The Costs of Testing in American Public Schools

D. Dorr-Bremme
J. Burry
J. Catterall
B. Cabello
L. Daniels

CSE Report No. 198
1983

Center for the Study of Evaluation
Graduate School of Education
University of California, Los Angeles

The research reported herein was supported in whole or in part by a grant to the Center for the Study of Evaluation from the National Institute of Education, U.S. Department of Education. However, the opinions and findings expressed here do not necessarily reflect the position or policy of NIE and no official endorsement should be inferred.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

INTRODUCTION

The Test Use Project in Overview

A broad range of educational testing issues confronts policy-makers at all organizational levels. One of these issues is the costs of testing. Federal, state, and local agencies -- together with professional and advocacy groups representing practitioners, parents and students, and test developers -- need to be aware of the costs and implications of different kinds of assessment practices and programs.

Over a three-year period from December 1979 to November 1982, CSE's Test Use Project gathered a large body of information on testing practices, uses and impacts, and costs in public schools across the nation. We focused on:

- achievement testing in reading/English language arts and mathematics,
- testing of these types as they occur in public schools at the upper elementary and high school levels, i.e., in grades 4-6 and 10-12,
- testing practices, uses and impacts, and costs within schools.

With the intention of informing the decisions that a variety of stakeholders may need to make about testing, our research followed from broad definitions of tests and testing. Our inquiry encompassed a wide range of formal assessment measures (e.g., commercially produced norm- and criterion-referenced tests and curriculum-embedded achievement measures; tests of minimum competency or functional literacy; district-, school-, and teacher-developed tests), as well as less formal assessment techniques (e.g., teacher's observations of and interactions with students in class).

The Test Use Project was conducted in two phases. Phase I research elicited a representative picture of achievement testing. Phase II, the subject of this report, explored the costs of testing. In this report we first present an overview of our Phase I research to provide a context for considering the time and effort associated with testing. Then we discuss the details about the direct and indirect costs, gained from Phase II research, associated with that effort. A description of the design for the three-year study follows.

Research in Phase I

Phase I research lasted two years, from the project's start-up to November 30, 1981. A planning year included a literature review, exploratory field work in three school districts, and re-analyses of data from an earlier CSE study of testing (Yeh, 1978). The results of this year's work directed us in following three central questions of the study and informed the survey design to address them:

1. With what frequency and distribution are tests given in the upper elementary grades and high school?

2. In what ways do tests and testing impact on schools and those within them,

   (a) through their very presence, as required or recommended,
   (b) through utilization of their results?

3. What factors influence:

   (a) where and how much particular types of testing are done?
   (b) the ways that types of tests, testing, and test score use impact upon schools and those within them?

The desired nationally representative picture of testing drew upon a probability sample of 114 school districts stratified by

geographical region, locale, SES, school district size and minimum competency testing policy. We obtained data from teachers and principals representing 91 of the selected school districts. The teachers consisted of fourth and sixth grade teachers reporting information on their testing practices in reading and math, and tenth grade teachers reporting their testing practices in English or math.

On the basis of the national survey, with strong substantiation from the fieldwork, the following picture of tests and testing appeared:

The typical elementary school enrolls about 500 students who are ethnically mixed but with caucasians in the majority. A significant minority of students receive federal aid and/or qualify for free lunch benefits. Almost two-thirds of the schools operate a school improvement program, and student achievement testing is typically required in these programs. More than half of the schools have minimum competency testing requirements. Secondary enrollments are typically around 1440 students. While other characteristics were quite similar to those at elementary schools, students in the typical high school appeared a bit more economically advantaged.

Three or four reading and math tests are typically required at the elementary and secondary levels. Most schools have district-required testing; school-required tests are administered in about 40 percent of the schools.

In making various school-level decisions -- curriculum needs, student assignments, resource allocation, student promotion/graduation, public information/parent communication -- principals seem to

consider a wide range of test and other information. Though none of these sources dominates, teachers' opinions and recommendations carry more weight than test results.

In general, tests results seem to have greater impact and wider consequences in lower SES schools than in higher SES settings.

The fourth or sixth grade elementary student is likely to spend about 10 hours a year on reading tests and somewhat more than 12 hours a year on math tests. The tenth grade student appears to spend more than 26 hours a year on English tests and about 24 hours a year on math tests. These figures include only time for taking tests -- and at that only in language arts and math -- and not the time spent preparing for the testing event or for scoring, recording, etc. after the test is given. As will be seen in a later section of this report, these test-taking figures increase significantly when the times associated with routine test-related activities are considered. The specific kinds of tests used, and the percentages of the total time devoted to testing in language arts/reading and math that they consume, appear in Table 1. Tables 2 and 3 show how the elementary and secondary teachers tend to use the various kinds of assessment devices they administer for planning, grouping, re-grouping, and grading. For these decisions teachers tend to rely heavily on their own and colleagues' judgment, and on commercial and teacher-constructed curricular measures.

For a more detailed treatment of the Phase I findings, see Lazar-Morrison et. al. (1980), and Burry et. al. (1981, 1982). For a

Table 1
Types of Test Used,
As a Percentage of the Total Time
Devoted to Testing

TYPE OF TEST	Elementary Teachers		10th Grade English Teachers	10th Grade Math Teachers
	Reading	Math		
Tests which form part of a statewide assessment program	3	3	3	1
Required minimum competency test	1	2	1	1
Tests included with curriculum materials	28	35	8	17
Other commercially published tests	17	18	6	3
Locally developed and district adopted tests	13	8	5	2
School or teacher developed tests	37	35	74	76

picture of teachers' reasoning processes as they use these tests, see Dorr-Bremme, (1983).

