

COMPENSATORY EDUCATION
PROGRAM EVALUATION MODEL
AND
FINDINGS

Jean Jolly
Palm Beach County Schools*

CONTEXT

The State Compensatory Education Act of 1977 provided special funding for remedial programs intended to assist students who had been identified by a substandard performance on appropriate assessment tests. Hillsborough County received funding for the 1977-78 school year and again for the 1978-79 school year. In the fall of 1977, Compensatory Education (C.E.) programs in math and communications were implemented in grades 10 and 11. Grade 12 was added in the fall of 1978.¹

When the program was implemented initially, locally developed criterion-referenced tests were used to select tenth and eleventh grade students. State Assessment scores were used for second semester exit from and entry into Grade 11. New criterion-referenced tests (CRT) measuring the same skills

*This paper was developed while the author was working in the Office of Testing and Evaluation, Hillsborough County Schools, Tampa, Florida

1

A.C.E. program was implemented in grades 7-9 in school year 1978-79. However, this paper presents only the findings for grades 10-12.

measured on the State Assessment Test were developed in the 1977-78 school year. The CRTs were used as a screening device for the 1978-79 program. Tenth, eleventh, and twelfth grade students in eleven senior high schools were participating in C.E. classes in January, 1979.

Product evaluation was required by the Department of Education. The funding proposal was to contain a program-based, measurable objective to be assessed in order to determine program accomplishment.

Hillsborough County chose to go beyond the state mandate by implementing a comprehensive, ongoing program evaluation which included process as well as product data. The primary purpose of the comprehensive evaluation was to serve as a guide to decision making by providing "timely" feedback to be used by program planners for ongoing program modification in curriculum, instruction, and management.

The evaluation was considered "ongoing" because it was not restricted to the C.E. program's funding cycle. The evaluation was planned to run, uninterrupted, for the life of the C.E. program in the Hillsborough County Schools.

EVALUATION DESIGN

The term "design" is used to describe the procedures employed in conducting a program evaluation. The design which was being used in the evaluation of the Hillsborough County Compensatory Education Program is cyclical. The cyclical nature of this evaluation design arises from the fact that a complex program was under scrutiny, and a complex system is not easily comprehensible. Questions about an ongoing program arise week by week, month by month, and year by year. Questions asked in one year may go unanswered until the following year. Answers presented to questions in one year may be found unsound when verification is attempted the following year. The evaluation must be continuous, year to year, proceeding in cycles which arise naturally from the process of evaluation inquiry.

In reality a cycle usually produces more questions than answers. The speculations which occur at the end of a cycle result in the formulation of new questions; thus, a new cycle begins. These cycles can be classified according to the mode of consideration to which the data are subjected. The analytic mode of the data-gathering cycle reduces the complexity of the system to relatively simplistic answers to relatively simplest questions as if the whole could be explained simply by aggregating the answers. In contrast, the synthetic mode of the integration cycle reassembles the findings from previously occurring cycles and considers the interrelationship of the findings.

The key concept of this design was the periodic integration of findings. The evaluation question for the integration cycle becomes "In light of what is known about the program to date, are there any tentative conclusions which might be drawn?" These periodic summaries serve as checkpoints for the program planner.

