FIELD TEST EVALUATION
OF THE CHILD DEVELOPMENT PROJECT

Napa Valley Unified School District

2000-2002

Department of Research, Evaluation, and Organizational Learning
Developmental Studies Center
Oakland, CA

John Thomas, Ph.D.
Independent Evaluator
Mill Valley, CA

October 2002
# Table of Contents

Introduction ............................................................................................................. 1

Program Overview................................................................................................... 2

Evaluation Questions .............................................................................................. 3

Evaluation Design and Methodology ................................................................. 3
  Schools................................................................................................................. 3
  Data Collection .................................................................................................... 5
    Surveys ................................................................................................................ 5
    Individual Interviews .......................................................................................... 5
    Focus Groups ................................................................................................. 5
    Student Achievement Measures ...................................................................... 6

Findings .................................................................................................................... 8
  Phonics ................................................................................................................. 8
    Patterns of Use ............................................................................................... 8
    Managing Multiple Groups ............................................................................ 12
    Successes and Challenges .............................................................................. 13
    Staff Development ......................................................................................... 15

Independent Reading .......................................................................................... 16
  Patterns of Use ............................................................................................... 17
  Staff Development ......................................................................................... 18

Comprehension ................................................................................................. 18
  Patterns of Use ............................................................................................... 19
  Successes and Challenges .............................................................................. 19
  Staff Development ......................................................................................... 23

Program Integration ............................................................................................ 24

Community .......................................................................................................... 27
  Patterns of Use ............................................................................................... 27
  Successes and Challenges .............................................................................. 28
  Staff Development ......................................................................................... 30

Implementation Support .................................................................................... 31
  Role of Principal ............................................................................................... 31
  Role of Reading Specialist ............................................................................. 32
  Staff Development ......................................................................................... 32
Student Outcomes........................................................................................................36
Decoding Abilities ........................................................................................................36
Standardized Achievement Test Scores .................................................................38
District Literacy Assessments ..................................................................................39
Student Attitudes and Behavior ..............................................................................39

Summary of Findings and Conclusions ..................................................................42
Implementation of SIPPS and Independent Daily Reading .......................................42
Growth in Student Decoding Ability .........................................................................42
Making Meaning Implementation ............................................................................43
Growth in Student Comprehension Skills ...............................................................43
Caring School Community Implementation ............................................................44
Growth in Sense of community and Social/Ethical Development ..............................44
Other Academic Outcomes ......................................................................................45
Other Affective Outcomes .........................................................................................45
Program Integration ..................................................................................................45
Support for Implementation .....................................................................................46

Appendices
A: Evaluation Instruments
B: 2001-2002 Teacher Implementation Survey Findings
C: 2000-2001 Teacher Implementation Survey Findings
D: 2001-2002 Summer Institute Evaluation Findings
F: 1998-2002 Student Outcomes
INTRODUCTION

This report describes the Developmental Studies Center’s (DSC) two-year effort to evaluate the literacy components and some of the community components of its K-6 comprehensive school reform program, the Child Development Project™ (CDP). The evaluation was conducted at two schools in the Napa Valley School District—Bel Aire Park Elementary, a racially-mixed, low income school, and Mt. George Elementary, an economically mixed school located in a rural area. The program was evaluated in relationship to effects on students at two comparison schools, matched to the program schools on student demographic characteristics and prior academic achievement. The program’s implementation was assessed at the two program schools.

Baseline assessments of students’ social and ethical development were conducted in Spring, 2000 at the program and comparison schools, and baseline assessments of academic achievement on the Stanford Achievement Test version nine (SAT-9) were assessed for three years prior to introducing the program, Spring 1998 through Spring 2000. The phonics program, SIPPS™, was field tested at Napa for two years at the program schools, 2000-2001 and 2001-2002. Teachers received staff development in comprehension strategy instruction during 2000-2001, and tried using these strategies informally with their students during the school year. Lessons from DSC’s newly-developed comprehension program, Making Meaning™, were pilot tested at the program schools during 2001-2002. Elements of the Caring School Community™ program were introduced in 2000-2001, and the community-building aspects of Making Meaning were introduced during 2001-2002.

DSC conducted a three-day summer institute for the two program schools in August of 2000 and again in August of 2001. The institute in 2000 focused on introducing the program elements and providing initial instruction on SIPPS, community-building, and comprehension strategies. The focus of the institute in 2001 was on improving the implementation of SIPPS for the 2001-2002 year (based on feedback from the previous year’s evaluation efforts), continuing implementation of the community-building program elements, and preparing for implementation of the Making Meaning program. On-site staff development focused on SIPPS lessons and comprehension strategy instruction during the 2000-2001 school year, and on Making Meaning lessons during the 2001-2002 school year.

During both years, DSC staff demonstrated lessons, observed teachers at each grade level implementing lessons, and met with teachers in grade-level groups to discuss implementation issues and get ideas for improving the program. DSC provided approximately 35 days of on-site support each year to the program schools by means of twice-a-month visits that included classroom observation, individual coaching, and periodic after-school workshops.

This two-year evaluation was conducted by a combination of internal and external evaluation staff. DSC’s Department of Research, Evaluation, and Organizational Learning designed and conducted the evaluation activities for 2000-01. In March, 2002 DSC contracted with John Thomas, a northern California-based independent evaluator, to provide an external perspective and particular expertise in school reform. In this role, he reviewed the findings from 2000-01 and the evaluation plans for 2001-02, and worked with DSC’s program and evaluation staff to develop instruments, collect and analyze data, and write this report.
The remainder of this report provides an overview of CDP, the evaluation questions and methods, and findings in each of six main areas (phonics, comprehension, independent daily reading, community, administrative support, and student outcomes). It concludes with a summary of findings and conclusions.

**PROGRAM OVERVIEW**

The Child Development Project is a research-based, comprehensive, school change program that supports students' academic achievement and fosters their social and ethical development. The CDP program addresses the three most pressing needs for students: decoding skills, reading comprehension skills, and caring connections among students, teachers, and parents. The CDP program includes three components:

*SIPPS* (Systematic Instruction in Phoneme Awareness, Phonics, and Sight Words). The *SIPPS* program is designed for students in kindergarten through second grade, but has also proven effective for students in the third grade or beyond who are experiencing reading difficulty. The program consists of three levels of instruction—"Beginning," "Extension," and "Challenge"—that progress from simple alphabetic relationships (i.e., letter-sound correspondence) to more subtle and complex spelling patterns, and from single-syllable to polysyllabic words. This individualized and streamlined approach to phonics provides a systematic scope and sequence of instruction that is teacher directed and allows for flexible grouping of students depending on their progress.

Teachers at the Napa program schools implemented the *SIPPS* program during each of the two intervention years.

*Making Meaning* is a year-long program for students from kindergarten through sixth grade. The program consists of multi-week units that focus on a comprehension strategy such as questioning, making inferences, or summarizing. Read-alouds of award-winning literature, teacher-facilitated discussions, explicit teaching of strategies, and guided and independent practice in the use of strategies form the core of instruction. The pattern of instruction is consistent across grade levels, and the strategies build in complexity and depth through the grades. The program is intended to supplement or replace existing reading comprehension curricula. It is intended to be accessible to all students and includes extensive support for English Language learners.

Teachers at the Napa program schools received initial instruction in comprehension strategies during the 2000-01 school year, preparatory for the introduction of *Making Meaning* the following year. In 2001-2002 they piloted *Making Meaning* lessons.

Both *SIPPS* and *Making Meaning* include individual daily reading in appropriately leveled texts to build reading fluency, enhance vocabulary, and increase comprehension.

*Caring School Community*. The *Caring School Community* program is designed to help schools create a caring learning environment and build in all students sense of community and connection to the school. The complete program includes classroom, cross-grade, family involvement, and schoolwide components.
Only the classroom component of the program was implemented in the Napa program schools. This consists of class meetings to build unity, set classroom norms, plan and make decisions, and solve classroom problems. In addition, the norm-setting and cooperative learning aspects of Making Meaning contribute to establishing a sense of community in the classroom, and were implemented as part of the piloting of Making Meaning lessons during 2001-2002.

EVALUATION QUESTIONS

The two-year evaluation of the CDP program in Napa was guided by the following questions:

- What has been the extent and quality of CDP implementation (SIPPS, Making Meaning, Community, Independent Reading) in participating classrooms?
- What successes and challenges have teachers experienced with CDP?
- How do the phonics and comprehension components relate to the school's ongoing literacy program?
- How well do the phonics and comprehension programs integrate?
- What roles have the principals and reading specialists played in CDP implementation?
- How effective were the initial and ongoing staff development for helping teachers to implement the CDP program?
- How has the CDP program affected student growth in reading?
- How has the CDP program affected student social and ethical development?
  - How has the CDP program affected school- and reading-related attitudes?

EVALUATION DESIGN AND METHODOLOGY

The sections which follow describe the schools involved in the study and data collection (from students, teachers, principals, reading specialists, and DSC staff developer) designed to address the evaluation questions above.
Schools

Four schools in the Napa Unified School District participated in the evaluation—Bel Aire Park Elementary School and Mt. George Elementary School implemented the CDP program, and two matched elementary schools served as a comparison group. The schools serve grades kindergarten through six. Bel Aire Park is divided into two schools within one school, a traditional school and an alternative school. The total enrollment at Bel Aire Park is 397, with 120 students in the alternative track and 277 in the traditional track. Its comparison school is similarly divided into a traditional and an alternative school. The total enrollment is 328, with 117 in the alternative track and 211 in the traditional track. For each school, students in the alternative track are predominantly White and are higher achieving than students in the traditional track, a substantial number of whom are Hispanic. Mt. George has an enrollment of 222 and its comparison school has an enrollment of 286. Students at both of these schools are predominantly White and high achieving. Demographic information for students in the four schools is presented in Table 1.