Phase II: Overview to The Costs Study

The cost study was designed to obtain an estimate of the direct and indirect monetary, opportunity, and psychological costs of testing in schools and districts.

Everything that we had discovered suggested that we would need considerable on-site work -- ongoing observation and interviewing conducted proximal to and focusing on particular assessment events -- to be able to locate and estimate important testing costs.

We considered four possible frameworks for analyzing testing costs: (1) cost accounting, which identifies costs and evaluates their magnitude; (2) cost-effectiveness analysis, which examines and evaluates costs and measures benefits in units (not necessarily

Table 2

Elementary Teacher Use of Assessment Information for Different Decision-making Purposes
(Percentages reporting use of this information as crucial or important for the specified purpose)

Source/Kind of Information	Planning Teaching at Beginning of School Year		Initial Grouping or Placement of Students		Changing a Student from One Group or Curriculum to Another		Deciding on Students' Report Card Grades	
	Reading	Math	Reading	Math	Reading	Math	Reading	Math
Previous teachers' comments, reports, grades	57	52	62	55	x	x	x	x
Students' standardized test scores	57	54	57	51	55	53	17	16
Students' scores on district continuum or minimum competency tests	51	47	50	45	45	39	20	18
My previous teaching experience	94	94	x	x	x	x	x	x
Results of tests included with curriculum being used	x	x	78	67	83	82	75	77
Results of other special placement tests	x	x	61	56	x	x	x	x
Results of special tests developed or chosen by my school	x	x	x	x	56	52	42	42
Results of tests I make up	x	x	80	86	78	85	92	95
My own observations and students' classroom work	x	x	96	97	99	99	98	98

Table 3

High School Teacher Use of Assessment Information for Different Decision-making Purposes
(Percentages reporting use of this information as crucial or important for the specified purpose)

Source/Kind of Information	Planning Teaching at Beginning of School Year		Initial Grouping or Placement of Students		Changing a Student from One Group or Curriculum to Another		Deciding on Students' Report Card Grades	
	Reading	Math	Reading	Math	Reading	Math	Reading	Math
Previous teachers' comments, reports, grades	28	29	34	40	x	x	x	x
Students' standardized test scores	47	29	49	30	62	39	12	8
Students' scores on district continuum or minimum competency tests	48	30	47	36	53	36	9	5
My previous teaching experience	99	97	x	x	x	x	x	x
Results of tests included with curriculum being used	x	x	45	35	58	43	44	31
Results of other special placement tests	x	x	42	26	x	x	x	x
Results of special tests developed or chosen by my school	x	x	x	x	50	31	28	34
Results of tests I make up	x	x	87	77	92	91	99	99
My own observations and students' classroom work	x	x	99	93	99	97	99	95

monetary) appropriate to the specific testing program; (3) cost-benefit analysis, which identifies each cost and benefit and then assigns (exclusively) dollar values to each; and (4) an economics of information paradigm, which estimates the proportion of resources that can justifiably be spent in the acquisition of information.

The more complex models -- cost effectiveness, cost benefit, and the economics of information paradigm -- were considered inapropriate at this early stage in the development of research on the costs of testing. A cost-effectiveness analysis would have required that we develop both a measure of a testing program's effectiveness and a total cost figure expressed in some appropriate unit. But the costs and benefits of testing are multiple and not directly comparable, and until a single total of costs can be associated with the effectiveness of the test or tests under scrutiny, the model is not strictly applicable. Cost-benefit analysis would have requireed the incorporation of cost and benefits in exclusively dollar terms. This requirement would apply to all testing costs, some of which have no conceivable dollar equivalents, and so we did not view cost-benefit analysis as a likely means of yielding useful insights. The economics of information paradigm would have presented even greater hurdles. In place of converting benefits and costs to dollar equivalents, this model would require benefits and costs to be directly associated with impact on pupil outcomes, including achievement. Relating elements of testing to schooling outcomes would have been problematic because both the costs and benefits of testing are likely to be difficult to define and their links to pupil outcomes may be remote.

Given these foregoing problems, we chose the cost accounting model for our initial research on testing costs, with the intention of (1) identifying the costs associated with testing for selected schools and districts, and (2) evaluating the magnitude of costs associated with testing for those selected schools and districts.

Summary of Methods

Because we had previously collected test-use data in both elementary and high schools, continuity suggested that we mount the costs study at these same grade levels. Phase II resources, however, were insufficient to fully examine testing costs at both school levels or even at the high school level alone, where we had already discovered a much greater variation in teachers' practices than among elementary school teachers. Our decision, therefore, was to focus our cost study on elementary school practice.

Two elementary schools were selected for study. One of these schools was an inner city elementary school which is part of a large metropolitan school district. The student population of this school was comprised predominantly of minority students of lower socioeconomic standing. This school participates in a large number of federal, state, and district funded programs, many of which require achievement testing. The second elementary school selected, in contrast, was part of a school district in a small suburban town. This school participates in no categorically funded programs and its student population consists largely of Asian and White, middle class students.

At the district level, we collected data on monetary costs of

basic skills by examining relevant district documents and in discussions with appropriate district officials. To determine opportunity costs at the district level, we conducted interviews with key personnel involved in basic skills testing and the use of test results.