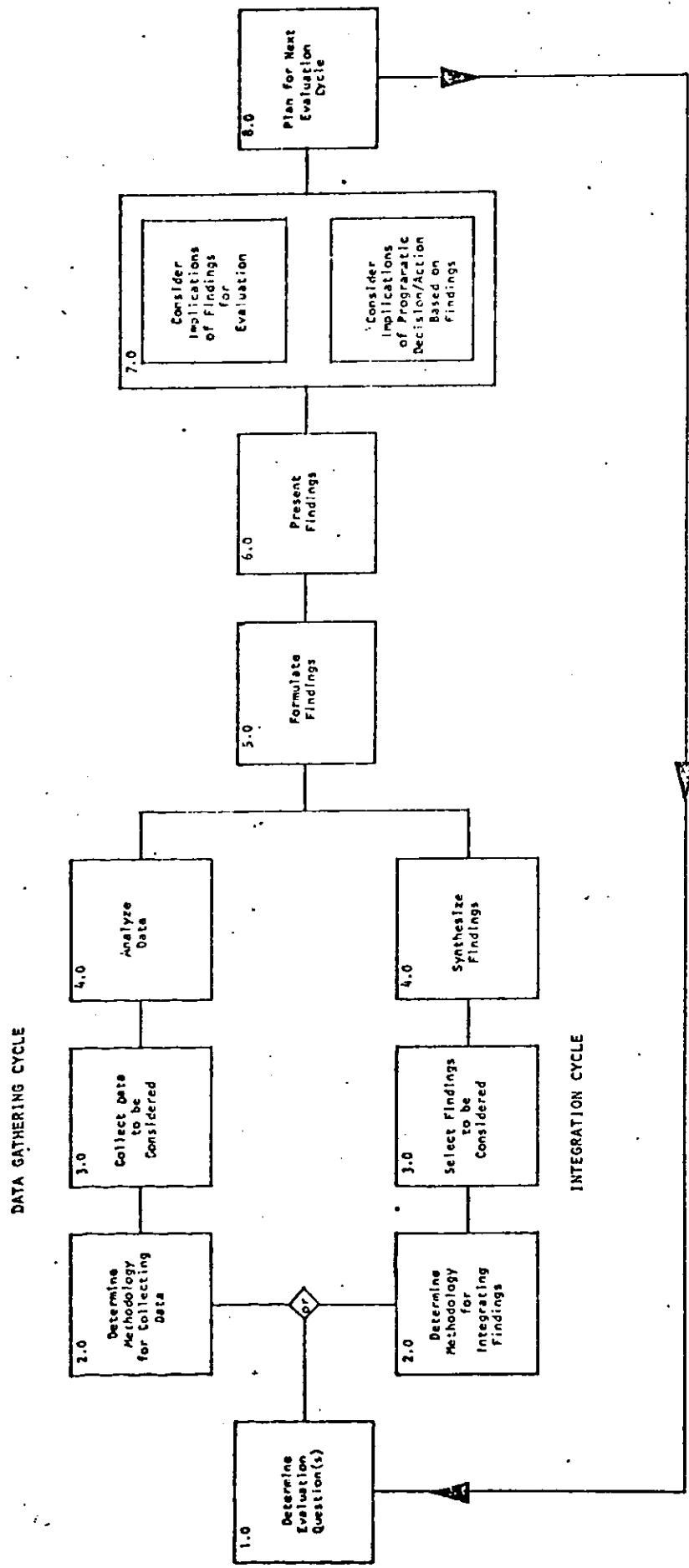


Fig. 1 The Design Used for the Evaluation of the Hillsborough County Compensatory Education Program.

At this point, the procedure used in either cycle is a familiar one. First, findings are formulated (5.0). Findings are then presented to appropriate decision makers (6.0). Of course, findings for and decisions based on an integration cycle are of a more complex nature than those for a data-gathering cycle. The evaluator must then become involved in a conceptual process (7.0) in which implications of the findings and implications of any decisions or actions based on the findings are considered.

The result of this intensive consideration is the germination of a new evaluation cycle. From the planning which occurs at the end of a cycle (8.0), a new evaluation question or set of questions is produced (1.0) and a new cycle begins.

In addition, the evaluator considers the need for what might be thought of as a research study related to the findings or decision/actions. In this design a study of this nature is called an "Allied Study."

Findings from an allied study may be integrated later with evaluation findings. Moreover, these findings often have implications for the total school curriculum as well.

