

About the Developmental Studies Center

Our Mission

Developmental Studies Center (DSC) is a nonprofit organization dedicated to children's academic, ethical, and social development. Since 1980, DSC has developed school-based and after-school programs that help children develop capacities to think deeply and critically so they will continue learning throughout their lives and strengthen their commitment to such values as kindness, helpfulness, personal responsibility, and respect for others.

DSC's Programs Develop Skills and Community

Programs for use in classrooms

Caring School Community™ • Grades K–6

The Caring School Community (CSC) program is a nationally recognized, research-based program that builds community—in the classroom, across grades, schoolwide, and with families.

Making Meaning® • Grades K–8

The Making Meaning program is a reading comprehension curriculum that teaches comprehension strategies through read-alouds, collaborative structures, and reflective partner work.

SIPPS® (Systematic Instruction in Phoneme Awareness, Phonics, and Sight Words) • Grades K–12

The SIPPS program teaches decoding systematically. It is designed specifically for intervention and covers single-syllable decoding, short vowels, simple consonants, complex vowels, consonant digraphs, polysyllabic strategies, and high-frequency sight words.

Being a Writer™ • Grades K–5

The Being a Writer program is a yearlong writing curriculum—the first program of its kind to bring together the latest research in teaching writing with support for students' social and ethical development. (Available August 2007)

Programs for out-of-school time

AfterSchool KidzLit® • Grades K–8

The AfterSchool KidzLit program is a literacy enrichment program consisting of terrific read-aloud books, and discussions and activities that help kids make connections between the stories, their own lives, and the world.

AfterSchool KidzMath™ • Grades K–6

The AfterSchool KidzMath program provides academic enrichment using cooperative math games and literature-based activities. Kids deepen their understanding and practice important math skills—and have fun.

Science Explorer • Grades K–6

Science Explorer is an inquiry-based, interactive program of experiments using ordinary materials that inspire students to explore scientific principles.

Math Explorer • Grades 6–8

Math Explorer invites children to fly planes, launch rockets, learn card tricks, and make cool stuff to take home—all while practicing the important math skills middle-school students need extra help with.

For more information, please visit www.devstu.org or contact us by phone at 800.666.7270.



**DEVELOPMENTAL
STUDIES CENTER™**

Nonprofit. Research Based. Mission Driven. Since 1980.

The Use of Implementation Data in Assessing the Effectiveness of the Child Development Project

Victor Battistich
Developmental Studies Center

May, 2000

Presented at the meeting of the Society for Prevention Research, June 2000, Montreal.

The Use of Implementation Data in Assessing the Effectiveness of the Child Development Project

The Child Development Project (CDP) is a comprehensive, whole school reform program that helps elementary schools to become caring communities of learners—environments characterized by supportive interpersonal relationships, shared goals, responsiveness to students' developmental and sociocultural needs, and an emphasis on prosocial values of personal responsibility, concern for others, and fairness, as well as a commitment to learning. The program includes classroom, school-wide, and family involvement activities that, working synergistically, are expected to foster students' positive development and resilience to risk. CDP is an ecological approach to intervention that is designed to influence all aspects of the school—curriculum, pedagogy, organization, management, and climate.

Information about CDP's theoretical bases, intervention model, and program elements is provided in Battistich et al (in press). The present paper focuses on the procedures used to measure program implementation in a four-year (baseline followed by three intervention years), multi-site demonstration trial involving a diverse group of 12 intervention and 12 comparison schools from six school districts across the US.

