>
<td>.37</td>
<td>.39</td>
<td>.34</td>
<td>.38</td>
<td>.23</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feasibility of the program (Y8)</td>
<td>.39</td>
<td>.37</td>
<td>.39</td>
<td>.31</td>
<td>.38</td>
<td>.33</td>
<td>.75</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extent of program implementation (Y9)</td>
<td>.44</td>
<td>.49</td>
<td>.45</td>
<td>.39</td>
<td>.42</td>
<td>.37</td>
<td>.59</td>
<td>.77</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of program implementation (Y10)</td>
<td>.48</td>
<td>.50</td>
<td>.49</td>
<td>.43</td>
<td>.47</td>
<td>.38</td>
<td>.62</td>
<td>.78</td>
<td>.93</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years since beginning teaching (X1)</td>
<td>-.13</td>
<td>-.04</td>
<td>-.01</td>
<td>-.08</td>
<td>-.10</td>
<td>.00</td>
<td>-.09</td>
<td>-.10</td>
<td>-.06</td>
<td>-.04</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Years at current school (X1)</td>
<td>-.05</td>
<td>-.02</td>
<td>-.09</td>
<td>-.17</td>
<td>-.12</td>
<td>-.06</td>
<td>-.05</td>
<td>-.04</td>
<td>-.01</td>
<td>.03</td>
<td>.66</td>
<td>1.00</td>
</tr>
</tbody>
</table>
REFERENCES


Brandon, P. R., Wang, Z., Chan, M. K., & Peecook, E. V. (1993) An evaluation of the implementation of the Special-Needs Schools Program in the Farrington-complex elementary schools in 1992-93. Honolulu: University of