At the school level, we collected information on the monetary and opportunity costs associated with all achievement testing via formal, comprehensive interviews with the building principals, instructional staff, and school specialists and resource personnel. We also conducted supplementary observation of testing in classrooms. Both procedures --- the comprehensive interviews and observation of testing events -- were also used to identify the psychological costs of testing for the schools' instructional staffs. Formal student interviews, supplemented by the classroom observations, provided the data base for estimating the psychological costs of testing for students in each school.

In the elementary school in Hillview, the small suburban district, we interviewed the building principal, all eleven teachers, and the single resource specialist who ran and taught in the school's learning laboratory. We conducted testing event observations in 2 classrooms at grades 2 and 5, and interviewed ten students each from grades 4, 5, and 6.

In the elementary school in Cityside, the large metropolitan district, we interviewed the building principal, sixteen teachers, three other administrators (special program coordinators), and two educational specialists. In addition, we conducted testing event observations in several classrooms, and interviewed ten students each from grades 4, 5, and 6.

MONETARY COSTS OF BASIC SKILLS TESTING IN TWO DISTRICTS

This section describes the basic skills testing practices in the two districts surveyed. We begin by providing district background information and the results of our data collection, followed by a profile of each district and its basic skills testing program, and then discuss the costs related to these testing programs. To facilitate comparisons and to help inform various policy issues, we discuss testing costs at the central district level and those incurred district-wide separately before attempting to construct overall cost totals. Following discussions of the two districts, a third section draws some limited comparisons from the two sets of data.

Case I: Littleton District

Littleton is a small, suburban district with four elementary schools, a junior high school, and a senior high school. The district is highly decentralized: the small central office has two certificated officials plus minimal support staff, and the six Littleton schools autonomously reach many decisions including some regarding their testing programs. Littleton's community is relatively stable, but has witnessed both an overall enrollment loss in recent years and a steady growth in Asian student population. A variety of data describing Littleton are presented in Table 4.

Table 4
Littleton District

Total Enrollment (1982-83 average daily attendance)		3354 pupils
High School (10 - 12)	1060 pupils	
Junior High School (7 - 9)	915 pupils	
4 Elementary Schools (K - 6)	1379 pupils	
Total Budget		$ 5.6 million
Per Pupil Spending		$ 1670
Other Significant funds		
Title I (Chapter I, ECIA)	$ 40,000	
PL 94-142	$ 40,000	
Percent Minority Pupils (Predominantly Asian) (range is 5% to 50% in elementary schools)		18 %
Number of Teachers		130

Littleton District's Testing Program

The schools administer an array of tests which meet both their own demands for pupil information and also various state mandates requiring particular tests at various grade levels. Because the district is small, there is no full-time testing coordinator in the central office nor anyone centrally responsible for testing. Responsibility for test coordination is a part-time responsibility of a counselor at the high school and at the junior high, and is one of

the principal's responsiblities at the elementary schools. Table 5 summarizes the basic skills testing activities in Littleton District, by type of test and grade level.

Table 5

Summary of Littleton District Basic Skills Testing

Level	Test	Basic Purpose
Elementary	Stanford Achievement Test	Cum records
	SRA Assessment Survey	Cum records
	Grade 4 Proficiency	State Required
	State Assessment (Grades 1,3,6)	State Required
	Metropolitan Achievement Test	Title 1 Evaluation
Junior High	SAT	Counseling/Curriculum review
	Gates MacGinitie	Placement
	Metropolitan Math	Placement
	L.A. County Proficiency (7,9)	State Mandate
Senior High	Differential Aptitude Tests	Counseling
	Iowa Test of Educational Development	Curriculum Assessment/ Counseling
	Strong Campbell	Interest Inventory
	Survey of Basic Skills	State Mandate
	Basic Skills Inventory	State Mandate (Required for Graduation)

The Costs of Testing in Littleton District

The principal goal of our research was to identify the various ingredients of the basic skills testing activities of the district, to estimate the magnitude of each of these activities, costs in their primary units (such as teacher or counselor hours devoted to testing, or direct dollar costs of materials and services purchased), and finally to convert all resource estimates to dollar equivalents. The overall level of resources committed to basic skills testing has meaning when compared to the total resources available to the district. From instructional and service standpoints, the time devoted to testing by pupils, teachers, counselors, administrators, and support staff may be important in the context of the overall allocation of time among tasks for district personnel.

We interviewed district personnel at all levels to identify the types of tests administered and the full range of district resources attached to their basic skills testing. We probed the nature of test administration, pretest and posttest activities of personnel, various analysis and dissemination activities at the classroom, school and central office levels, and the types of materials and services purchased from outside vendors. After achieving a base picture of Littleton's testing, we surveyed district personnel -- an assistant superintendent, principals, counselors and teachers -- to generate estimates of dollars expended or time involved in testing activities. Table 6a presents a summary of th types of costs identified, and the actual estimates for each of these costs in theirprimary units.

It is apparent that basic skills testing is not a major activity in Littleton's central office, since such testing only occupies between 3% and 8% of work time for central staff. Table 6a displays this finding as fractions of time spent on all testing matters by three individuals at the central level--the assistant superintendent, a program coordinator, and a clerical staffer. Because they are responsible for all testing functions, none of the respondents was able to suggest a finer breakdown of time, such as significant allocations to one particular test or to testing at particular grade levels.