Procedures

We anticipated that there would be substantial variability in the adequacy of program implementation among schools and teachers, and therefore conducted systematic classroom observations each year—as well as included items in teacher and student questionnaires—to assess progress toward implementation. Each classroom in the 24 schools (n=541-582) was observed four times during each school year, using a structured observation system. There were separate teams of four observers in each of the six districts. In order to maintain observer blindness as to condition, new groups of observers were hired and trained each year. Each year, each of the observers made one visit to each classroom in his/her district. The observers watched class activities during each of 12 four-minute periods, and, after each, made a series of ratings of the occurrence of various activities, practices and qualities during the period, using three-point scales (0 = no occurrence or intensity, 1 = minimal occurrence or intensity; 2 = more than minimal occurrence or intensity). The observers focused on the teacher and classroom activities for half of the four-minute periods (these are called "classroom focus" periods), and on the behavior and activities of small groups of students for the other half ("student focus" periods), alternating between the two in successive periods. Each classroom visit took about 90 minutes in total.

To help train the observers in all districts a set of videotapes of a broad range of class sessions, divided into four-minute segments, was created. Criterion scoring of these videotapes was done by project staff members. Another set of similar videotapes was used to help maintain observers' accuracy and consistency during the course of the year. Observers in all districts periodically scored the same videotape segments independently, and their scoring was compared with the criterion scoring of the same segments. Average overall observer agreement with the criterion scores was 74.61% over all four years (using an index that controlled for the likelihood of agreement on infrequent events: Clement, 1976).

Constructing an Index of Implementation

Scores for each item in the classroom observation form were aggregated across the segments and visits for each classroom each year. Items relating to several program-relevant aspects of teacher practices were identified from the observation data, and scales representing each were constructed after consideration of inter-item correlations, exploratory factor analyses, and the conceptual meaning of each construct. To confirm the distinctness of these scales, we factor analyzed the final set of selected teacher practices items, using principal axes factor analysis. These analyses indicated that the items could reasonably be organized into seven teacher practices factors that appeared to represent important aspects of the intended CDP program.

Because we felt that the classroom observations alone might be missing some of the subtleties of CDP program implementation, and, particularly, that effective program implementation involved teachers' *intentions* and *orientations toward students* as well as their overt *behaviors*, we decided to include several measures of these teacher attitudes (from the teacher questionnaire) in our final measure of overall program implementation. In creating the index of implementation we therefore combined the seven scales derived from the classroom observations with four teacher attitude scales.¹ The observation-based and teacher attitude scales used in the implementation index are described in Table 1.

The mean internal consistency reliability of the implementation index over the four years, assessed by considering each of the 11 component subscales as an "item," was .74.

Validating the Index

A discriminant function analysis was conducted using the implementation scores of three groups of teachers: those nominated by CDP program staff as particularly good implementers, the rest of the program teachers, and the comparison teachers. For this analysis, the set of observational and questionnaire scales was used, rather than the final implementation index score, to examine the extent to which the individual measures contributed to discriminating the groups. The groups were found to differ significantly on all of the measures, and the discriminant function scores strongly differentiated the three groups. A multiple classification analysis produced a "correct" classification for 90% (18 of 20) of the nominated teachers, indicating that the measures had adequate validity.

Limitations of the Index

Although the measures of implementation appeared to do a good job of discriminating among teachers with respect to their program-relevant practices and attitudes, measures of implementation of the school-wide and family involvement components of the CDP program were not included. The implementation index thus only measured the classroom components of CDP. Moreover, one important aspect of the classroom program—the use of CDP's literature-based reading program—also was

¹ In computing the index, the observation and questionnaire scales were first re-scaled onto a common metric, and then the scales within each set (i.e., observation and questionnaire) were then averaged. The final implementation score was formed by averaging the observation and questionnaire "subscales," so that classroom practices and teacher attitudes contributed equally to the measure of program implementation.

not included in the implementation index. Although a number of items representing the use of literature were included in the classroom observation system, these items were found to be incapable of reliably distinguishing the use of CDP's literature program from other classroom approaches that involved the use of literature.