Table 1

Demographic Breakdown of Students at the Participating Napa schools: 2001-2002

<table>
<thead>
<tr>
<th>Student Demographics</th>
<th>Bel Aire Park</th>
<th>Comparison</th>
<th>Mt. George</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alternative</td>
<td>Traditional</td>
<td>Alternative</td>
<td>Traditional</td>
</tr>
<tr>
<td>American Indian/Alaskan</td>
<td>1%</td>
<td>9%</td>
<td>1%</td>
<td>7%</td>
</tr>
<tr>
<td>Asian</td>
<td>1%</td>
<td>1%</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td>African American</td>
<td>1%</td>
<td>1%</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>Filipino</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>White</td>
<td>88%</td>
<td>46%</td>
<td>85%</td>
<td>30%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3%</td>
<td>42%</td>
<td>9%</td>
<td>59%</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

As shown in Table 2, students at Mt. George, in the alternative track at Bel Aire Park, and in the alternative track of Bel Aire Park’s comparison school are high achieving, ranking at or near the top of the distribution of California elementary schools on the state’s Academic Performance Index. Students in the traditional track at Bel Aire Park and their comparison group are average in academic achievement.

Table 2

Academic Performance Scores for the Participating Napa Schools

<table>
<thead>
<tr>
<th>Academic Performance</th>
<th>Bel Aire Park</th>
<th>Comparison</th>
<th>Mt. George</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alternative</td>
<td>Traditional</td>
<td>Alternative</td>
<td>Traditional</td>
</tr>
<tr>
<td>2001 Academic Performance Index (API)</td>
<td>872</td>
<td>660</td>
<td>846</td>
<td>686</td>
</tr>
<tr>
<td>2001 Statewide Rank</td>
<td>10</td>
<td>5</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>
Data Collection

In all, as described below, data collection for this two-year evaluation included student and teacher surveys, principal and staff developer interviews, teacher focus groups, and student achievement measures. Copies of the survey instruments are included in Appendix A.

Student Survey

The student survey was administered to 3rd through 6th grade students at the program and comparison schools in the Spring of the baseline year (2000) and each of the two intervention years (2001, 2002). The survey assessed students' school-related attitudes (i.e., sense of school community, liking for school), reading-related attitudes (i.e., liking for reading, liking for read-alouds, sense of efficacy as a reader, and amount of reading for pleasure), and social/ethical attitudes and behaviors (i.e., concern for others, commitment to democratic values, conflict resolution skills, and altruistic behavior).

Summer Institute Evaluation

At the close of each summer institute participants were asked to provide feedback on the institute via a written survey. It included questions on the quality of their staff development experience, the extent to which each component of the institute contributed to their learning and understanding, and the motivation and preparation to implement the CDP program.

Teacher Implementation Survey

Program teachers completed a survey on instruction practices and staff development in phonics, comprehension, and community in the Spring of each of the two intervention years (2000-01 and 2001-02). With the exception of new questions added to the 2002 survey in order to evaluate the Making Meaning lessons and staff development related to the Making Meaning program, the two surveys asked the same questions in the same way both years. The surveys included both fixed-response and open-ended questions. Response rates for the surveys each year were as follows: Bel Aire Park (100%, 2000-01; 100%, 2001-02) and Mt. George (100%, 2000-01; 83%, 2001-02).

Principal Interview

The principal interview was an approximately one-hour session held with each of the two principals of the program schools each Spring. It included questions on implementation of the CDP components, staff development, the role of the principal and reading specialist, and integration of the components.

Staff Developer Interview

The staff developer who provided ongoing staff developer at both schools each Spring was interviewed for approximately one hour. The interview included questions relating to implementation of each of the CDP program components, staff development activities and effectiveness, the role of the principal and reading specialist, and integration of the components.
Teacher Focus Groups

*SIPPS teachers:* The *SIPPS* focus groups, conducted in Spring 2002, involved four teachers at Mt. George and three teachers at Bel Aire Park. These focus groups addressed how teachers have incorporated *SIPPS* into their overall literacy program and what is challenging with respect to the implementation of *SIPPS*.

*Teachers of both *SIPPS* and Making Meaning:* These focus group involved three teachers at Mt. George and six teachers at Bel Aire Park, also in Spring 2002. It focused on the challenges involved in implementing both *SIPPS* and *Making Meaning* at the same time, and on advice these teachers would give another teacher about using the two programs.

Student Achievement Measures

*Decoding:* Slosson Oral Reading Test (SORT). In order to evaluate the effectiveness of the *SIPPS* phonics program, the Slosson Oral Reading Test of decoding ability was administered as a pretest to students in grades 1-3 in the fall of 2001, and to kindergarten students in the winter of 2002 at both the program and comparison schools, prior to the beginning of phonics instruction. The SORT was administered as a posttest to K-3rd grade students in the Spring of 2002, after approximately seven months of phonics instruction for 1st through 3rd grade students, and approximately three months of phonics instruction for kindergarten students.

*Reading:* SAT-9 Vocabulary and Comprehension. Scores on the state-mandated SAT-9 academic achievement test for students in grades 2-6 were examined from Spring 1998 (three years prior to program implementation) through Spring 2002 (after two years of program implementation at the two program schools) in order to assess program effects on vocabulary and reading comprehension.

*Reading:* District Literacy Assessments. In addition to the SAT-9 data, scores on district assessments of literacy administered to grades 3 though 6 were examined from Spring 2000 (baseline) through Spring 2002 (after two years of implementation at the two program schools) at the program and comparison schools to assess program effects on literacy.

Student academic achievement was examined using scores from the California statewide testing system (*SAT-9*, from 1998 through 2002 and California Standards Test in Language Arts in Spring 2002) and district- and state-administered language arts tests (Spring 2000 through Spring 2002).

A summary of the data collection activities for the two-year evaluation is shown in Table 3.
<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring 1998</td>
<td>Baseline assessment of students’ academic achievement (SAT-9, grades 2-6)</td>
</tr>
<tr>
<td>Spring 1999</td>
<td>Baseline assessment of students’ academic achievement (SAT-9, grades 2-6)</td>
</tr>
<tr>
<td>Spring 2000</td>
<td>Baseline assessments of students’ academic achievement [SAT-9 (grades 2-6) and District Literacy Assessments (grades 3-6)] and school-related and social/ethical attitudes and behaviors (grades 3-6)</td>
</tr>
<tr>
<td>Summer 2000</td>
<td>Evaluation of summer institute</td>
</tr>
<tr>
<td>Spring 2001</td>
<td>Teacher implementation survey  [Principal interview  [Staff developer interview</td>
</tr>
<tr>
<td>Spring 2001</td>
<td>Assessment of students’ academic achievement [SAT-9 (grades 2-6) and District Literacy Assessments (grades 3-6)] and school-related and social/ethical attitudes and behaviors (grades 3-6)</td>
</tr>
<tr>
<td>Summer 2001</td>
<td>Evaluation of summer institute</td>
</tr>
<tr>
<td>Fall 2001</td>
<td>Baseline assessment of students’ decoding ability (Slosson Oral Reading Test, grades 1-3)</td>
</tr>
<tr>
<td>Winter 2002</td>
<td>Baseline assessment of students’ decoding ability (Slosson Oral Reading Test, grade K)</td>
</tr>
<tr>
<td>Spring 2002</td>
<td>Teacher implementation survey  [Principal interview  [Staff developer interview  [Teacher focus groups</td>
</tr>
<tr>
<td>Spring 2002</td>
<td>Assessment of students’ academic achievement [SAT-9 (grades 2-6); District Literacy Assessments (grades 3-6); Slosson Oral Reading Test (grades K-3)] and school-related and social/ethical attitudes and behaviors (grades 3-6)</td>
</tr>
</tbody>
</table>
FINDINGS

The sections which follow present the findings of this evaluation of CDP related to each of the following: phonics, independent reading, comprehension, program integration, community, implementation support, and student outcomes. Where appropriate, to aid understanding and interpretation of the findings, each section begins with further details on the component. Within sections, findings are presented in a question-and-answer format. Graphics are included to highlight the most noteworthy findings. In some cases, these explanations were developed with further input from the reading specialist at one of the schools (Bel Aire Park), and the two DSC staff developers who worked with the Napa schools.

Phonics

SIPPS is an individualized approach to phonics instruction. The program that recognizes that within any classroom, students will differ greatly in their decoding skills, and thus different students need instruction at different levels of decoding. SIPPS is intended primarily as a phonics program for beginning readers at the primary grades, but also can be used effectively as an intervention program for upper-grade students who are struggling with decoding. For daily phonics instruction, students are assigned to groups that are relatively homogeneous in decoding abilities, based on teacher administered assessments. These groups can be formed within or across classrooms. Students are reassessed and regrouped periodically during the school year to keep instruction at the optimal level for each student.

Patterns of Use

What percentage of teachers have students who receive SIPPS instruction?

Finding: During 2001-02 all primary teachers and almost half of upper-grade teachers had students who received SIPPS instruction. As shown below, there was a slight drop-off from the previous year in the upper grades, due primarily to teachers at Bel Aire Park.
2001-2002: At both schools, all primary teachers, 50% of upper-grade teachers at Mt. George, and 38% of upper-grade teachers at Bel Aire Park had students who received SIPPS instruction. (See Appendix B, p. 4.)

2000-2001: At both schools, all primary and over 50% of upper-grade teachers had students who received SIPPS instruction. (See Appendix C, p. 7.)

The primary reason for the drop-off was that upper-grade teachers said that their students no longer needed SIPPS. This is to be expected because most students at the Napa program schools who needed instruction in phonics received it during 2000-2001, with upper-grade phonics students completing SIPPS challenge level. In addition, the Bel Aire Park reading specialist reported that some 6th grade teachers whose students were still below grade-level in reading did not use SIPPS because they felt that the manner of instruction (e.g., use of flash cards) was developmentally inappropriate.

How frequent are SIPPS sessions?

**Finding:** In 2001-02, on average, teachers implemented SIPPS four times a week for 20-35 minutes. This represents a slight increase in the frequency of SIPPS sessions compared to the previous year.

2001-2002: On average, teachers at both schools implemented SIPPS four times a week for 20 minutes (challenge), 25 minutes (beginning), or 35 minutes (extension). (See Appendix D, pp. 5-8.)

2000-2001: On average, teachers at both schools implemented SIPPS three to four times a week for 20 minutes (challenge), 30 minutes (beginning), or 40 minutes (extension). (See Appendix C, pp. 8-11.)