FINDINGS

There are four primary uses for the findings of a data-gathering cycle: (1) providing feedback to the program planners for modification in management procedures, (2) providing the focus for a subsequent data-gathering cycle, (3) forming the data base for a future integration cycle, and (4) suggesting

DATA GATHERING CYCLE #1

Time Frame	General Evaluation Question(s)	Methodology	Major Findings	General Plans for Next Evaluation Cycle	
				Programmatic Decision/Action	Dissemination
January, 1978	What are the issues and concerns related to C.E. which exist at present?	Interviews with program planners Study of C.E. documents	a) Teachers state that they need a document containing comprehensive guidelines for C.E. Group Interviews with C.E. teachers in each school	<p>a) Guidelines for Implementation of the Hillsborough County Education Program was developed and disseminated.</p> <p>b) Committee of school-based teachers and administrators was formed for making grading recommendations.</p> <p>c) Planning began for workshops to be conducted in spring and summer for C.E. in-service</p> <p>d) C.E. teachers and school-based administrators see the need for one county-level person for whom coordination of the program is a major responsibility.</p>	<p>General Director of Secondary Education Secondary English Supervisor (Math Supervisor position vacant at that time)</p> <p>b) Teachers are not sure if county grading policy applies to C.E.</p> <p>c) Teachers feel the need for thorough servicing before teaching in the C.E. Program.</p> <p>d) Supervisor of C.E. position was developed for the 1978-79 school year.</p>

DATA GATHERING CYCLE #2

Time frame	General Evaluation Question(s)	Methodology	Major Findings	Dissemination	Programmatic Decision/Action	General Plans for Next Evaluation Cycle
February, 1978	<p>What program modifications have taken place?</p> <p>What is the general attitude of teachers and students toward C.E.?</p> <p>What additional issues and concerns exist?</p> <p>Is there much variation in programs school to school?</p>	<p>Teacher survey</p> <p>Interviews with department heads and assistant principals for curriculum</p> <p>Interviews with teachers and students</p> <p>Data summary of State Assessment scores by school</p>	<p>a) All schools had copies of the Guidelines, but not all teachers had a copy. Also, all teachers did not have a copy of the Minimum Student Performance Standards for Florida Schools.</p> <p>b) The C.E. math program was well ahead of the communications program in material development.</p> <p>c) Both teachers and students felt that the small classes and individualization by pacing were allowing students to experience their first school success.</p> <p>d) The variance from school to school was great both in student population and in program structure; problems were frequently unique to a school; consequently summary data had less utility than school specific data.</p> <p>e) A number of schools had a policy that C.E. students should not be given grades higher than C. However, most teachers felt that an A-F range was appropriate.</p>	<p>General Director of Secondary Education</p> <p>Secondary Supervisors of Math and English</p> <p>Assistant principals for curriculum</p>	<p>a) Both Guidelines and Performance Standards were distributed to all teachers.</p> <p>b) Workshops were planned for development of material for the communications program.</p> <p>c) A C.E. grading policy of A-F was approved.</p>	<p>Obtain school and program specific data on teachers and students as well as total program data.</p>

DATA GATHERING CYCLE #3

Time Frame	General Evaluation Question(s)	Methodology	Major Findings	Dissemination	Programmatic Decision/Action	General Plans for Next Evaluation Cycle
March, 1978	Is there a difference in the attitude of black and white students?	Teacher and student survey	<p>a) A significantly higher proportion of black than white students state that they work "much harder" in C.E.</p> <p>b) A significantly higher proportion of black than white students state that they would be willing to go to summer school.</p> <p>c) Both groups of students choose to remain in C.E. than to return to regular classes.</p> <p>d) Math C.E. students feel that C.E. will help them get a diploma.</p> <p>e) Math C.E. students report more one-to-one instruction and individualization than do communication students.</p> <p>f) Students report that paper-and-pencil instruction predominates.</p> <p>g) Math C.E. students are more likely to self-pace their work than communications students.</p>	<p>General Director of Secondary Education</p> <p>Verification of differences from school to school and program to program.</p> <p>Workshop planned for materials (other than paper-and-pencil type).</p>	<p>a) Workshop planned for materials (other than paper-and-pencil type).</p> <p>b) More C.E. Math than communications teachers stated a desire to continue teaching in the C.E. program.</p> <p>c) Teachers having only one (1) C.E. class were less likely to desire to continue teaching C.E. than teachers having two (2) or more C.E. classes; however, full-time C.E. teachers desired to have 1 or 2 regular classes also.</p>	