Utility of the Index

Study-wide comparisons revealed a number of statistically significant differences between students at the 12 intervention and 12 comparison schools, but about as many favored comparison students as favored program students. Absent implementation data, the conclusion of the study would be one of program ineffectiveness. However, as noted above, we anticipated that there would be substantial variability among teachers and schools in program implementation, and examination of the implementation data showed that only five of the 12 intervention schools had accomplished widespread changes in program-relevant classroom practices and teacher attitudes, relative to observed practices and attitudes at their comparison schools, by the end of the third intervention year. For these five schools, t -values for planned contrasts on linear changes from baseline ranged from 1.77 ($p < .09$) to 4.83 ($p < .001$), with effect sizes ranging from .41 to 1.10. At the remaining seven intervention schools, the t -values for the planned contrasts were all less than 1.00, and the effect sizes ranged from -.06 to .20. Figure 1 illustrates these differences in observed changes in practices at a representative "low" and "high" change school.

Figure 1 here

In contrast to study-wide findings, comparisons of student outcomes at the five "high change" program schools versus at their matched comparison schools revealed a large number of positive program effects, including increases in students' *sense of the school as a community*, *school-related attitudes and academic motivation* (e.g., liking for school, task orientation toward learning, academic self-esteem), and *social attitudes, values, and behavior* (e.g., concern for others, conflict resolution skills, commitment to democratic values, altruistic behavior), and reductions in their involvement in *drug use and other problem behaviors* (e.g., use of alcohol and marijuana, involvement in gang fighting). Overall, more than 50% of the outcome variables examined showed significant effects favoring program students at these five schools, and there were no significant differences favoring comparison students. Ignoring significance levels, fully 93% of the outcome variables showed changes from baseline favoring program students.

We also have used the implementation data to conduct explicit tests of CDP's theoretical model (in which effects on student outcomes are hypothesized to be mediated by effects on students' sense of the school as a community). These analyses have been conducted both at the individual level (i.e., growth curve modeling) and at the classroom level (using aggregate data), and have provided strong support for the program theory. An illustration of the complete model examined at the classroom level and a summary of the findings from some of these analyses is shown in Figures 2 and 3, respectively.

Figures 2 & 3 here

Finally, the implementation data also has proved to be important in examining program effects at follow-up. Currently, we are conducting a study of students from three of the six districts while they are in middle schools. Half of the former program students are from three of the “high change” elementary schools, and half are from three of the “low change” schools. Interim analyses of middle school data revealed a number of program effects at follow-up, particularly for students from the “high change” elementary schools. Study-wide, former program students scored significantly higher than comparison students in *self-esteem*, and were more involved in *positive group activities* (e.g., organized sports, clubs, community youth groups) than comparison students. In addition, former program students had higher *educational aspirations* than comparison students, perceived *student-teacher relationships* in their schools to be more positive than did comparison students, reported engaging in fewer acts of *misconduct at school* than comparison students (e.g., breaking school rules, showing disrespect for teachers, cheating on an exam), and reported less *use of marijuana* than comparison students. These significant differences represent approximately 21% of the outcome variables examined, and in contrast to the study-wide findings for the full sample during elementary school, there were no significant differences favoring comparison students.

A much larger number of effects favoring program students was found for the subsample of students from the three high change schools: 59% of the outcome variables examined showed differences favoring the former program students, and there were no significant differences favoring comparison students. With respect to school-related attitudes, perceptions, and performance, former program students had higher *course grades* and *achievement test scores* than comparison students; scored higher in both *educational aspirations* and *educational expectations*; had more positive perceptions of the *supportiveness of the school environment* for students and of the *quality of student-teacher relationships*; scored higher than comparison students in *trust in and respect for teachers* and *liking for school*; and scored lower than comparison students in *loneliness at school*. Former program students from the high change elementary schools also scored higher than comparison students in *self-esteem* and *sense of efficacy*. In terms of involvement in positive and negative activities, former program students were more involved in *positive group activities* and reported engaging in less *misconduct at school* than comparison students, and also reported that they *attended religious services* more frequently and were less involved in *serious delinquency* (e.g., carrying a weapon, using a weapon in a fight, selling drugs, committing burglary) than comparison students. Finally, they reported that more of their friends *attended religious services* and were *positively involved in school* (e.g., completed their assignments, liked school most of the time, tried to do good work in school, planned on graduating from school), and that fewer of their friends were involved in *misconduct at school* or in *delinquent behaviors* than did comparison students.