The teachers were in their second year of using SIPPS in 2001-2002. They had mastered the routines and structure of the program and had seen that the program was effective, yielding greater fidelity in implementation. In addition, the increased frequency of instruction may be due to changes in teachers or teacher acceptance. According to the Bel Aire Park reading specialist some 3rd grade teachers who had been resistant to the program during 2000-2001 did SIPPS with greater fidelity during 2001-2002; other teachers left and the new teachers did SIPPS; and in one case, a teacher was pressured by other teachers to use the program.

What SIPPS levels are being implemented and how many groups are there per class at each level?

**Finding:** In line with the results from the previous year, all levels were implemented at both schools. In 2001-02 however, teachers formed fewer groups per level compared to the previous year.
2001-2002: All three SIPPS levels were being implemented at both schools, with an average of 1-2 groups per level. (See Appendix B, pp. 5-8.)

2000-2001: All three SIPPS levels were implemented at both schools, with more beginning level groups implemented at Bel Aire Park. At beginning and extension levels, most classrooms have 2-3 groups. At the challenge level, most classrooms have one group. (Appendix C, pp. 8-11.)

Additional input from the reading specialist at Bel Aire Park suggests some reasons for this reduction in the number of groups per level. At Bel Aire Park primary teachers did more sharing of students for SIPPS instruction during 2001-2002, especially with the lowest and highest readers in their classes. This yielded fewer groups at each level in any one classroom. Also, the reading specialist indicated that teachers had a wider range of abilities in their classes, resulting in more SIPPS levels per classroom (e.g., a beginning, extension, and challenge group), but with fewer groups needed at each level.

Did teachers make modifications to the SIPPS lessons and, if so, what kinds?

**Finding:** Over one-half of Bel Aire Park teachers reported making modifications, which was about the same percentage as last year. At Mt. George, the percentage of teachers making modifications dropped from 64% to 29%. At both schools, the most common type of modification was to introduce supplemental activities.

![Percentage of Teachers at Mt. George and Bel Aire Park Making Modifications to SIPPS Lessons, 2000-01 and 2001-02](chart)

2001-2002: Over half (59%) of Bel Aire Park teachers and nearly one-third (29%) of Mt. George teachers made modifications to SIPPS lessons. Ways in which teachers modified the lessons included (the frequency of each response is shown in parentheses): (See Appendix B, p. 10.)
Major Theme
• Introduced additional activities and/or personalized SIPPS for motivation purposes: new words, partner reading, sight word sound cards, dictionary, cut-up sentences, games (9)

Minor Themes
• Dropped aspects of SIPPS: omit spelling and/or do spelling as part of workshop (2)
• Break lessons into smaller parts, interspersed throughout the day, just do half (2)
• Introduced other language arts components: guided spelling, comprehension (2)
• Modified higher extension lessons to make them easier (1)

2000-2001: Over half (57%) of Bel Aire Park teachers and almost all (83%) of Mt. George teachers made modifications to the SIPPS lessons. (See Appendix C, p. 13.)

The decline in teacher modifications to the lessons at Mt. George may be due to the principal's instructions. According to teachers participating in the SIPPS focus group, Mt. George teachers in 2001-2002 were instructed by the principal to do SIPPS "by the book" because of the evaluation being conducted. (Focus Group)

What percentage of SIPPS teachers used multiple groups?

**Finding:** Most teachers taught multiple SIPPS groups in 2001-02. Overall, there was a 23% increase, compared to last year, in the percentage of teachers who use multiple groups, due primarily to Bel Aire Park.
2001-2002: Almost 70% of SIPPS teachers at both schools taught multiple SIPPS groups in their classrooms. (See Appendix B, p. 9.)

2000-2001: Just over one-half of teachers at both schools taught multiple SIPPS groups in their classroom. (See Appendix C, p. 12.)

According to the Bel Aire Park reading specialist, primary grade teachers at her school shared small numbers of their lowest and highest readers, but generally had a wider range of abilities in their classes during 2001-2002, yielding multiple levels of SIPPS groups in one classroom (e.g., beginning, extension, and challenge level groups).

Managing Multiple Groups

To what extent do teachers share instruction of SIPPS groups?

Finding: In 2001-02, over one-half of teachers with beginning and extension level groups provided SIPPS instruction with another teacher; nearly all challenge groups received instruction by one teacher alone. Compared to last year, there was a reduction in the sharing of instruction for challenge groups.

2001-2002: Over half (60%) teachers reported that both they and another teacher provided SIPPS instruction to their students at the beginning and extension levels. At the challenge level, nearly all (92%) of teachers reported that either they or another teacher exclusively provided SIPPS instruction. (See Appendix B, pp. 5-8.)

2000-2001: With the exception of one teacher at Mt. George, all teachers reported that both they and another teacher provided SIPPS instruction to their students. (See Appendix C, pp. 8-11.)

The reduced sharing of instruction at the challenge level was due to both an increase in whole-class or large-group challenge-level instruction and to classes where students no longer needed phonics instruction. In addition, at Bel Aire Park, the two 1st grade teachers reconfigured their classroom such that one 1st grade had all beginning level students and the other had all extension level students. This meant these teachers had no need to share instruction of SIPPS.

How difficult was it to manage multiple groups?

Finding: Teachers find managing multiple groups to be moderately difficult. These findings are consistent with last year’s findings. The most common difficulty was keeping non-SIPPS students busy during SIPPS instruction.

2001-2002: Teachers at both schools found managing multiple SIPPS groups to be moderately difficult, with Mt. George teachers finding it slightly more difficult (3.7 on a scale of 1 to 5) than Bel Aire Park teachers (3.1). Teachers described the following kinds of difficulties: (See Appendix B, p. 9.)
Major Theme
- Classroom management: trying to keep students working who are not in a SIPPS group (11)

Minor Themes
- Monitoring students’ work, giving help to all students (3)
- Not having sufficient materials: need two manuals, charts for wall (2)
- Dealing with student behavior problems (1)
- Pace: not going too fast (1)
- Takes time away from guided reading (1)

Teachers in the 2001-02 SIPPS focus groups explained that while it is difficult to manage multiple groups, they are also concerned about losing some students when you try to ease management problems by using larger groups. Exchanging students between teachers can help, but the teachers then don’t know if students are getting what they need. Additionally, managing multiple groups is most difficult in multi-grade classrooms and in kindergarten. (Teacher Focus Groups)

2000-2001: Teachers at both schools found managing multiple SIPPS groups to be moderately difficult, with Mt. George teachers finding it slightly more difficult (3.5 on a scale of 1 to 5) than Bel Aire Park teachers (3.1). (See Appendix C, pp. 12-13.)

Managing multiple groups is an inherent difficulty with the program, because of the individualized nature of its instruction.

Successes and Challenges

What were the major successes associated with implementing SIPPS?

Finding: For teachers in 2001-02, observations of academic improvement continued to be the most popular indicator of SIPPS success.

2001-2002: Teachers at both schools reported successes with SIPPS:

Major Themes
- General positive effect on student reading competence: instant success, kids can transition to regular text, kids are using rules, kids are reading, writing, spelling (15)
- Acquisition of knowledge: phonics, segmentation, learning new words, “syllabification” (7)

Minor Themes
- It’s easy to teach (1)
- Kids love the stories (1)
- Kids like to work in small groups (1)
- Working together as a whole class (1)
Finding: Both principals and teachers report that SIPPS now runs smoothly.

2001-2002:
• SIPPS now stands alone. This year, everyone felt more comfortable with it.
• Some teachers are not used to scripted instruction, but they are seeing the response in terms of how well students are doing.
• Teachers are doing SIPPS in summer school. The API went up 52 points last year and people started to look at SIPPS.
• SIPPS will continue at both schools despite a district-wide adoption of a basal program.

2000-2001:
• SIPPS is widely regarded to be effective and exceeded expectations.
• Getting SIPPS started took a lot of effort.
• Time required is an issue, but teachers have followed SIPPS lessons as written and have tended to drop other parts of the curriculum for SIPPS.
• Multiple groups was problematic initially, but there has been much progress toward resolution (exchanging students, using reading specialists to provide instruction, etc.).

In addition, during the 2002 interviews teachers and principals identified the following signs of success:

• The Second Year Effect: SIPPS seems to run smoothly now that teachers have had a year to work things out. (Teacher Focus Groups, Principal Interviews)

• Generalizability of Methods: SIPPS techniques can be used whenever a new word pops up in subject matter areas. (Teacher Focus Groups)

Not surprisingly, in the second year of implementation, teachers were more familiar and confident with the organization, content, and routines of SIPPS and were better able to manage multiple groups of students. Also, teachers were convinced of the effectiveness of SIPPS.

What were the major challenges associated with implementing SIPPS?

Finding: Having sufficient time and classroom/group management are the predominate issues for implementing SIPPS.

2001-2002: Teachers reported several challenges with SIPPS:

Major Themes
• Managing different levels, meeting everyone’s needs, handling different learning styles (7)
• Time: scheduling, keeping time, finding time, keeping up (6)
• Managing the classroom: managing students while trying to teach, keeping other students engaged (5)
Minor Themes
- Keeping it interesting, maintaining enthusiasm (3)
- Lack of application to other reading (1)
- Lack of spelling component (1)
- Transition from beginning to extension (1)

Teachers in focus groups and the Staff Developer also report challenges related to time and to implementing the extension level:

- **Time Commitment:** SIPPS represents a large portion of the language arts curriculum. Other practices have suffered. A lesson a day may be too much for some children (Teacher Focus Groups). Preparing for lessons continues to be challenging and preplanning can be time consuming. (Teacher Focus Groups, Staff Developer Interview)

- **Extension:** At a certain point in the lessons (e.g., lesson 6, adding a consonant, long vowels), the lessons become too difficult for some students; it is not clear how to proceed. (Teacher Focus Groups)

2000-2001: Teachers reported several challenges with SIPPS, in decreasing order of frequency:

- Time required
- Learning the program and getting SIPPS started
- Engaging other students in meaningful learning activities
- SIPPS is boring for students, teachers

It is unlikely that issues of time and management will ever be completely eliminated, but it is clear from teacher and principal reports that they become considerably less problematic by the second year of using the program. In the second year, both the teachers and the students know the routines, there is more sharing of students for instruction, and teachers often have other students do independent reading while the teacher is working with a SIPPS group.

**Staff Development**

How confident were teachers following the Summer Institute concerning their use of SIPPS?