Table 6a

Littleton District Central Office Testing Costs

Personnel	Time Allocation
Assistant Superintendent	5% FTE
Special Programs Coordinator	3% FTE
Secretary	8% FTE
Purchases	No Central Purchases Reported

The central school-level costs displayed in Table 6b refer to those testing costs above the classroom level at the six schools in the district. At the elementary schools, these costs are for the time of principals and clerical staff at each school; at the junior high, costs included the time of a counselor who is responsible for test coordination and who is assisted by clerical staff, and also some additional dollar costs for scoring services and materials; at the

high school level, costs included counselor time, clerical staff time, and material and service purchases.

The classroom level costs reported in Table 6c include the hours teachers devoted to testing and the amount of pupil time spent in testing. One apparent fact of Littleton basic skills testing is that time spent in district-mandated, basic-skills testing appears to be rather negligible at the high school levels in comparison to the earlier grades.

Table 7 essentially replicates Tables 6a, 6b and 6c, except that Table 7 converts district personnel time devoted to testing to dollar equivalents. This was done by applying estimated annual personnel cost figures for each category of staff involved in testing (teachers, principals, administrators, counselors, and clerical staff), and then estimating the value of the time devoted to testing by each staff category as an appropriate share of their annual cost to the district. Table 7 thus presents dollar estimates for the costs of each test, at each level, and affords some detail in showing just where these costs occur. For instance, the SAT test in the elementary schools commands the personnel resources of principals ($692), clerical staff ($3694), and teachers ($7776). This can be contrasted with the 4th grade proficiency test which engages comparatively few resources in its administration and handling (clerical costs of $77 and teacher costs of $518). Many similar comparisons can be drawn with these data.

Pupil time that was shown in Table 6 was not converted to dollar estimates, and so does not appear in Table 7. Since pupils do not

Table 6b

Littleton District School Level Resources for Testing

	Test	Grades Admin.	Principal Hours	Counselor Hours	Clerical Hours	Purchases
Elementary Schools	SAT	1,2,3	48	-	384	$0
	State Asses.	3	36	-	40	0
	Profic. 4	4	0	-	8	0
	Profic. 6	6	4	-	0	0
	Subtotals (Elem)		88	-	432	0[1]
Junior High School	SAT		0[2]	0	0	0
	Gates/Mac		0	0	Pre- 14 Post- 1	$40[3]
	Metro/Math		0	Pre- 90 Post-60	0	0
	Profic.			Post-60	Pre- 10 Post-48	$1800[4]
	Subtotals (Jr. High)		0	210	73	$1840
High School	Diff. apt.		0[2]	Pre- 4 Post-25[5]	Pre- 2 Post- 4	$500[4]
	SBS		0	Pre- 3 Post- 5	Pre- 3 Post- 5	0
	BSI		0	Pre- 4 Post-10	Pre- 10 Post-10	900[6]
	Subtotals (High School)				34	$1400
Grand Totals			88 hrs.	261 hrs.	539 hrs.	$3240

Notes: [1] No respondents reported awareness of specific purchases for elementary school tests. Unknown costs of answer sheets and replacements books are buried in supplies budgets and not reported here.

[2] Junior high and high school principals delegate testing matters to a counselor.

[3] Costs of replacement books.

[4] Costs of scoring services.

[5] These hours are allocated @ 20 hours for student conferences and 5 hours for parent communications.

[6] Costs of answer sheets and scoring services.

Table 6c
Littleton District Classroom Level Resource Testing

	Test	Hours Per Teacher Admin.	Other	Total	Number of Classes	Total Hours	Pupil Hours/ Pupil[1]
Elementary Schools	SAT	18	12	30	24	720	18.0
	one lab teacher		30	30		30	
	State Assess. (3)	6.5	8	14.5	8	116	6.5
	Profic/4 (4)	4	2	6	8	48	4.0
	Profic/6 (6)	2.5	2	4.5	8	36	2.5
	Subtotal (4 elementary schools)					950	
Junior High	SAT	13	3	16	50	800	13
	Gates/Mac	7.5	4.5	12	29	348	7.5
		(plus estimated total pre-test time)				10	
	Metro Math	2.5	1	35	7	24.5	2.5
	Profic.	9	2.6	11.6	8	93	9
	Subtotal					1275.5 hours	
High School	Differential Aptitude Test (10)[2]					10	1
	Survey of Basic Skills					12	1
	Basic Skills Inventory					6	1
	Subtotal					28	
	GRAND TOTAL:					2253.5 hours	

Notes:

1 Pupil time reported is for test administration. Negligible time before or after tests reported.

2 High school basic skills testing is conducted en mass (i.e., in the auditorium for entire subject population.

Table 7

Littleton District Testing Costs in Dollar Approximations
(Note that this table replicates Table 6 but replaces hour estimates with dollar equivalents)

Central Office Costs	
Assistant Superintendent[1]	$ 2000
Coordinator[4]	$ 750
Secretary[2]	$ 1600
Total	$ 4350

Central School Level Costs

ELEMENTARY (K-6)

TEST	Principal[3]	Clerical[2]	Totals
SAT (1-3)	$ 692	$ 3694	$ 4386
State Assess.(3)	519	385	904
Profic/4 (4)	0	77	77
Profic/6 (6)	58	0	58
Totals	$ 1269	$ 4156	$ 5425