DATA GATHERING CYCLE #4

Time Frame	General Evaluation Question (s)	Methodology	Major Findings	Dissemination	Programmatic Decision/Action	General Plans for Next Evaluation Cycle
April-May, 1978	Are there differences from school to school and program to program in availability of material, instructional procedures, program management or other variables?	Classroom audits and observations	<p>There are major differences from school to school and program to program, and class to class.</p> <ul style="list-style-type: none"> a) In one case students work independently, using their CRT maps for program planning while in another case all students work together on all objectives. b) Most math programs use some form of cumulative testing while communication programs test on one skill at a time. c) The teachers In some schools have worked together in developing materials keyed to the State skills while teachers at other schools have relied on materials distributed by the district office. d) In nearly all cases records of student progress were complete. e) Unique problems exist in some schools. 	Supervisors of Secondary Math and English	Supervisors discussed findings with Department Heads in each school.	Analyze test scores to determine if there are differences in achievement school to school and program to program.

DATA GATHERING CYCLE #5

<u>Time Frame</u>	<u>General Evaluation Question(s)</u>	<u>Methodology</u>	<u>Major Findings</u>	<u>Dissemination</u>	<u>Programmatic Decision/Action</u>	<u>General Plans for Next Evaluation Cycle</u>
June, 1978	How did each program and each school score on the CRT posttest? How did each program score on the CTBS?*	Collection and analysis of test data.	<p>a) The math program overall showed a statistically significant gain in CTBS math scores. The null hypothesis of no change in percentile rank was rejected.</p> <p>b) The communication program overall did not show a statistically significant gain in CTBS reading and language scores (.01 level). The null hypothesis of no change in percentile rank was accepted.</p> <p>c) The change in math percentile rank ranged from a loss of 2 percentile ranks to a gain of 7 percentile ranks.</p> <p>d) There are three different "populations" in C.E.: students taking C.E.-math only, students taking C.E. communications only, and students taking both. The "math" group only entered the program at the 28%ile rank. The "communications" only entered with a 17%ile rank in</p>	Supervisors of Secondary Math and English	Supervisors discussed findings with Department Heads in each school	Plan to integrate quantitative and qualitative data to determine if differences in test scores are related to differences in programs.

*CTBS grade equivalent gain was reported to DOE to satisfy the evaluation requirements.

DATA GATHERING CYCLE #5

Major Findings

(continued from
page 9)

- reading and a 15%ile in language. The "both" group entered with a 14%ile in math, a 10%ile in reading, and a 11%ile in language.
- e) In three schools approximately one-fourth of the C.E. communications students scored at the 18%ile rank in reading.
- f) Eleventh grade math students achieved a mean number of 7.0 skills on the October, 1977, State Assessment Basic Skills Test. They achieved a mean number of 9.33 skills on the CRT posttest administered May, 1978.
- g) Eleventh grade math students equalled or exceeded the district State results (mastery level) on 5/13 skills on CRT tested in May, 1978.
- h) All schools had a statistically significant gain in number of objectives (math) achieved (State Assessment Basic Skills Test and local CRT Minimal Skills Test).
- i) C.E. Eleventh Grade Communications students equalled the district State Assessment results (mastery level) on 1/19 skills on CRT tested in May, 1978.

Major Findings

(continued from
page 9)

INTEGRATION CYCLE #1

Time Frame	General Evaluation Question(s)	Methodology	Major Findings	Dissemination	General Plans for Next Evaluation Cycle	
					Programmatic Decision/Action	
July-August, 1978	After all previous evaluation cycle findings are integrated, what statements can be made about the program?		<p>a) Categorization of schools scoring above, at, or below expectations on the math CRT posttest, eleventh grade</p> <p>b) Display of data in matrix form to identify common elements in schools which performed above expectations</p> <p>c) Integration</p>	<p><u>MATH</u></p> <p>a) The use of spiral (cumulative) testing appear related to achievement.</p> <p>b) Some schools are pretesting on all skills (objectives) and are not using the CRT diagnostic maps for individualizing instruction. This does not appear to be related to achievement.</p> <p>c) C.E. workshop participants</p>	<p><u>MATH</u></p> <p>a) The Secondary Math Supervisor General Director of Secondary Education incorporated spiral testing (cumulative) into the program design for 1978-79.</p> <p>b) A C.E. workshop was conducted which included dissemination of evaluation results.</p>	<p>Look at the 1978-79 C.E. program in light of the integrated findings for the 1977-78 program.</p> <p>Verify findings</p> <p>Plan study for normative information (Allied Study #1)</p> <p>Plan study for math objectives 11, 14, and 16 (Allied Study #2)</p>