**Finding:** Teachers at both schools felt prepared and motivated to use SIPPS at the beginning of the 2001-02 school year, a moderate to substantial increase from last year on both ratings.
2001-2002:

- Following the Summer Institute, teachers at both schools felt prepared (4.0) and motivated (4.3) to use SIPPS.

```
<table>
<thead>
<tr>
<th>Items</th>
<th>1 not at all</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 to a great extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mt. George</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepared to Implement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIPPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivated to Implement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIPPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bel Aire Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepared to Implement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIPPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivated to Implement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIPPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

- K-1 teachers at the Summer Institute rated the component SIPPS—Issues and Questions, highest in terms of helping their learning and understanding (4.9); grades 2-6 teachers rated this component lower (3.5) (See Appendix D, p. 4.)

2000-2001:

- Teachers at both schools felt that they understood SIPPS to a moderate extent (3.8), were only somewhat motivated to use SIPPS (3.2), and were not particularly well prepared to do so (2.6).
- Mt. George ratings for these items were slightly higher than Bel Aire Park ratings. (See Appendix E, p. 1.)

These findings about motivation and preparation to do SIPPS by the end of the second summer institute further support the second-year effect described above. Teachers had seen such success with the program during the first year that they has mastered its routines and wanted to continue using it.

**Independent Reading**

Daily independent reading is an important component of both SIPPS and Making Meaning. Independent reading is essential to students becoming fluent readers and to increasing vocabulary. Independent reading also can provide opportunities for student practice using comprehension strategies and for teachers to monitor their rate, accuracy and comprehension. It is important to note that the Napa teachers were using Guided Reading prior to the introduction
of CDP, and were not willing to do to Independent Daily Reading in place of it. As a result, most did not use this approach, an integral part of SIPPS and Making Meaning, where students read for 20-30 minutes each day in leveled texts (i.e., books that are neither too easy nor to difficult) and the teacher actively monitors students' reading rate, accuracy, and comprehension. Findings about the amount of some type of independent reading students did and about teachers’ use of leveled books are presented below.

Patterns Of Use

How extensive was the implementation of some type of independent reading?

Finding: In 2001-02 almost all teachers reported having students do independent reading of some kind (whisper reading, sustained silent reading, reading aloud to teacher, book clubs, reading with a partner) on a daily basis. Compared to the previous year, there was a slight increase in average minutes devoted to independent reading.

2001-2002: Virtually all teachers (92%) at both schools reported having students do independent reading (of some kind) on a daily basis, typically for 20-25 minutes. (See Appendix B, p. 23.)

2000-2001: Virtually all teachers (97%) at both schools reported having students do independent reading (of some kind) on a daily basis, typically for 15-25 minutes. (See Appendix C, p. 14.)

Related to this finding, the reading specialist from Bel Aire Park noted that teachers often had other students doing independent reading while some students were doing SIPPS.

How extensive was the use of leveled books?

Finding: Consistent with the previous year, in 2001-02, about three-fourths of teachers reported using leveled books selected by both teacher and students, with a sizeable minority reporting that students select their own books.

2001-2002: Most teachers (74%) reported using leveled books that were jointly selected by teachers and students, but a sizeable minority only did student-selected “free reading.” (See Appendix B, p. 23.)

2000-2001: Most teachers (73%) reported using leveled books that were jointly selected by teachers and students, but a sizeable minority only did student-selected “free reading.” (See Appendix C, p. 14.)

The reading specialist from Bel Aire Park indicated that teachers worked hard during 2001-2002 on leveling books for independent reading. Although the use of leveled texts is consistent with the program, it is important to keep in mind that
most teachers were not using the program's approach to Independent Daily Reading.

**Staff Development**

Following the Summer Institute, how confident were teachers about using independent reading?

**Finding:** Following the Summer Institute for 2000-2001, on average, teachers at Bel Aire Park, in particular, indicated that they had a moderate understanding of the program's approach to Independent Daily Reading. Teachers at both schools felt moderately prepared and motivated to use Independent Daily Reading at the beginning of the 2001-02 school year.

**2001-2002:**
- Teachers at the Summer Institute rated Independent Reading among the lowest of the institute components in terms of helping their learning and understanding (3.1). (See Appendix D, p. 4.)
- Following the Summer Institute, teachers at both schools felt moderately prepared (3.8) and motivated (4.1) to use Independent Reading.

**2000-2001:** Teachers at Mt. George understood Independent Reading to a moderate extent (3.6); Bel Aire Park teachers reported that they understood Independent Reading to a lesser extent (3.1) (See Appendix E, p. 1.)

The increase observed in motivation and preparation to use Independent Reading is understandable since this was the second time the strategy was covered in summer institutes. Teachers' moderate response regarding the extent to which the Independent Daily Reading component of the summer institute helped their learning and understanding likely reflects the conflict teachers acknowledged at the outset of the program (between their desire to continue using Guided Reading and DSC's emphasis on Independent Daily Reading). Again, both this conflict and lack of understanding/learning accounts for their not using all aspects of Independent Daily Reading as prescribed by the program.

**Comprehension**

The *Making Meaning* program consists of week-long units composed of multi-day reading comprehension lessons for students in grades K-6. The lessons include teacher read-alouds so that all students can access content regardless of reading ability, teacher-facilitated discussions of books and stories, explicit teaching of comprehension strategies that build in depth and complexity from lower to upper grades, and both guided and independent practice in using the strategies. Teachers at the two Napa schools pilot-tested draft *Making Meaning* lessons as they were available during the 2001-2002 school year.
Patterns Of Use

How extensive was the implementation of *Making Meaning*?

**Finding:** Although almost all teachers taught *Making Meaning* lessons, relatively few of the units at any grade level were taught by all of the teachers, and relatively few teachers taught all days of the weekly units.

2001-2002:
- Most upper and lower grade teachers at Bel Aire Park (88%) used *Making Meaning*.
- Only 1 in 5 specialty teachers at Bel Aire Park (20%) used *Making Meaning*.
- All teachers at Mt. George (100%) used *Making Meaning*.
- The reasons given for not using *Making Meaning* were not having time in the school day, or not being able to fit *Making Meaning* into the curriculum. (See Appendix B, p. 12.)

According to the DSC staff developers, the teachers at both schools were given only a few *Making Meaning* lessons at the beginning of the year. Because they were accustomed to planning out their year in September and because the remainder of the lessons were delivered in draft form and in batches, over the course of the year, *Making Meaning* teachers never had the opportunity to plan or prepare for the majority of lessons. Additionally, there was no administrative mandate at these schools to allot a block of time for *Making Meaning* lessons. As a consequence, teachers did not follow a systematic implementation schedule, nor did they tend to implement units or lessons in their entirety.

Successes and Challenges

What were the successes associated with the implementation of *Making Meaning*?

**Finding:** For teachers, observation of student-to-student communication and the successful teaching of strategies were popular signs of the success of *Making Meaning*. Principals and focus group teachers pointed to the value of the *Making Meaning* framework as well.

2001-2002: Signs of success listed by teachers included the following:

**Major Themes**
- *Instructional techniques and strategies*: stories, thoughtfulness, building ways to talk about books, "think-pair-share," "turn to your partner" (14)
- *Children's responses to the program*: kids enjoyed it, were engaged, loved the books, latched on to strategies, transferred strategies to other reading; kids were able to listen to one another, take part in discussions (11)
Minor Themes
- Emphasis on expository text (2)
- Grade level planning (2)
- Classroom community (2)

Teacher Focus Groups and Principal Interviews also yielded signs of success:
- Generalizability of Methods: The strategies and skills can be applied to other subject matter areas by students and teachers; to some extent this is happening spontaneously. (Teacher Focus Groups, Principal Interviews)
- Benefits of the Program Structure and Vocabulary: The program provides an organizing framework and vocabulary that is useful for children's socialization and community building and for giving teachers a common language to talk about instruction and a way to teach comprehension skills directly. (Principal Interviews)

What were the specific student benefits of Making Meaning as rated by participating teachers?

**Finding:** Growth in partner work and sense of community were the most highly rated benefits attributable to Making Meaning for both ELL and English-speaking students.

![Teachers' Mean Ratings of Extent of ELL and English Speaking Students' Growth in Particular Areas Attributable to Making Meaning](chart)

<table>
<thead>
<tr>
<th>Area of Growth</th>
<th>1 not at all</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 to a great extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner Work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sense of Community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discussion Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2001-2002:
The highest rated benefits for ELL students (both lower and upper grades) were:
- Partner work (4.28)
- Sense of classroom community (4.00)
- More caring attitude on the part of students (3.88)
- Discussion skills (3.72)

The highest rated benefits for English-speaking students (both lower and upper grades) were:
- Partner work (4.34)
- Discussion skills (4.24)
- Motivation to read (4.10)
- Sense of classroom community (4.03) (See Appendix B, pp. 14-21.)

This finding suggests that the impact of the program is not confined to one demographic group. It is also interesting that growth in discussion skills was seen as an outgrowth of the program for ELL students as well as for English-speaking students. According to the staff developers, teachers reported that ELL students took part in program-related classroom discussions to a significant extent.

What were the major challenges associated with the implementation of *Making Meaning* this year?

**Finding:** Finding time for the program and fitting it into the curriculum were the salient issues for teachers. Staff development/teacher support and assessment difficulties were stressed by principals and focus group teachers.

2001-2002: Challenges listed by teachers on their end-of-year survey included the following:

**Major Theme**
- *Time*: finding time, finding consistent time, time to plan, time to get to know materials, length of lessons (15)
- *Incorporating it into the program* with other subjects; doesn’t match thematic sequence/other lessons in the curriculum (9)

**Minor Theme**
- *Getting past superficial level of engagement on the part of students*, getting students to share (3)
- *Logistics*: getting substitutes, not knowing when materials were arriving (3)
- *Keeping ELL learners engaged* (2)
- *Feeling as though I had to teach all the lessons from one book* (1)

Teachers in the focus groups, the staff developer, and principals also mentioned sources of challenges:
• **Time Commitment**: Making Meaning poses a substantial time commitment, yet it still does not fulfill the language arts requirement; some teachers ended up not putting much time into Making Meaning. (Teacher Focus Groups, Staff Developer Interview)

• **Support Requirements**: Making Meaning needs to be supported in order to be successful and teachers need to understand and accept the constructivist model of teaching and that reading does not teach comprehension. New teachers need modeling. (Teacher Focus Groups, Principal Interviews, Staff Developer Interview)