JUNIOR HIGH SCHOOL (7-9)

TEST	Counselor[4]	Clerical	Purchases	Total
SAT	$ 0	$ 0	$ 0	$ 0
Gates/Mac	0	150	40	190
Metro Math	1803	150	0	1953
Profic.	721	557	1800	3078
Totals	$ 2524	$ 857	$ 1840	$ 5221

HIGH SCHOOL (10-12)

TEST	Counselor	Clerical	Purchases	Total
Differential Aptitude Test	$ 349	$ 58	$ 500	$ 907
Survey of Basic Skills	96	77	0	173
Basic Skills Inventory	168	192	900	1260
Totals	$ 613	$ 327	$ 1400	$ 2340

Totals: Principals $ 1269 Clerical $ 5340
Counselors $ 3137 Purchases $ 3240

Total School Central Level Costs: $ 12,986

Notes:

1 Based on $ 40,000* salary and fringes
2 Based on $ 20,000 salary and fringes
3 Based on $ 30,000 salary and fringes
4 Based on $ 25,000 salary and fringes

* Estimates per state-wide averages. Actual Littleton costs may be slightly higher.

Table 7 (Continued)
Littleton District Testing Costs in Dollar Approximations

Classroom Level Costs

ELEMENTARY (K-6)

TEST	Teacher Cost
SAT[1] (1-3)	$ 7776
State Assess. (3)	1253
Profic/4 (4)	518
Profic/6 (6)	389
Total	$ 9,936

JUNIOR HIGH (7-9)

TEST	Teacher Cost
SAT (7,8,9)	$ 8640
Gates/Mac (7,8,9)	4023
Metro Math	265
Profic.	1004
Total	$ 13,932

SENIOR HIGH (10-12)

TEST	Teacher Cost
Differential Aptitude	$ 108
Survey of Basic Skills	130
Basic Skills Inventory	65
Total	$ 303

Totals:
Cost of Teacher Time: $ 24,171 (Based on Grand total of 2253 teacher hours)

engage fractions of the district's budget in the manner that staff involved in district activities do, they do not represent direct or indirect costs to the district that have a meaningful dollar interpretation. Nevertheless, the amount of time pupils spend in various activities can be thought of as having various costs and benefits, particularly those accruing to the effectiveness of the instructional programs of the district.

Table 8 summarizes the preceding cost data in several ways. It shows how the costs of various tests are distributed, the degree to which testing costs are incurred as a result of outside mandates for assessments, the total cost of each test per pupil tested, and the testing cost per pupil at different levels.

The costs per pupil tested for each test and at each level are shown immediately to the right of the total monetary costs. These costs range from a high of $28.33 for the SAT test at the junior high to a low of $0.86 for the SBS at the high school. In addition, the total costs of testing per pupil enrolled at each level are shown at the extreme right of the table. The central office resources devoted to testing translate to $1.30 per pupil district-wide. The junior high devotes the most resources to testing ($20.93 per pupil), and this amount is just slightly under 1.3 percent of the district's $1670 average per pupil expenditure. Overall, it appears that Littleton testing costs amount to about three-quarters of one percent of the overall total of district expenditures.

Table 9 shows what fraction of the testing costs per pupil at each level can be accounted for by direct versus indirect costs. Here

Table 8
Littleton Testing: Costs Per Pupil, and Cost Summary, by Level

Level	Total Monetary Costs Central	Teacher	Total	Total Costs Per Tested Pupil	Costs Per Pupil At Level
Central Office	$ 4350		$ 4350		$ 1.30 per pupil
Elementary					
SAT*	$ 4386	$ 7776	$ 12162	$ 20.27	
State Assess.*	904	1253	2157	3.60	
Prof 4*	77	518	595	2.98	
Prof 6*	58	389	447	2.24	
All Tests	$ 5425	9936	$ 15361		$ 11.13 per Elementary pupil
Junior High					
SAT	$ 0	$ 8640	$ 8640	$ 28.33	
Gates	190	4023	4213	4.61	
Metro	1953	265	2218	3.39	
Prof.*	3078	1004	4082	6.60	
All Tests	$ 5221	$ 13932	$ 19153		$ 20.93 per Junior high pupil
High School					
DAT	$ 907	$ 108	$ 1015	$ 2.88	
SBS*	173	130	303	0.86	
BSI*	1260	65	1325	1.25	
All Tests	$ 2340	$ 303	$ 2643		$ 2.49 per High school pupil

* State mandates

we have included as direct costs those items for which the district incurs an expenditure of funds, such as the cost of test booklets,

Table 9
Littleton District: Direct vs. Indirect Cost of Basic Skills Testing, by Level

Level	Testing Costs Per Pupil	Direct Share	Indirect Share
Central Office	$ 1.30		100%
Elementary	$ 11.13	negligible	100%
Junior High	$ 21.93	9.7%	90.3%
High School	$ 2.49	53%	47%

answer sheets, and scoring services. The indirect costs represent the share of personnel time (or its dollar equivalent) devoted to testing activities. With the exception of the high school testing, it appears that the vast majority of testing costs are bound up in the time of district personnel who administer the tests and who analyze and disseminate the results. In contrast, the high school testing program experiences relatively high direct costs since the activities occupy comparatively few teachers, who are needed for few hours, and at the same time incur comparatively high costs for scoring services.