(continued)

INTEGRATION CYCLE #1
Page Two

Major Findings	Programmatic Decision/Action
<u>MATH</u>	
f) C.E. students have difficulty in taking tests; teaching "test taking" techniques may be related to achievement.	c) The spiral approach was incorporated into the C.E. communications program. d) School Department Heads made many assignments in the 1978-79 C.E. program on a volunteer basis.

COMMUNICATIONS

a) The low gain in mastery of skills may be related to low measured reading achievement of many C.E. students.

b) Teacher attitude concerning C.E. teaching assignment may be related to students' low achievement.

c) Communications may not lend itself to a "skill" approach as well as math does.

ALLIED STUDY #1 Gain Expectation

PROBLEM

In order to draw conclusions about program "success", the evaluator integrated findings from five data-gathering cycles. The starting point for this integration was the ranking of schools according to "gain scores". At the time there appeared to be no reliable alternative. A method for setting reasonable expectation of gain in number of objectives or in total items, pretest to posttest, was needed.

STUDY

A computer program is being developed, the output of which will provide answer to the following questions for subject and grade level by school and district.

1. At each pretest level, what is the average gain made of students pretest to posttest (Part A)?
 2. At each pretest level, what is the percentage of students scoring at each posttest level (Part B)?
 3. What is the average gain (objective raw score) pretest to posttest (Part C)?
 4. What is the percentage of students scoring at each posttest level (Part D)?

Posttest # Objectives Mastered			
	0 1 2 3 4 5 . . . X	1R	X
Pretest # Objectives Mastered			
1			
2			
3			
4	(B)		(A)
5			
.			
.			
.			
.			
.			
X			
	N		GX
	(D)		(C)

X = The highest objective number on CRT
 TN = Total Number of students per row
 \bar{X} = Mean of each row
 N = Number of students per column
 GX = Mean of Column \bar{X} .

ALLIED STUDY #2 Math Objectives

PROBLEM

The high school math Compensatory Education Program judged to be most successful showed moderate to high CRT posttest level of mastery (72%-98%) on all skills except #11, #14, and #26 (64%-66%). The district Compensatory Education average on the CRT ranged 63% to 95% on all skills except #11, #14, and #16 (56%-57%). These results paralleled the results on the State Assessment Test which measured the same skills.

STUDY AND FINDINGS

The fifteen items measuring skills #11, #14, and #16 were administered to 1,240 students in six levels of senior high school mathematics. It can be seen in Table 1 that low level of mastery exists at higher mathematics levels as well as in Compensatory Education.

After further investigation, two possible sources of this problem were uncovered: (a) students who enter accelerated math programs in the seventh grade have no instruction in the three skills as part of the formal curriculum throughout junior and senior high school, and (b) percents are listed in the back of the sixth grade and junior high school math books.

TABLE 1
Percentages of Students
Not Achieving Math Skills

Skill Number Description	C.E. MATH	GENERAL MATH CAREER MATH PRE-ALGEBRA	ALGEBRA 1 CONSUMER MATH	GEOMETRY	ALGEBRA 2 COLLEGE REVIEW	ALGEBRA 3 TRIGONOMETRY MATH ANALYSIS	MATH & CALCULUS
Skill #11 (subtract 2 mixed numbers)	46%	37%	28%	34%	16%	.08%	.05%
Skill #14 (divide decimal numbers)	44%	27%	24%	24%	2%	.05%	.05%
Skill #16 (Multiply a decimal number by a percent)	42%	25%	31%	71%	.08%	.05%	.02%

ACTION

Senior high schools periodically introduced "review" lessons in the three objectives in the algebra and geometry classes. Junior high school math department heads have been instructed to see that the three skills are taught in all math classes. The elementary supervisor has held one workshop on percents for sixth grade teachers.