• **Assessment**: Assessments were hard to do, logistically, in groups or partners; it is difficult to observe comprehension, especially with students who don’t participate in discussions; most teachers did not pick up on the assessment suggestions or did not try; nevertheless, teachers want to be able to produce student grades. (Teacher Focus Groups, Staff Developer Notes, Staff Developer Interview)

• **Student Engagement**: Teachers were challenged by trying to get all students to be more active, share their thinking, “dig deeper” in discussions. (Staff Developer Notes)

• **Student Partnerships**: Teachers were concerned about the variation in effectiveness across partnerships, off task behavior in partnerships, the need for additional partnership skill training, and the lack of guidance in the program about how and when to change partners. (Teacher Focus Groups, Staff Developer Notes)

The difficulties associated with finding time for the program and fitting it into the curriculum can be viewed as a direct result of the conditions of the field test. First, teachers were not given an overall plan for integrating the program within the larger curriculum. Second, they were not given a mandate by administrators or by CDP to devote a particular block of time to the program. And third, since they received the majority of lessons sporadically, in draft form, throughout the year, they probably did not give the lessons the planning time and priority they might have given them under different conditions.

What percentage of teachers modified *Making Meaning* and what were the most common modifications?

<table>
<thead>
<tr>
<th>Finding:</th>
<th>Most teachers modified the <em>Making Meaning</em> lessons, either by shortening or cutting out program parts.</th>
</tr>
</thead>
</table>

*2001-2002*: The majority of teachers (74%) reported that they modified the *Making Meaning* lessons. These modifications included: (See Appendix A, p. 13)
Major Themes
- Shortened lessons: already covered some things so left out some activities, seemed too redundant, skipped things students have mastered (11)
- Added things to the lessons: added stop and jot before sharing, used other books, added other strategies, vocabulary (10)
- Dropped components: IDR, community building, discussion of books (5)

Minor Theme
- Made modifications based on the time of the year and curricular themes (2)

The Making Meaning lessons have a certain amount of built-in redundancy. Whereas in other pilot tests of the program, teachers began to see the value of this redundancy in terms of student learning, in an implementation such as this one where instructional time was an issue, teachers may have seen redundancy as an opportunity to save time by dropping lessons or parts of lessons.

Staff Development

How confident were teachers about using Making Meaning at the end of the Summer Institute?

Finding: Teachers at both schools felt prepared and highly motivated to use comprehension lessons; furthermore, they said that they came away with ideas and strategies that would help them in their teaching.

2001-2002:
- Following the Summer Institute, teachers at both schools felt prepared (4.0) and highly motivated (4.5) to use comprehension lessons.
- Following the Summer Institute, teachers at both schools agreed that they had learned ideas and strategies to help in their teaching of reading (4.0)
- Grades 2-6 teachers at the Summer Institute gave an overall high rating to the component Comprehension Lesson Demonstration and Discussion highly in terms of helping their learning and understanding (4.3); K-1 teachers rated this component even higher (4.7). (See Appendix D, p. 4.)

The high level of motivation and feelings of preparedness on the part of teachers was probably due to the amount of guidance teachers received during the Institute for administering the first batch of Making Meaning lessons.
Program Integration

How effective do teachers view the combination of SIPPS and Making Meaning?

Finding: SIPPS and Making Meaning are largely seen as separate programs with different purposes and approaches. At the same time, almost all teachers indicated that the combination of the two programs was effective.

2001-2002:
Teachers rated the effectiveness of the SIPPS and Making Meaning programs quite highly and made comments like the following:

Major Theme
- Great match, good combination (25)

Minor Theme
- Grammar and writing are missing, it is difficult to fit in IDR and guided reading (1)
- SIPPS could be shorter (1)
- Making Meaning is difficult for younger students (1)

In addition, principals, and the staff developer made the following comments about the integration of SIPPS and Making Meaning:

- From an administration view, SIPPS leaves nothing to figure out. If you make a commitment, it happens, even with teachers who are hesitant or who have never taught reading before. With regard to Making Meaning, I think it definitely needs to be supported in order to be successful. (Mt. George Principal Interview)

- SIPPS is very effective for 1st, 2nd, and 3rd grade students. For kindergartners, they have a short day to start with, so SIPPS groups are a challenge. Making Meaning is very powerful for 3rd, 4th, 5th, and 6th grade students. It fits in with previous packages [at the school]. (Principal Interview)

- There was not enough time for full implementation of Making Meaning this year. Because Making Meaning was not regimented [like SIPPS], it got shoved aside. (Staff Developer Interview)

- In lower grades, K-1, there was less emphasis on reading comprehension and more emphasis on fluency, so SIPPS and Making Meaning seemed separate and not blended. As you go up in grades, you didn't see as much conflict. (Staff Developer Interview)

Staff development for the two programs did not include any discussion of philosophical differences in the programs or how the two programs could be or should be implemented. It is thus interesting that the teachers indicated that the programs are a "good match." It is likely that this finding is indicative of teachers’ beliefs that there is a need for both types of programs.
How was SIPPS integrated into the Language Arts block?

**Finding:** The most common method of integration (at Bel Aire Park only) appeared to be doing SIPPS with one group while doing some other language arts activity with another.

2001-2002: Teachers reported numerous integration strategies. They are roughly categorized as follows:

**Major Themes**
- *Simultaneous:* Do SIPPS with one group while doing other language arts activities with another group (Bel Aire Park only) (9)
- *Complete integration:* SIPPS is not separate, integrated into everything, SIPPS is language arts (7)
- *Separate block:* start day with SIPPS for one hour then do other language arts activities (5)

**Minor Themes**
- SIPPS students are pulled out of the classroom (2)
- Stress spelling/sight words (1)
- "A version of SIPPS" (1)

Responses to this item represent 25 of the 29 teachers who used SIPPS. However, teachers responded to this item in different ways. Some addressed the question of when SIPPS was used. Some indicated how they integrated SIPPS with other components of language arts. Others addressed the question of how central SIPPS has become in their language arts program. Thus, the results are only able to be roughly categorized as shown above, and remain somewhat difficult to interpret. Presumably, some of the teachers who indicated that they devoted a separate block of time to SIPPS could have administered SIPPS to one group while another group engaged in some other language arts activity. Some of the teachers who indicated that SIPPS was their entire language arts program may not have had time to do other activities, but this is not necessarily the case for all of these respondents. Accordingly, more detailed inquiry on how teachers integrate SIPPS into their language arts block is needed.

How did teachers integrate Making Meaning into the Language Arts block?

**Finding:** More than one-third of teachers adapted Making Meaning to their ongoing language arts program in one way or another. The most popular strategy was to be selective in the use of the program; that is, using only parts of the program or using it for particular purposes.

2001-2002: Teachers indicated that they integrated Making Meaning in the following ways:
Major Themes
- Selected aspects of the program and applied it to ongoing language arts curriculum, e.g., to model comprehension strategies, replaced basal work with Making Meaning stories (9)
- Used it sporadically whenever there was time, e.g., try to leave Friday open for Making Meaning, used occasionally but try to work strategies in (8)
- Inserted program within existing block, e.g., between read alouds and book clubs (7)

Minor Themes
- New block of time: added time rather than integrated it (4)
- Have not integrated it (1)

As suggested earlier, the influence of the prior year’s staff development activities could have carried over into this year. Because they had been trained in the previous year in the flexible application of teaching and learning strategies, teachers this year may have been more likely to view the reading comprehension strategies as tools they could select out and apply as they saw fit. In addition, the fact that lessons were sometimes not available to them in a timely fashion may have contributed to their view that the lessons were resources rather than prescribed activities to be administered routinely as written.

Did teachers use Making Meaning in other areas of teaching?

**Finding:** Almost all teachers at both schools reported using Making Meaning in other areas of teaching—in particular, using the communication and comprehension strategies in other subject matter areas.

2001-2002:
- 81% of teachers at Bel Aire Park used Making Meaning in other areas of teaching.
- 90% of teachers at Mt. George used Making Meaning in other areas of teaching. (See Appendix B, p. 22.)
- Teachers reported the following uses for Making Meaning:

  **Major Theme**
  - Applied Making Meaning to other books/subjects: all books, social studies, mathematics, non-fiction, expository text, core literature (13)

  **Minor Themes**
  - Applied “sharing with your partner” in other contexts: journal writing, student led discussion (5)
  - Book clubs (3)
  - Read alouds (2)
  - IDR (1)
In the prior year, CDP staff development activities focused on reading comprehension strategies and on the application of these strategies across the curriculum. It is possible that teachers' propensity to use the *Making Meaning* strategies in other areas of teaching is in some part attributable to what they learned and practiced during the 2000-01 school year.

**Community**

Teachers at the Napa program schools implemented the classroom component of the *Caring School Community* program, but did not implement the cross-grade, home-school, or schoolwide components. The classroom component consists of teacher-facilitated class meetings to build unity, establish classroom norms, plan and make decisions, and solve problems. In addition, the class meetings and cooperative structures used in the *Making Meaning* program contribute to building community in the classroom.

**Patterns Of Use**

**How frequently did teachers implement class meetings?**

**Finding:** In both 2000-01 and 2001-02, teachers reported using most types of class meetings at least monthly.

*2001-2002: Teachers at both schools reported using most types of class meetings at least monthly. (See Appendix B, pp. 2-3.)*

*2000-2001: Teachers at both schools reported using most types of class meetings at least monthly. (See Appendix C, pp. 2-3.)*

**How frequently did teachers implement norm-setting class meetings?**

**Finding:** In both 2000-01 and 2001-02, teachers reported using norm-setting class meetings on a quarterly basis.

*2001-2002: Teachers at both schools reported they typically did norm-setting class meetings quarterly. (See Appendix B, pp. 2-3.)*

*2000-2001: Teachers at both schools reported they typically did norm-setting class meetings quarterly. (See Appendix C, pp. 2-3.)*
Successes and Challenges

What were the signs of success associated with implementation of the Community component?

**Finding:** Class meetings, establishing norms, improvements in student communication patterns, and the formation of a sense of community were popular indications of success for teachers during both the 2000-01 and 2001-02 school years.

**2001-2002:** Teachers identified the following indicators of success:

**Major Themes**
- Class meetings, norm setting, discussions, vocabulary for talking together (8)
- Consistency, clear expectations, setting standards (Bel Aire Park only) (6)
- Children feel safe and positive about expressing themselves, sharing feelings and ideas, complimenting others (4)
- Unity building, sense of community (4)
- Solving issues, problem solving (4)

In addition, principals identified another sign of success related to teacher community:

Staff Development and Teacher Community: Staff Development this year has compensated to some degree for the absence of focus on Community by helping to build a community of learners among staff. (Principal Interviews)