Some tests administered in Littleton result from the district's own discretionary assessment needs, while others are administered to satisfy state requirements. Table 10 shows the share of testing costs at the elementary, junior high, and high school levels resulting from each of these two types of tests. Again, a contrast is apparent between the high school and lower levels. About a fourth of Littleton

testing below grade 10 is done in response to outside mandates, while more than half of the costs of testing in the high school are tied directly to such mandates.

Table 10
Littleton District: Mandated vs. District Discretionary Testing Costs, by Level

Level	Overall Basic Skills Testing Costs Per Pupil	Mandate Share	Discretionary Share
Elementary	$ 11.13	24.6%	75.4%
Junior High	$ 20.93	21.5%	78.5%
High School	$ 2.49	61.6%	38.4%

Summary Comments: Littleton District Testing Costs

Some overall observations emerge from the profile of Littleton's testing costs.

First, the central office testing costs are minimal--about $1.30 per pupil. Second, the magnitude of testing costs overall is small in comparison to overall resource expenditure, on the order of three-quarters of one percent of the total. And within this small total cost for testing, a generally small fraction is accounted for by direct dollar expenditures for such things as tests, materials, and scoring. As such, from a budgetary standpoint, Littleton's testing occupies a fairly negligible portion of its total resources, and of those costs that are attributable to testing, by far the most important are the costs of teacher and administrator time. This suggests to us that the dollar costs of testing may be less important than other considerations attached to the personnel time that

generates most of those costs, such as effective use of teacher or principal time. Overall, it appears that the testing "budget" per se, even when construed to include personnel time allocations, is not a potential gold-mine should Littleton seek to divert testing resources to other endeavors.

Case II: Metro District

Metro is a major urban school district. The pupil population is diverse, the district maintains hundreds of schools and employs thousands of teachers, and its budget is a complex mix of general support and state and federal categorical programs. Table 11 highlights some of Metro District's salient characteristics.

Table 11
Metro District: Descriptive Data

Total Enrollment (1981-82)		543,791
High School (10-12)	127, 221 pupils	
Junior High School (7-9)	120,337 pupils	
Elementary School (K-6)	291,632 pupils	
Schools for Handicapped	4,601 pupils	
Total Budget		$ 1.84 billion
Per pupil spending includes:		
Basic State Aid per pupil	$ 1,890	
Local revenues per pupil	409	
Federal Programs per pupil	330	
State Categoricals per pupil	320	
Other Revenues per pupil	351	
Student Racial/Ethnic Composition		
American Indian	0.37%	
Asian/Pacific Islander	7.5 %	
Black	22.2 %	
Hispanic	47.4 %	
White	22.5 %	
Number of Schools		
Elementary	427	
Junior High Schools	75	
High Schools	49	
Magnet Schools/Centers	84	
		Average Per
Number of Classroom Teachers	Total	Grade Level
Elementary	9721	1389
Junior High	3539	1180
High School	3742	1247

On average, Metro spends approximately twice as much money annually per pupil than Littleton, but most of the difference is accounted for by the presence of Metro's specially funded programs. The district pupil population is largely non-white, with significant representation from several minority groups.

Metro District's Testing Program

As we found with Littleton, Metro District administers a variety of basic skills tests for different internal and external purposes. The tests administered, at which grade levels, and for which reasons are outlined in Table 12. The largest single testing effort is the skills test given to all children in grades 1 through 6, the Continuum-Based Skills Survey (CBSS). This test was developed by the district over a period of several years and is used primarily to provide teachers with good information about their students' performance. The test also satisfies state and federal reporting requirements for the Chapter I ECIA (formerly Title I, ESEA) program for grades 3 and 5.

Another dominant testing effort is seen in the grade 7 and grade 10 proficiency assessments which are given to students initially at these levels and repeatedly (if necessary) until they are passed. Three tests--one each for math, writing, and reading--are administered for these proficiency assessments at each level. The high school assessment meets a state mandate requiring districts to establish such testing as a requirement for graduation. The junior high proficiency tests represent a district decision to assure appropriate pupil performance prior to high school entry, although pupils may enter 10th

Table 12
Summary of Metro District Basic Skills Testing

	Test	Grades	Type	Rationale
Elementary	CBSS	1-6	Criterion-referenced	Pupil diagnosis, curriculum planning, 3-5: Chapter I reports to State/Fed
	CTBS	3,5 (6 optional)	Norm-referenced	Instructional program assessment.
	CTBS Espanol	1-6	Norm-referenced	Individual tests for all children receiving Spanish reading instruction.
	CAP	entry,1,3,6		State Assessment
Junior High	ASC	7 plus retakes	Proficiency	Pupil progress, math
	Writing Profic.	7 plus retakes	Proficiency	Pupil progress, language, writing
	PAIR	7 plus retakes	Proficiency	Pupil progress, reading
	CTBS	8	Norm-referenced	Instructional program assessment.
(Chapter I schools)	CTBS	7,8,9	Norm-referenced	State/Federal reports.
Senior High	Math Profic.	10 plus retakes	Proficiency	H.S. graduation requirement, math.
	Writing Profic.	10 plus retakes	Proficiency	H.S. graduation requirement, writing, language
	READ Sr.	10 plus retakes	Proficiency	H.S. graduation requirement, reading
(Chapter I schools)	CTBS	10-12	Norm-referenced	State/Federal reports (10 out of 49 schools)

grade without having passed the junior high battery of proficiency tests.