ALLIED STUDY #3
Mastery Stability

PROBLEM

Compensatory Education teachers had noted that "mastery" of skills was not always longlasting and that the CRT diagnostic map did not always reflect a student's current competence.

If a study can "master" a skill one month but not the next, one must ask questions about the utility of a diagnostic-prescriptive test.

STUDY

Consideration of C.E. pretest and posttest scores will result in the following table:

TABLE 2

**Percentages of Students Maintaining Mastery
and Not Maintaining Mastery by Skill**

Skill #	Pretest Mastery	Posttest Mastery	Pretest Mastery	Posttest NonMastery
Skill #	% of Cases		% of Cases	

Results will be analyzed to determine if pretest mastery to post-test nonmastery exists in any substantial percentage of cases and if so, whether the mastery to nonmastery is common to all skills or primarily restricted to certain skills.

DATA GATHERING CYCLE #6

Time Frame	General Evaluation Question(s)	Methodology	Major Findings	Dissemination	Programmatic Decision/Action	General Plans for Next Evaluation Cycle
September-October 1978	<p>a) Do teachers and administrators agree with the integrated findings for the 1977-78 program?</p> <p>What are the current issues and concerns related to C.E.?</p>	<p>a) Interviews with representatives of administrators involved in C.E., counselors, department heads and C.E. teachers</p> <p>b) Open-ended questionnaires for C.E. teachers</p>	<p>a) 83% of the C.E. teachers find the CRT profile map inadequate as the base of a diagnostic, prescriptive instructional program. These teachers teach all skills to all students, reteaching skills numerous times until the degree of retention is deemed satisfactory.</p> <p>b) C.E. math teachers suspect that the local screening test, which is limited to computational exercises, does not identify a large group of students who will fail the SSAT-Part II.</p>	<p>Assistant Superintendent for Instruction.</p> <p>General Director of Secondary Education</p> <p>Supervisor of C.E.</p>	<p>a) Diagnostic maps from the March screening test will not be transferred to C.E. teachers in the fall.</p> <p>b) New procedures for documentation are being considered.</p> <p>c) Program planners are to re-emphasize the inclusion of all minimum performance standards in the C.E. curriculum.</p> <p>d) C.E. teachers and school based administrators view the procedures required for documentation of skill mastery excessively time-consuming as are the procedures required for screening new entries in the school for C.E.</p> <p>e) Most teachers and administrators found the integrated findings to have reasonable validity.</p>	<p>Obtain school and program specific data on teachers, students and administrators.</p> <p>Plan study on stability of "mastery" scores (Allied Study #3)</p>

DATA GATHERING CYCLE #7

Time Frame	General Evaluation Question(s)	Methodology	Major Findings	Dissemination	Programmatic
					Decision/Action
November, 1978	How do C.E. teachers describe their own C.E. program? How do C.E. students describe their own C.E. program?	Teachers Interview Student Group Interview	<p>a) Teachers and students do not have a common view of the purpose of C.E.; the two views most often expressed are "to pass the state test" and to strengthen math and communications skills.</p> <p>b) C.E. teachers must have a thorough knowledge of content because in many cases enabling skills must be addressed as prerequisite to direct teaching of minimal skills. Students appear to be aware of teachers' knowledge in the content area.</p>	(Planned for) General Director of Secondary Instruction Secondary English and Math Supervisors Principals and Assistant for Curriculum	<p>Repeat this cycle each month from January through May.</p> <p>Introduce an Integration cycle when appropriate.</p> <p>d) Students complain of boredom with paper-and-pencil activities. Teachers state that audio-visual aids and instructional games help break the monotony.</p> <p>e) Because many schools have three grade levels in a C.E. class, it will be difficult to talk about a tenth, eleventh, or twelfth grade C.E. program.</p>