**2000-2001:** Teachers identified the following indicators of success:

**Major Themes**
- Overall successful use of class meetings (5)
- Establishing norms (5)
- Student kindness, respect for each other (5)
- Student input/participation (5)
- Building classroom community (5)

**Major Themes**
- Students working cooperatively (2)
- Problem solving (2)

What were the major challenges associated with the implementation of the Community component?

**Finding:** Maintaining norms, dealing with student conflicts, finding time for the program, and maintaining student engagement were issues for teachers in both 2000-01 and 2001-02.
2001-2002: Challenges reported by teachers included the following:

**Major Themes**
- Instruction is difficult for teachers: finding time, involving students (16)
- Student behavior: competitiveness, tattling, anger, egocentricism, bullying, disruptive behavior, lack of motivation, unsupportiveness (8)
- The learning is difficult for students: young kids don’t get it, attention problem, ability problem, lack of comprehension (7)

In addition, principals and the Staff Developer reported other challenges:

- **Inadequate Implementation:** The Community piece has not been addressed even though there is a need for Community lessons to get the most out of Making Meaning. Additionally, some teachers don’t wish to give up instructional time to teach these skills. (Principal Interviews)

- **Questions at the End of Lessons:** Teachers and students had difficulty with Community-focused, “Discuss the story” questions at the end of lessons; they ignored them or treated them cursorily, and students did not tend to give authentic replies to these questions. (Staff Developer Interview)

2000-2001: Various challenges were cited for community, most relating to time required, and student interaction and engagement:

**Major Themes**
- Time to fit it in (6)
- Establishing and maintaining norms (5)
- Dealing with particular problem student(s) (5)

**Minor Themes**
- Student conflicts/not wanting to work together (3)
- Multiple skill/maturity levels with class (3)
- Students listening to each other/taking turns (3)
- Difficult for youngest students (3)
- Student participation/engagement (2)

The staff development activities, and to some degree the lessons on norms, emphasized building a sense of community and resolving classroom conflicts. However, this material was not accompanied by the introduction of a classroom management system or management tools for the teacher. Some of the challenges listed by teachers may reflect their inexperience with and lack of training in managing the new classroom context associated with class meetings and maintaining norms.
How do principals feel about the relative progress of the Community component?

**Finding:** In both years, Community is viewed as an important, but the least implemented, aspect of the program.

2001-2002:
- *This piece has not been addressed in either of the two years. It's the least implemented piece.* (Principal Interviews)
- *The Community piece is going along, but you don't learn how to run a class meeting overnight.* (Principal Interviews)
- *We need to work on this more. There is a need for Community lessons in order to maintain a sense of community among the staff and get the most out of Making Meaning lessons.* (Principal Interviews)

2000-2001:
- *Classroom community and school-wide community were seen as weak at both schools.* (Principal Interviews)
- *There was concern about the need to build community among staff.* (Principal Interviews)

As described earlier, community was addressed to some extent during both summer institutes. However, it was not emphasized during the year-long staff development teachers received.

**Staff Development**

How confident were teachers following the Summer Institute about implementing Community?

**Finding:** To a degree similar to last year, teachers at both schools felt motivated and prepared to use class meetings.

2001-2002:
- *Following the Summer Institute, teachers at both schools felt motivated (4.4) and prepared to use (4.0) the class meetings.*
- *Following the Summer Institute, teachers developed ideas about how the program in general and class meetings in particular build community in their classrooms (4.0).*
- *Teachers at the Summer Institute rated Community Building Strategies for Starting the Year moderately high in terms of helping their learning and understanding (3.4).* (See Appendix A, p. 41.)

2000-2001:
- *Teachers at both schools felt motivated to use norm-setting class meetings (4.4) and were prepared to use these meetings (3.9).*
• Teachers developed ideas about how the entire program in general and class meetings in particular build community in their classrooms (Mt. George: 4.5; Bel Aire Park: 4.0). (See Appendix B, p. 33.)

Implementation Support

This section presents findings on various supports for program implementation: the role of the principal, the role of the literacy coordinator, and during-the-year staff development and classroom support for teachers.

Role of Principal

What was the role of the Principal in the implementation of the Child Development Project and how did it differ this year compared to last?

**Finding:** In 2001-02, the Bel Aire Park principal was active in the staff development process all year; in contrast, the Mt. George principal “stepped back” from active involvement as the year progressed.

**2001-2002:**
• I’ve been involved a lot less than last year. I have stepped back to see who’s using what and how. I work closely with the two reading specialists. (Mt. George Principal Interview)
• At Bel Aire Park, the principal was on top of the staff. She gave support but did not demand and that was what was needed. (Staff Developer Interview)

**2000-2001:**
Both principals actively supported the program by:
• Attending staff development
• Arranging to purchase class-level sets of program books
• Facilitating SIPPS assessments
• Planning and problem-solving with teachers. (Staff Developer Interview, Principal Interviews)

Principal styles vary:
• The Mt. George principal (with nine years experience as Mt. George principal) coordinates the program and occasionally implements program components as a substitute teacher at the school. (Staff Developer Interview, Principal Interviews)
• The Bel Aire Park principal (in her first year as principal in the district) actively directs, organizes, and plans with teachers. Bel Aire Park is more than twice the size of Mt. George with 26 teachers. (Staff Developer Interview, Principal Interviews)

Overall, the data suggest that the effect of the principal on implementation success is more salient in year one than in subsequent years. That is, no
differences in the quality of implementation between Bel Aire Park and Mt. George could be traced to the observed differences in principal involvement levels during 2001-2002.

Role of Reading Specialist

What was the role of the Reading Specialist in the implementation of the Child Development Project?

**Finding:** As was true in 2000-01, the role of the reading specialist was central in the implementation of SIPPS for 2001-02.

2001-2002:
- *She has helped us think through organization and restructuring to accommodate leveled SIPPS groups; this was very important. We could not do SIPPS unless the reading specialist was involved.* (Bel Aire Park Principal Interview)
- *Mt. George had two reading specialists, both 40% time. They gave full support to SIPPS in the first block of time each day and did parallel teaching in the fall.* (Principal Interview).

2000-2001:
- *Both schools' reading specialists implemented SIPPS in the classroom.* (Principal Interviews)
- *The Bel Aire Park reading specialist was involved with other aspects of the program as well* (Staff Developer Interview, Principal Interviews):
  - Coordinates staff development activities within the school
  - Works with teachers on instructional conversations
  - Works with the principal on assessment.
- *Mt. George had two reading specialists, both 40% time. They gave full support to SIPPS in the first block of time each day and did parallel teaching in the fall.* (Principal Interview)

Information on the role of the reading specialist was limited to summary statements on the part of the principals, who were in agreement that reading specialists play a key role in the implementation of SIPPS. Future field tests might include survey and interview questions for teachers and others that would help determine the specific ways in which reading specialists contribute to the success of the CDP components.

Staff Development

How adequate was the amount of staff development received?

**Finding:** Teachers at both schools indicated that the amount of staff development was about right and, the majority of their comments on this question were positive.
2001-2002:

- Teachers at both schools indicated that the amount of staff development received was about right (3.2). (See Appendix B, pp. 27-30.)
- Teachers written responses to the question of the amount of staff development received were largely positive:

**Major Theme**
- Just right, good staff developer, good communication, good readings, good teaching and debriefing, much better this year (16)
- Teachers written responses indicated just a few issues with the amount of staff development:

**Minor Themes**
- Too repetitive: do things once; only teach and be observed once (3)
- More would have been better, more time to implement (2)
- More grade level time (2)

Overall, teachers received a fairly generous amount of staff development. It should be kept in mind that the amount of staff development they received was considerably greater than teachers typically receive, even in schools where a new program is being adopted.

**How positive were Bel Aire Park teachers about the structure of their new (for 2001-02) staff development design?**

A new structure for ongoing staff development was tried out in the 2001-02 school year at Bel Aire Park school. Staff development was provided on two consecutive days every other week (once per month for a particular teacher). In contrast, at Mt. George in 2001-02 and at both schools in the 2000-01 school year, staff development was provided on one day every other week.

**Finding:** Bel Aire Park teachers rated the structure of their staff development as very effective (more organized than last year and allowed for a more focused planning process, better discussions, and immediate feedback).

2001-2002:

Compared to last year, Bel Aire Park rated this year's structure for staff development as very effective (4.5). They explained their ratings as follows: (See Appendix B, p. 31.)

**Major Themes**
- Good format: two-day structure is better, more organized than last year, loved the post lesson discussion, easier to stay focused, good for grade level buy-in, easier to prep for (10)
- Good process: planning, observing, and debriefing, immediate feedback, more concrete (6)
Minor Theme

- More comfortable staff developer; excellent communicator (3)

The two-day staff development structure was sufficiently well received to become the model for staff development activities at subsequent field test sites.

What suggestions did teachers have about ways that staff development could be improved?

Finding: Teachers had a number of suggestions including making the process more grade-level specific, reducing the reading and time allotment, and expanding the application of Making Meaning to reading in the content areas. The principals and the Staff Developer agreed that Making Meaning will continue to require support to be effective.

2001-2002:
Teachers made the following suggestions on the teacher survey:

Major Theme

- Make it more grade level specific, more grade level time (6)

Minor Themes

- Clone the staff developer (3)
- Deal with special kids, struggling readers (2)
- Less reading, fewer articles (2)
- Less time: once a quarter, shorter observations, less time out of classroom (1)
- A writing component could be helpful (1)
- Watching each other got too repetitive (1)
- Involve families so it carries over into the home (1)
- We each should teach the lesson; we all plan then come back and discuss it (1)
- Keep staff development going next year (1)

In addition, focus group teachers, principals and the staff developer had other suggestions about improving staff development:

- Making Meaning will Require Further Staff Development: Principals inquired about the availability of staff development in the future; they believe that Making Meaning requires collaboration and support to be successful. (Principal Interviews)

- Reading in the Content Areas: This may be an area worthy of future staff development, especially involving expository texts and the role of Making Meaning in meeting state standards. (Teacher Focus Groups, Principal Interviews, Staff Developer Interview)
• *Time Allocation for Making Meaning Lessons:* More work is needed in helping teachers teach Making Meaning more efficiently. *(Staff Developer Interview)*

According to the staff developers, some of the difficulties teachers had with the current staff development model were a carry over from the previous year when teachers were given guidance on how to pick out and apply reading comprehension strategies across the curriculum. This past year, some of the teachers were frustrated with a staff development model designed to help them administer lessons on a regular basis. It may have been helpful for the principals or the staff developers to get a commitment in advance for a more systematic implementation of the program as written.