Concluding Comments

In spite of limitations in the extent to which all components of the CDP program are represented, the implementation index has obviously been invaluable for

demonstrating program effectiveness. Without such data, the conclusion from a “black box” evaluation would have been that the program was not effective (or, at best, had mixed outcomes). Only by using the implementation data to distinguish schools that had made significant progress toward program implementation by the end of the third intervention year from those that had not were we able to provide clear evidence that, *when implemented widely throughout a school, CDP results in a broad range of benefits for students*. Perhaps equally important, the implementation data allowed us to conduct direct tests that confirmed that the data were consistent with the program’s theoretical model, further bolstering the claim of program effectiveness.

References

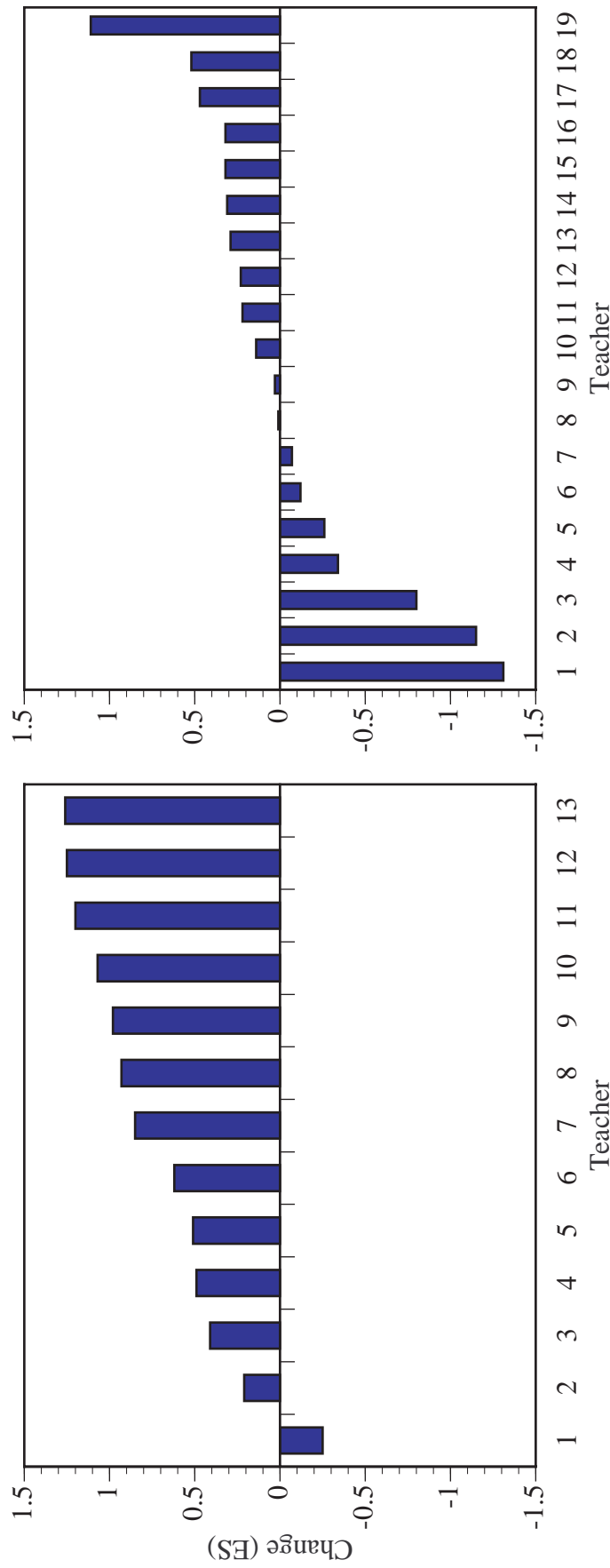
- Battistich, V., Schaps, E., Watson, M., Solomon, D., & Lewis, C. (in press). Effects of the Child Development Project on students' drug use and other problem behaviors. *Journal of Primary Prevention*, **21**.
- Clement, P. W. (1976). A formula for computing inter-observer agreement. *Psychological Reports*, **39**, 257-258.