DATA GATHERING CYCLE #8

<u>Time Frame</u>	<u>General Evaluation Question(s)</u>	<u>Methodology</u>	<u>Major Findings</u>	<u>Dissemination</u>	<u>Programmatic Decision/Action</u>	<u>General Plans for Next Evaluation Cycle</u>
December-January 1978-1979	How did each program and school score on the SSAT-I and II and on the CRT pretest?	Analysis of SSAT and CRT pretest scores	a) The distribution of CRT pretest scores is relatively normal in some programs, positively skewed in some and negatively skewed in others. b) Few students would exit the math C.E. program (grade II) at the end of first semester, but a large number of students who failed math Part II would enter the program.	Assistant Superintendent for Instruction General Director of Secondary Education Supervisor of Compensatory Education	The Testing and Evaluation Office was asked to study the question of whether a math functional literacy test would predict which students were likely to fail SSAT-Part II in Math.	Examine the 1978 SSAT scores in light of the evaluation findings for the 1977-78 Tenth and Eleventh grade C.E. Program. Plan allied study on math, Part II failures. (Allied Study #4)

Plan to integrate quantitative and qualitative data to determine if differences in test scores are related to differences in programs.

ALLIED STUDY #4
Prediction of Functional Literacy Math Failures

PROBLEM

A Compensatory Education teacher suggested that a test measuring math skills A-N was not appropriate for predicting which students would fail the functional literacy (O-X). Furthermore, it was suggested that no "computational" test would be appropriate because the student did not have to operationalize. Consequently, many students who score well on computations but not on "word problems" are not screened into C.E., and many of these same students might be predicted to fail SSAT-Part II math.

After the 1978 SSAT scores were analyzed, there were found to be 2,170 failures on the SSAT-Part II in math. Of the 2,170 failures, 549 students had not been placed in C.E. because their screening score was too high; of the 2,170 failures, 731 had been placed in C.E. on the basis of their screening test score. Students without spring screening scores equalled 445.

The question became, would the 549 students who passed the computational type screening test have failed the local functional literacy CRT? As a preliminary study, it was decided to select one high school and compare scores on the locally developed CRT functional literacy pretest with scores in SSAT-Part II (math) given two weeks later after the pretest was administered to eleventh grade students.

STUDY

A comparison of math functional literacy pretest scores (given two weeks earlier) with SSAT-Part II scores showed a strong relationship between the two tests. One school was selected for the study ($N=77$ eleventh grade students). The correlation between the two tests was .617.

ACTION

The subsequent decision was to use the local math functional literacy test as a C.E. screening device for ninth and tenth grade students in March, 1979.

INTEGRATION CYCLE #2

<u>Time Frame</u>	<u>General Evaluation Question(s)</u>	<u>Methodology</u>	<u>Major Findings</u>	<u>Dissemination</u>	<u>Programmatic Decision/Action</u>	<u>General Plans for Next Evaluation Cycle</u>
January, 1979	After all previous evaluation cycle findings are integrated, what statements can be made about C.E.?	Schools ranked according to retrospective evaluation of 1977-78 C.E. programs in light of 1978 SSAT scores.	(In process)	Description of each school program is being developed including CRT pre-test scores, SSAT scores, teacher questionnaires data, teacher and student interview data, and retrospective evaluation of 1977-78 program.		

topics for allied studies. There are three primary uses for the findings of an integration cycle: (1) providing feedback to the program planners for curricular and management decisions, (2) providing the focus for a subsequent data-gathering cycle, and (3) suggesting topics for allied studies.

By January, 1979 the Compensatory Education Program Evaluation was in its second integration cycle. Eight data-gathering cycles had been completed. Two allied studies had been completed, and two were in progress.

The pages which follow provide a brief description of the application of the evaluation design in a specific C.E. Program. General Evaluation questions, methodologies, major findings, etc., are presented. Also included are four examples of allied studies.

SUMMARY

The examples presented herein should be considered from the standpoint of their utility to the decision makers in Hillsborough County as they apply to the Compensatory Education Program there. There is little possibility that the findings would generalize to other districts.

What will generalize, however, is the evaluation design. Since program evaluation has been mandated, some model for evaluation is appropriate. The model used in Hillsborough would appear to be applicable to almost any C.E. program. The primary value in such a model is its cyclical nature which supports the generally accepted notion that evaluation is not only a process but an ongoing process.