The comments made by some teachers concerning the need for more grade-level-specific staff development probably reflects frustration about the after-school meetings during which time teachers read and talked about articles. These articles did not always pertain to teachers' grade level to the extent they wished.
Student Outcomes

As described below, this evaluation assessed students' academic outcomes in terms of their decoding ability, and performance on the SAT-9 and district literacy assessments. Affective outcomes addressed attitudes and behaviors related to school, reading, community, and social/ethical development.

Decoding Abilities

What were the effects of the program on students' decoding abilities?

**Finding:** First through third grade students at the program schools showed significantly greater gains in decoding scores than students at the matched comparison schools ($p < .01$, ES = .24). There were no significant differences between kindergarten program and comparison students in decoding gains.

*The greater gains in decoding scores for first through third grade program students were found at both program schools, relative to their matched comparison schools, and at both the traditional and alternative tracks at Bel Aire Park and its matched comparison school. However, within subgroup, only the differential gains for students in the traditional track at Bel Aire Park were statistically significant. (See Appendix A, p. 32.)*
Thus, the largest positive program effects on decoding were found among student populations that were largely Hispanic and socioeconomic disadvantaged.

It also should be noted that kindergarten students had only received 2-3 months of SIPPS instruction at the time of the posttest, in contrast to the 7 months of instruction received by students in grades 1-3. The amount of instruction in kindergarten may not have been enough by the time of the posttest to yield significant gains in decoding relative to comparison students.

- **Students at Mt. George had the highest decoding scores of students in the study, scored significantly higher than students at the other three schools on decoding scores at pretest, and scored significantly higher than students at both comparison schools (but not Bel Aire Park) on decoding skills at posttest.**

The fact that Mt. George students did not have significantly greater gains than their comparison students may be due, at least in part, to a ceiling effect.

**Finding:** Among program students, English language learners showed significantly greater gains in decoding abilities than English-proficient students \((p < .03, ES = .22)\).

The particular effectiveness of SIPPS with English language learners is consistent with teacher reports and the findings from the earlier pilot study of SIPPS.
Standardized Achievement Test Scores

What were the effects of the program on students' standardized achievement test scores?

**Finding:** Program and comparison students did not differ significantly in changes from baseline (1998-2000) to the intervention years (2001-2002) in total reading or total language scores on the SAT-9, nor did they differ significantly in changes from baseline in the spelling or reading comprehension subtests of the SAT-9.

- This absence of effect held across both program schools, relative to their comparison schools, and for students in the alternative and traditional tracks at Bel Aire Park, relative to comparison alternative and traditional students.

  Overall, this absence of observed effects on the SAT-9 need not be regarded as a program failure. SIPPS is, primarily, a decoding program, whereas the SAT-9 is a test of components of reading other than decoding, namely, vocabulary, reading comprehension, grammar, spelling and word usage. Additionally, it may be that SAT-9 total reading and total language scores would be impacted by the combined implementation of SIPPS and Making Meaning. This possibility should be evaluated following a more complete and systematic implementation of the Making Meaning program.

- Students in both the alternative and traditional tracks at Bel Aire Park performed significantly lower on the SAT-9 than their matched comparison students during all three baseline years as well as the two intervention years. However, the size of the difference in reading scores decreased somewhat during the two intervention years, relative to the three baseline years.

  Although this decrease from baseline to intervention years on the part of Bel Aire Park students was not significant, it does suggest that the program had some effect.

- Students at Mt. George generally did not differ significantly from their matched comparison students in reading or language scores during baseline or intervention. However, they scored significantly higher than comparison students in reading scores in 2001 (p < .05, ES = .21). This increase in reading achievement was not maintained in 2002. (See Appendix A, p. 33)

  It may be important to keep in mind that Mt George is a high achieving school. Since Mt. George students are performing at the top of the distribution of scores on the SAT-9, they would be more likely to experience a “ceiling effect.” In other words they have less “room” to exhibit improvement between baseline and intervention.

  A valid test of the effect of the CDP program on SAT-9 reading comprehension scores should await full and systematic implementation of the Making Meaning program.
District Literacy Assessments

What were the effects of the program on tests of language ability?

**Finding:** Program students showed significantly greater gains from baseline than comparison students in *language mechanics* ($p < .05$, $ES = .16$) and *language expression* ($p < .06$, $ES = .14$).

- *These differences generally held across both program schools, relative to their comparison schools, and for students in the alternative and traditional tracks at Bel Aire Park, relative to comparison alternative and traditional students.* (See Appendix A, pp. 34-35.)

**Finding:** Program and comparison students did not differ significantly in changes from baseline in *narrative, textual, or functional language, vocabulary, or spelling.* (See Appendix A, pp. 34-35)

It is worth noting that teachers at the Napa schools did not use the Independent Daily Reading component of SIPPS (or *Making Meaning*), but rather used other approaches to independent reading (Guided Reading, sustained silent reading). Thus, the increases in vocabulary that would be expected with the IDR component of the program would not necessarily be expected from the Napa intervention.

It is puzzling that the differences between program and comparison students would be in language mechanics and expression, which have to do with capitalization, punctuation, word usage, sentence structure, and paragraph content and organization – none of which is taught by SIPPS or *Making Meaning*.

Student Attitudes and Behavior

What were the effects of the program on students' reading-related attitudes and behavior?

**Finding:** Program students increased significantly from baseline in their *liking for reading*, compared to a decline among students at the matched comparison schools ($p < .03$, $ES = .23$). (See Appendix A, pp. 36-38)

**Finding:** Program students increased significantly from baseline in their *frequency of reading outside of school*, compared to a decline among students at the matched comparison schools ($p < .01$, $ES = .26$). (See Appendix A, pp. 36-38)

- *These differences generally held across both program schools, relative to their comparison schools, and for students in the alternative and traditional tracks at Bel Aire Park, relative to comparison alternative and traditional students.* (See Appendix A, pp. 36-38.)
These results are consistent with findings from the teacher surveys and principal interviews. Teachers and principals in Napa and during an earlier field test of SIPPS reported that students increased in their enjoyment of reading and motivation to read as a result of the SIPPS program.

**Finding:** Relative to their comparison students, students in the traditional track at Bel Aire Park also showed a significant positive effect on *liking for read-alouds* \((p < .02, \text{ ES} = .42)\). Students at the alternative track at Bel Aire Park showed a significantly negative effect on *liking for read-alouds* \((p < .04, \text{ ES} = -.41)\). (See Appendix A, pp. 36-38)

The reason for this difference is not known. During 2000-2001, all teachers at both schools reported doing Read Alouds, although mostly using non-DSC books. Because of the piloting of *Making Meaning* (which also uses Read Alouds) during 2001-2002, teachers were not asked about their use of other Read Alouds.

**Finding:** Program and comparison students did *not* differ significantly in *sense of efficacy as a reader, the number of books read since the start of the school year, or the amount of time spent reading in the last week.* (See Appendix A, pp. 36-38)

The absence of significant effects on the number of books read and time reading during the past week is not consistent with the difference in frequency of reading outside of school reported above. A program effect on sense of efficacy at reading also would be expected, although primarily among beginning readers at the primary grades. Primary grade students were not administered the student surveys.

**What were the effects of the program on students’ school-related and social/ethical attitudes and behavior?**

**Finding:** Program students increased significantly from baseline in their *sense of school community*, compared to a decline among students at the matched comparison schools \((p < .01, \text{ ES} = .27)\). Program and comparison students did *not* differ significantly in their *liking for school*. (See Appendix A, pp. 36-38)

- **The significant positive gain in sense of school community held across both program schools, relative to their comparison schools, and for students in both the alternative and traditional tracks at Bel Aire Park, relative to comparison alternative and traditional students.** (See Appendix A, pp. 36-38)

- **Students at all four Napa schools scored quite high in liking for school at baseline and during each of the intervention years:**

**Finding:** Relative to their comparison students, students in the traditional track at Bel Aire Park also showed significant positive effects on *conflict resolution skills* \((p < .04, \text{ ES} = .34)\). Students at the alternative track at Bel Aire Park did not differ from their comparison students in *conflict resolution skills*. (See Appendix A, pp. 36-38)
• Bel Aire Park traditional students had the lowest scores in conflict resolution skills at baseline of any group, and increased during the intervention years, compared to a decrease among traditional track comparison students.

**Finding:** Relative to their comparison students, students at Mt. George increased significantly in commitment to democratic values \((p < .01, \text{ ES} = .57)\) and altruistic behavior \((p < .05, \text{ ES} = .35)\). (See Appendix A, pp. 36-38)

• Students at Bel Aire Park also increased from baseline in their commitment to democratic values, but so did students at their comparison school. Students at Mt. George's comparison school declined from baseline in commitment to democratic values.

• Mt. George students were the only group that increased from baseline in altruistic behavior. Students at Bel Aire Park and both of the comparison schools showed declines from baseline in altruistic behavior.

Overall, the positive program effects on students’ sense of community and on social/ethical attitudes and behavior are consistent with other CDP findings. Additionally, although the Caring School Community program was not fully implemented in this field test, teachers reported regular use of class meetings and they received Unit One of Making Meaning, which focused on building a “community of readers.”
SUMMARY AND CONCLUSIONS

This section summarizes this evaluation’s findings and conclusions in ten inter-related areas: implementation of SIPPS and Independent Daily Reading, growth in student decoding ability, Making Meaning implementation, growth in student comprehension skills, Caring School Community implementation, growth in sense of community and social/ethical development, other academic outcomes, other affective outcomes, program integration, and support for implementation. In addition, where appropriate indicated program and staff development revisions, and evaluation activities being undertaken by DSC are described.

Implementation of SIPPS and Independent Daily Reading

The SIPPS program was successfully implemented at both of the Napa program schools. It was implemented at all primary grades and as an intervention program for those students who needed additional phonics instruction at approximately half of the upper-grade classrooms. All three levels of SIPPS were implemented at both of the program schools. The program was implemented as intended with the exception that teachers did not follow all aspects of Independent Daily Reading as prescribed. Students typically engaged in some type of independent reading for approximately 20-25 minutes daily. Teachers used leveled books in their classrooms but did not necessary assure that students were reading in books appropriate for their level or monitor their comprehension during independent reading.