Table 1
Measures Included in Implementation Index

<u>Classroom Observation Scales</u>	<u>Sample Items</u>
Promotion of student autonomy, freedom and influence (11 items, $\alpha = .63$)	Teacher gives students choice of activities Students participate in planning
Use of cooperative learning activities (6 items, $\alpha = .72$)	Teacher encourages helping and cooperation % of periods with cooperative groups % of periods with structured cooperative activity
Promotion of social understanding and prosocial values (14 items, $\alpha = .83$)	Teacher mentions, discusses, encourages understanding of others' motives, feelings, perspectives Teacher reads, uses media, assigns/leads activity focused on prosocial values
Level of personal relationship between teacher and students (6 items, $\alpha = .69$)	Teacher makes effort to relate personally to students Teacher shows interest in student feelings about tasks
Minimization of external control (7 items, $\alpha = .78$)	Teacher uses or promises rewards/grades/points Teacher uses or threatens punishment Pervasiveness of teacher's control over students
Emphasis on intrinsic motivation (5 items, $\alpha = .65$)	Teacher talks about the inherent interest of academic activities Teacher talks about relevance of academic task/activity to other or long-range goals
Elicitation of student thinking and active discussion (13 items, $\alpha = .78$)	Teacher asks for inferences/hypotheses Teacher draws connections between students' ideas Teacher encourages students to follow-up on each others' ideas
<u>Teacher Questionnaire Scales</u>	
Optimism about students' learning potential (10 items, $\alpha = .67$)	There is really very little I can do to insure that most of my students achieve at a high level. (R) If students aren't given the right values and habits at home, there is little that schools can do for them. (R)
Trust in students (3 items, $\alpha = .61$)	Students can be trusted to correct their own tests. Students can be trusted to work together without supervision.
Belief in importance of student self-direction, exploration, collaboration, and understanding. ("constructivist" beliefs) (9 items, $\alpha = .80$)	Students learn best when they are involved in exploring things, inventing and trying out their own ways of doing things. Students need to talk about their ideas in order to really learn. Students are strongly motivated by the need to understand things for themselves. Hearing each others' ideas is essential for student learning.
Belief in promoting student autonomy. ("control ideology") ^a	Responses about appropriateness of alternative ways for teachers to handle four described student problems

Note. (R) = reverse-scored item.

^aBecause this measure is a combination four of differentially-weighted subscores, a measure of internal consistency of the overall scale could not be calculated.



School 1

$M = .73, SD = .45$
 $t(38) = 4.68, p < .001$ vs. Comparison

School 2

$M = .00, SD = .58$
 $t(56) < 1.00, p = .96$ vs. Comparison

Figure 1. Observed changes from baseline in attitudes and practices (relative to comparison sites) in a "high change" (School 1) and "low change" (School 2) CDP program school. (From Watson, M. S., Battistich, V., & Solomon, S. (1997). Enhancing students' social and ethical development in schools: An intervention program and its effects. *International Journal of Educational Research*, 27, 571-586.)