Finally, some of Metro's testing is done to satisfy reporting requirements for federal and state aid programs. The CTBS is administered to fulfill these requirements at various levels and it complements the grades 3 and 5 CBSS test which also doubles for district and federal purposes.

Metro District Central Office Testing Costs

The size and organization of Metro District dictated a somewhat different approach to the assessment of district testing costs than the one we pursued in Littleton, though the guiding questions were constant: What range of elements constitutes the costs of conducting basic skills testing in the district? Which tests are accompanied by which types of costs? What is the magnitude of these costs? What is the importance of these costs from the standpoint of overall district resource management? But since there are hundreds of schools and thousands of teachers and other individuals involved in Metro's testing, our research necessarily could not take the microscopic look at testing that we were able to do in Littleton.

The first problem we faced in Metro was that testing responsibilities lay in many offices throughout the district, and that no one person had a complete view of the full array of testing practices and related activities. The second problem was that the various officials responsible for testing were not accustomed to thinking about the various costs of what they oversee. The district does not budget for testing in ways that correspond to our questions of interest. We

therefore confronted a substantial amount of detective work, and the results left us with a great many partial perspectives of the objects of our inquiry. What follows is a report of our attempts to reconcile these views onto an overall ledger.

In contrast to Littleton, Metro assigns significant amounts of central resources to its basic skills testing programs, both in the form of personnel who administer and coordinate the testing programs, and in direct purchases of services and materials. The central office houses five professional and five clerical staff who work exclusively with district tests. One professional oversees the entire testing program, one administers Chapter I (compensatory education) testing programs, and the other three divide responsibility for the remaining tests. The activities of these individuals have largely predictible descriptions--scheduling tests and all related activities, coordinating purchase and delivery of materials, arranging for test scoring, writing reports of test results, and ongoing development of the testing programs.

District testing coordinators also conduct inservice training for principals, coordinators of testing at the school level, and area directors of instruction. The training schedule is heaviest in October and January when two or three day-long sessions per week are conducted by one or more of the five central office coordinators.

The central office also houses two automated scoring machines which are used whenever machine scorable answer sheets accompany tests. These machines require between four and six operators when tests are being scored--and these personnel devote themselves to other

test-handling duties when not operating scanners. In addition, the central office requires some programmer/consultants to assist in its information processing.

Table 13 summarizes the costs Metro incurs to maintain its central testing related services. These costs are predominantly found in the various personnel allocated to testing in the central office. The total central cost, $ 482,000, represents a cost of just under one dollar per pupil enrolled in Metro District.

Table 13
Metro District: Central Costs Not Specific to Particular Tests
($ in 1000's)

Job Identification	Number FTE	Annual Cost
Basic Skills		
Professional/coordinator	4.1	$ 150
Clerical	4.0	80
Compensatory Education		
Professional/coordinator	1.0	35
Programmers	1.9	65
Clerical	1.0	20
Scanning		
Operator/handlers	5.0	100
Programmer/consultant	.2	7
		$ 457
Office Space		$ 10
Transportation		10
Warehousing		5
		$ 25
Total Central Office		$ 482
Total Cost per pupil		$ 0.89

In addition to maintaining a central coordination and administration staff for its basic skills testing, Metro incurs significant central costs for testing through a variety of services and purchases outside of the central office which nevertheless remain above and beyond any costs incurred in the schools themselves. These costs are summarized in Table 14.

Table 14
Metro District: Summary of Annual Costs
Above School Level, Outside Central Office
($ in 1,000's)

Cost	Amount
Development of CBSS	$ 120
Area Scoring Centers	$ 400
Supplies	$ 120
Test Processing and Handling	$ 103
Contract Scoring	$ 211
Total	$ 954
Average cost per pupil	$ 1.75

The most significant testing program cost outside of the central office costs stems from the operation of the area scoring centers in the district's 10 regional offices. The 1981-82 estimate of these costs was $400,000 which is allocated primarily to "seasonal" employees hired temporarily during peak test scoring times. Metro District also contracts with outside vendors for test processing and handling. Supply costs for all tests (booklets, answer sheets, pencils) are estimated to total $120,000 annually. Finally, in 1976 Metro entered into a long-term contract with an outside laboratory for the development of its elementary skills assessment CBSS test. In

1981-82 this service cost about $120,000 and the total spent for this contract since its inception is about $1,000,000.

The total cost of these additional services and purchases ($954,000) represents about $1.75 per pupil district wide. The grand total of testing costs in Metro District which occur above the school level ($1.436 million) represents about $2.64 per pupil. These estimates are highlighted in Table 15.

Table 15

Total Metro District Testing Costs Above the School Level
($ in 1000's)

Central Office Costs	$ 482
Other Central Costs	$ 954
Total	$ 1,436
Average cost per pupil	$ 2.64

The Costs of Specific Testing Conducted in Metro District

Table 16 shows the costs incurred by Metro District for each of its basic skills tests. These figures represent a mixture of direct budgeted costs, the estimated costs of personnel assigned to specific tests, and the pro-rating of central testing functions costs not specifically attributable to any one particular test or group of tests. We attained the direct costs for materials and contract scoring in district accounting records, which also provided estimates of processing and handling costs. Area scoring center costs were estimated in interviews probing share-of-activity devoted to the various tests. District office personnel costs were assigned on the

basis of reported share of personnel time devoted to specific tests. The remaining costs of testing ($307,000) were allocated across tests according to the number of pupils actually tested in each assessment during the school year.