Overall the strength of teachers’ implementation of SIPPS improved in the second year, as evidenced by: increased motivation and preparation to implement SIPPS by the end of the second Summer Institute, increased frequency of SIPPS sessions, increased use of multiple SIPPS groups, and principal testimony. Further, teacher modifications to the program primarily involved additional language arts activities, and the percentage of teachers reporting program modifications decreased from the first to the second intervention year, particularly at Mt. George (where the principal was more explicit about teachers using SIPPS as intended).

Implementing SIPPS is not, however, without challenges. These challenges include difficulties associated with managing multiple groups (especially, keeping non-SIPPS students engaged during SIPPS activities), finding time for the program in the already crowded curriculum, and managing multiple levels of SIPPS groups in a single classroom.

Based on these findings, DSC in incorporating more information on managing multiple groups into the SIPPS staff development. Further, in order to emphasize implementation of Independent Daily Reading as prescribed, DSC is presenting it to schools as an integral part of both SIPPS and Making Meaning.

Growth in Student Decoding Ability

Teacher reports almost universally supported the success of SIPPS at improving students’ decoding abilities. When asked what were the major successes associated with SIPPS, 22 out of 26 responses (85%) on the Teacher Survey cited observations of students’ reading competence or knowledge of phonics. Both principals of the program schools also indicated that SIPPS was effective at improving students’ decoding abilities.
These testimonials were supported by the findings from comparative analyses of gains in decoding for students at the program and comparison schools—program students in grades 1-3 improved significantly more than students at the matched comparison schools in scores on the Slosson Oral Reading Test from pretest to posttest during 2001-2002. Further, among program students, English language learners showed significantly greater gains in decoding ability than English-proficient students. This finding is consistent with teacher reports as well as findings from an earlier pilot study of SIPPS, and further confirms that SIPPS works particularly well with students in most need of decoding instruction.

Making Meaning Implementation

Making Meaning was not fully implemented as a reading comprehension program in the Napa program schools. Instead, teachers informally did instruction in comprehension strategies with their students during 2000-2001, and piloted some Making Meaning lessons (as they became available) during 2001-2002. None of the teachers implemented all of the lessons, and the majority (74%) modified the program by shortening the lessons or not implementing some portion of them. This is not surprising, given that the Napa teachers were only participating in a pilot test of the program, and were not provided with the lessons long enough in advance of when they were to be implemented to allow for adequate time for planning and preparation.

Given these circumstances, it is not surprising that when asked about challenges in implementing the program, over 80% of teachers’ responses had to do with time constraints or the difficulties of fitting the program within other subjects and the language arts curriculum. In addition, teachers and administrators expressed the need, in order for the program to be successful, for additional staff development and ongoing support. Consequently, the Napa study cannot be considered an adequate evaluation of the program.

Despite these caveats, there were clear signs of success related to the implementation of Making Meaning during this pilot test year:

- The teacher survey offered several opportunities for teachers to express their opinions about Making Meaning and one opportunity to express their opinions about the overall effectiveness of Making Meaning (and SIPPS). Although they could have used these opportunities to be dismissive about the program or to be critical of program features, no such negative comments were found in any survey.

- Almost all of the teachers at both program schools reported using Making Meaning strategies in other areas of teaching, and many indicated that their students were spontaneously applying program strategies and skills to their reading in other subject areas.

- Teachers reported feeling confident about using the lessons prior to implementation and thought that the staff development provided during the year was productive and helpful.

- Principals reported that the program provided a “framework and vocabulary” that gives teachers a common language to talk about instruction and facilitates both community-building and students’ socialization.
• Although *Making Meaning* represents a departure in methods and philosophy from those of *SIPPS*, almost all of the teachers described the combination of *SIPPS* and *Making Meaning* as a “good match” for reading instruction.

**Growth in Student Comprehension Skills**

When asked about the successes of *Making Meaning* on the end-of-year surveys, almost all of the teachers’ responses focused on the effectiveness of the instructional strategies, perceived student growth in reading comprehension, and student enjoyment/engagement during the lessons. Teachers also indicated that the program was effective with both English-speaking and ELL students. These reports suggest that the program is working as intended and supports both student growth in comprehension and classroom community.

However, teacher perceptions of increases in student comprehension were not supported by analyses of findings from standardized achievement tests. Program and comparison students did not differ significantly in changes from baseline (1998-2000) to the intervention years (2001-2002) on the reading comprehension subtest of the SAT-9. However, given the partial and limited duration of implementation in the Napa schools, significant growth in SAT-9 scores was not expected.

An adequate assessment of the *Making Meaning* program must await the findings from the current field test of the entire program in schools in Newark, California. It is being field-tested as a complete program at all elementary schools in the Newark Unified School District in California. This field test evaluation has two goals. The evaluation will assess how well the program is working in terms of its impact on students’ reading comprehension (via growth on the SAT-9 and on the district’s own comprehension measure). A second goal of the evaluation is to solicit teacher feedback: on issues, concerns, and successes related to classroom implementation of the program; on how well the initial and ongoing staff development is meeting their needs; and for program revisions. Results from this field test will be available in fall, 2003.

**Caring School Community Implementation**

The *Caring School Community* program was not fully implemented in the Napa schools. The primary challenge to this aspect of the CDP program in Napa was the difficulty of building community among their students. The majority of teachers on the Teacher Survey commented on how hard it was to get students involved in community-building activities or the extent to which student behavior interfered with the accomplishment of community-building objectives. Principals similarly commented that community-building was very important, but was the least implemented component of the CDP program during both years.

**Growth in Sense of Community and Social/Ethical Development**

Although the *Caring School Community* program was not fully implemented in the Napa schools, the evaluation findings indicate that the classroom aspects that were implemented had a significant positive impact on students’ social/ethical development at the program schools. The positive effects shown by the results from analyses of the student survey data are supported by
teacher reports of the program’s effectiveness and that it is seen as important by both teachers and the principals of the program schools.

The positive effects on students' sense of community and social/ethical attitudes and behaviors observed in Napa are consistent with the findings from the extensive previous research on the old CDP program, as well as with the more recent findings from the third-party evaluation of the Caring School Community program in several elementary schools in St. Louis. These findings suggest that the revisions to CDP undertaken to enhance its effects on reading and literacy will not reduce the program's demonstrated effectiveness at building school community and enhancing students' social and ethical development.

**Other Academic Outcomes**

Similar to findings on SAT-9 reading comprehension, there were no significant differences between program and comparison students on the spelling subtest or in total reading or language of the SAT-9.

On the other hand, scores on district-administered literacy assessments showed significant differences favoring program students in language mechanics and expression. However, as these language skills are not a focus of either SIPPS or Making Meaning, it is unlikely that these differences can be attributed to CDP.

**Other Affective Outcomes**

In addition to the significant effects on sense of community and social/ethical development there were significant differences favoring program students in liking for reading and frequency of reading outside of school.

**Program Integration**

Not surprisingly, teachers generally saw SIPPS and Making Meaning as distinct programs with different purposes and approaches. At the same time, teachers rated the two programs as very effective and commented that they were a good combination. A larger problem remains, however. How to integrate the two programs into the larger language arts curriculum given existing time constraints.

Teachers at the program schools integrated SIPPS into their language arts curriculum in one of three ways:

* simultaneous implementation whereby other language arts activities were carried out at the same time as SIPPS using multiple groups

* separate implementation whereby SIPPS was introduced during one block of time and other language arts activities were covered during another time of the day

* complete infusion, whereby SIPPS was the focus of all language arts activities.
Further inquiry is necessary to determine more detail about SIPPS integration strategies, including the strategies teachers are most likely to choose.

For *Making Meaning*, the most popular strategy was to be selective in the use of the program; that is, using only parts of the program or using it for particular purposes. This finding is not necessarily reflective of expected integration strategies, given that it is based on a pilot implementation where teachers faced time constraints.

In the 2002-03 field test evaluation of *Making Meaning*, DSC is investigating how teachers who have an opportunity to integrate a full version of *Making Meaning* into their language arts curriculum are doing so. DSC will be using this information to develop a better understanding of how both SIPPS and *Making Meaning* can be fit into the language arts block.

**Support for Implementation**

Not surprisingly, principals’ and literacy coordinators’ direct involvement and support for the CDP components contributed to their success in the program schools. Although their styles varied, principals provided direct support for SIPPS in particular -- by attending staff development, teaching lessons themselves, helping with assessments, and planning with teachers. The reading specialists were even more directly involved, providing ongoing classroom support. Principal and literacy coordinator support and involvement are key requisites for current and prospective schools and districts working with DSC to implement the full version of CDP.

Teachers generally responded favorably to the staff development they received. Following each year’s Summer Institute, they indicated that they were both motivated and prepared to implement the program components and, in the end-of-year surveys, they indicated that the amount of staff development they had received was appropriate and made many favorable comments about it. That being said, it also is clear that the on-site support provided during 2001-2002 was better than that provided during 2000-2001. In particular, the structure of this support was changed at Bel Aire Park (but not at Mt. George) in 2001-2002 to two consecutive days once per month rather than one day every other week. The Bel Aire Park teachers rated this structure as very effective, and indicated that it allowed for a more focused planning process, better discussions of lessons, and more immediate feedback than the earlier approach, and that it was easier to stay focused and contributed to grade level “buy in.” Given this clear improvement, this same approach is being used at all schools during this year’s field test of Making Meaning.