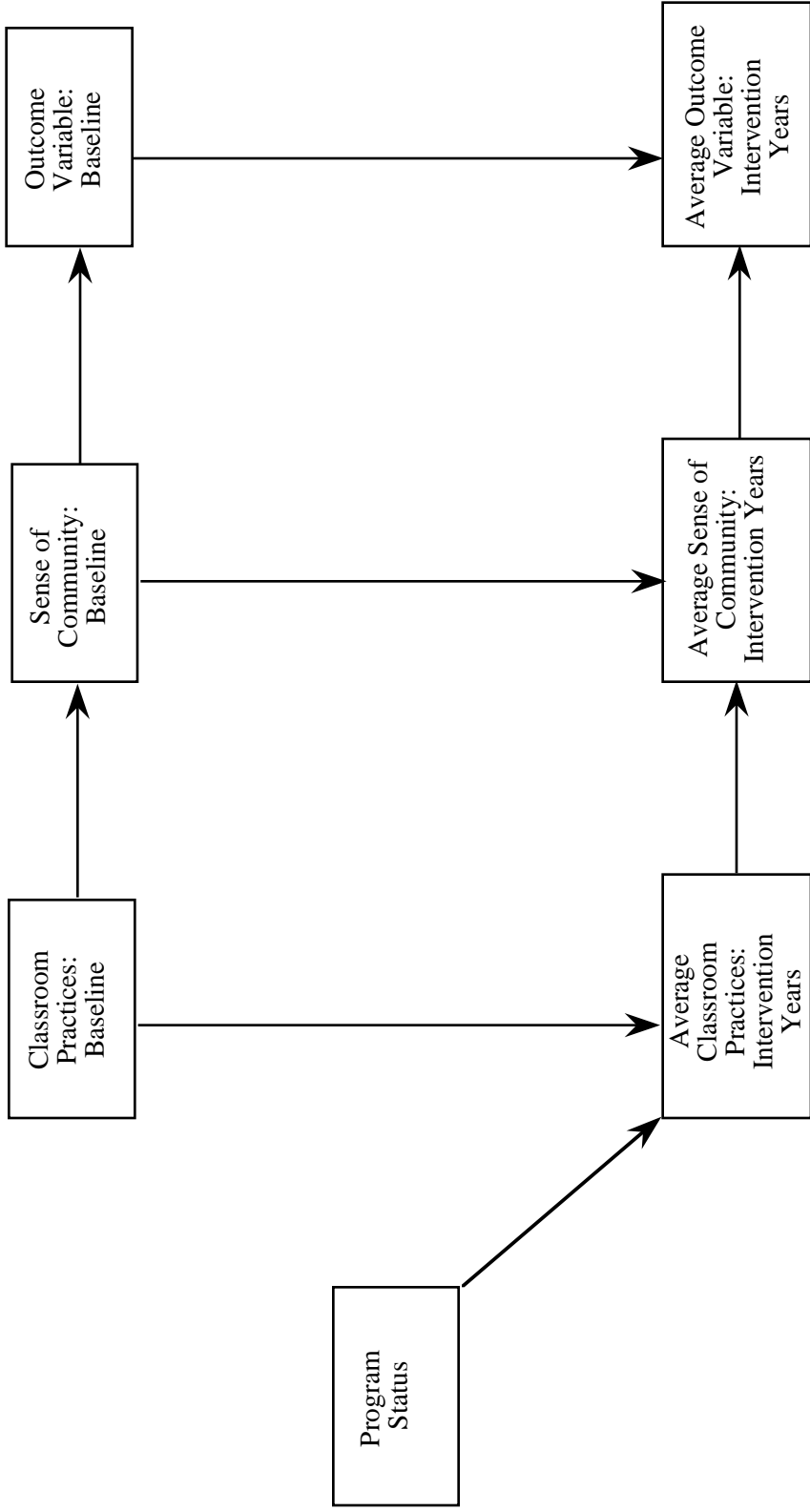


Figure 2. Full Classroom-Level Model for Examining Intervention Effects

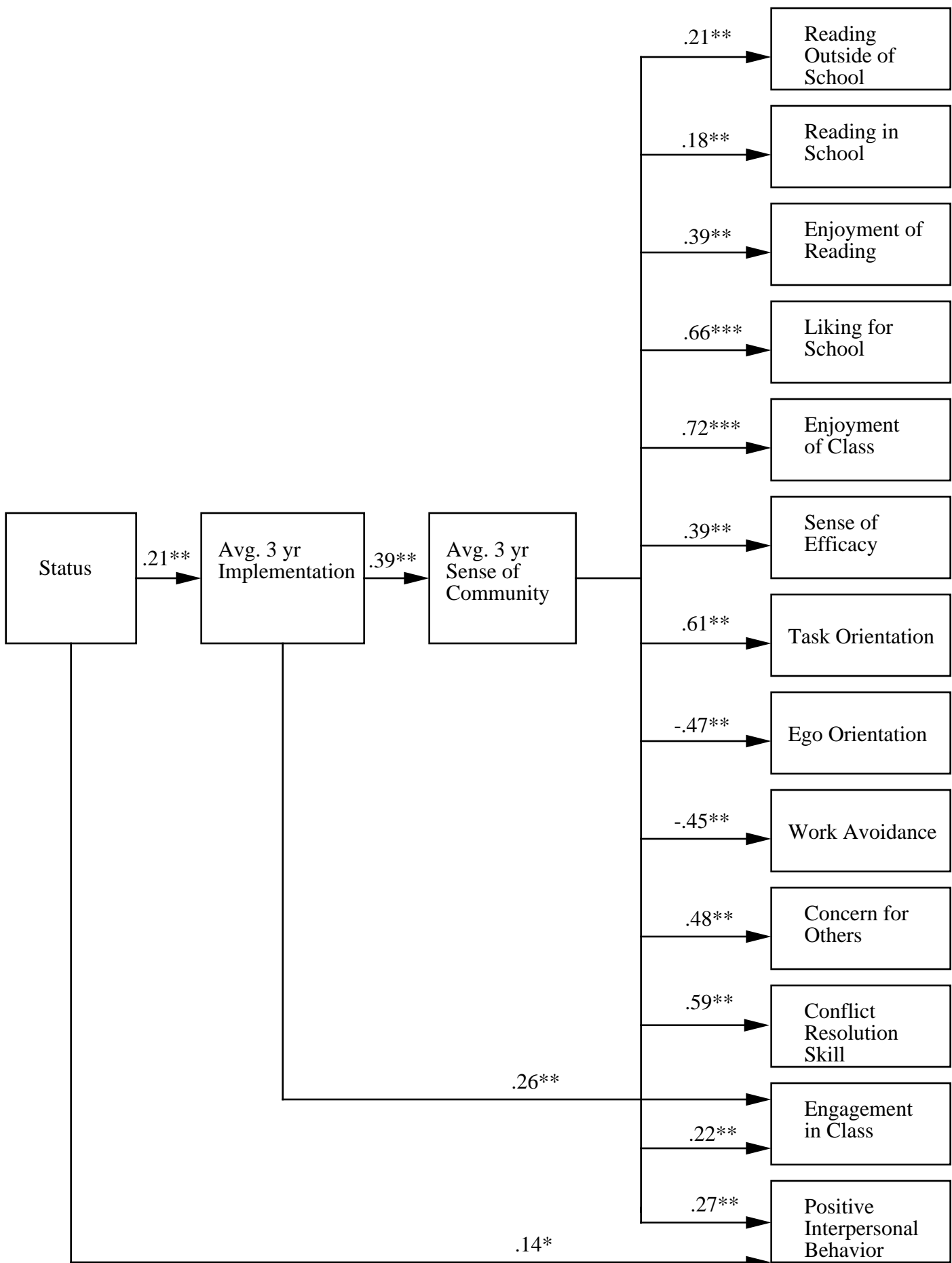


Figure 3. Summary of Classroom-Level Modeling Analyses

(From: Solomon, D., Battistich, V., Watson, M., Schaps, E., & Lewis, C. (in press). A six-district study of educational change: Direct and mediated effects of the Child Development Project. *Social Psychology of Education*.)