Table 16
Metro District: Central Costs by Test
($ in 1,000's)

	DIRECT COSTS				DISTRICT OFFICE				
Test	Materials	Contract Scoring	Processing & Handling	Area Scoring Center	Profess.	Clerical	Share of Unallocated Costs[1]	Contract Development	TOTALS
CBSS	$ 5	$ 0	$ 15	$ 200	$ 19	$ 10	$ 98	$ 120	$ 467
CTBS	3	0	0	50	19	10	80	--	162
ASC	20	83	5	25	12.5	7	21	--	173.5
Writing Proficiency (Jr. High)	teacher graded		36	25	12.5	7	21	--	101.5
READ Jr.	60	45	11	25	12.5	7	21	--	181.5
Math Proficiency	20	83	6	25	12.5	7	22	--	175.5
Writing Proficiency (Sr. High)	6	teacher graded	30	25	12.5	7	22	--	102.5
READ Sr.	6	0	0	25	12.5	7	22	--	72.5
	$ 120	$ 211	$ 103	$ 400	$ 113	$ 62	$ 307	$ 120	$ 1436

[1] Based on share of total pupils tested for each test.

The total costs for each test are again displayed in Table 17, along with per pupil costs for each.

Table 17

Metro District: Costs of testing Per Pupil Tested, by Test

TEST	TOTAL COSTS ($ in 1,000's)	COSTS PER PUPIL TESTED[1]
CBSS	$ 467	$ 1.60
CTBS	162	1.55
ASC	173.5	3.55
Writing Proficiency (Junior High)	101.5	2.13
READ Jr.	181.5	3.68
Math Proficiency	175.5	2.97
Writing Proficiency (Senior High)	102.5	1.79
READ Sr.	72.5	1.26
Total	$ 1436	Average $ 2.64

[1] Numbers of pupils tested estimated using enrollments by grade level, plus estimates of test retakes for proficiency tests.

School Level Testing Costs

We now turn to the costs of testing in Metro District that lie beyond the district's central office. Recall that the district spends about $2.64 per pupil for these activities. Now we investigate testing costs incurred in the schools themselves, which involve administrators, counselors, coordinators, secretaries, and teachers. We should note that our researchers went beyond basic skills testing in their analysis of testing costs in the Metro District study schools. This was done with hopes of facilitating comparisons to our recent national survey which probed the nature of all testing activities in elementary and secondary schools. These comparisons are reported, as well as some comparisons to the Littleton District basic skills testing costs reported above.

What follows is a suggestion of what the total Metro District cost patterns would look like if testing practice similarities were to obtain between the schools we observed and those in the balance of the district's schools, which we could not observe. At the elementary level, we studied the testing costs in a "typical" Metro District school, Cityside. We extended these findings across all of the district's elementary schools to estimate the total of resources devoted to testing at this level. At the junior high and high school levels we do not even have limited field work to draw from, (since project resources precluded fieldwork at that level.) To project total costs at the secondary level, we examined what we learned abut testing costs in our other study district (Littleton), and calculated what must be considered, at best, illustrative (and at worst simply

non-representative) figures for the much larger Metro District. At both the elementary and secondary levels, we used information derived from our national survey of test use to suggest what types of tests may account for the costs we identified.

Elementary testing costs: Our case study of Cityside afforded us a rich view of its various costs related to testing of all types conducted during the 1981-82 school year.

Table 18 draws on our findings regarding Cityside's distribution of testing costs (Table 38) that year and projects these cost findings across the remainder of the district's elementary schools. The table shows our case study findings regarding the central office costs as well as the schools' direct and indirect costs associated with all testing over the 1981-82 school year. These tests include basic skills tests and the various tests that teachers use solely for curricular or pupil progress assessments. Column (A) presents the costs for all contributing personnel, services, and materials in per-pupil terms. The cost per pupil at Cityside school for all testing activities is estimated at $130, or almost 4.0 percent of the district's total general expenditures ($3,300) per pupil.

Estimates of the total testing cost across the district's 427 elementary schools, which are displayed in column (C) of Table 18, were calculated by means of a linear extrapolation from what we observed in the case study. The projected grand total of testing costs for Metro District elementary schools is about $38 million, which represents about 3.9 percent of the district's total expenditures for elementary education ($960,000,000).

TABLE 18

Estimates of Total METRO DISTRICT Elementary Level Testing Costs Per Cityside School Case Study

TYPE OF COSTS	Hours/Year (% Work Time)	Dollar Equivalents	(A) Total at Cityside (Enrollment = 830)	(B) Per Pupil Cost	(C) Estimated Total Costs All Elementary Schools (Enrollment = 291,000)
District-Office Costs:					
$2.64 per pupil x 830 pupils			$ 2191	$ 2.64	$ 768,000
Direct Costs to School:					
Purchase of Metropolitan Achievement Test			1200		
Purchase of Curricular Reading Tests			5000		
Purchase of Scantron Scoring Machine Forms			200		
			$ 6400	$ 7.71	$ 2,244,000
Indirect Costs for School (Personnel Time):					
Administrators/Coordinators -					
Reading Resource Teacher	328.5 (19.3%)	$ 5790			
Title I Program Coordinator	11.5 (0.7%)	210			
Teacher Testing Coordinator	35.0 (2.1%)	472			
	375.0	$ 6472		$ 7.80	$ 2,270,000
Clerical/Secretarial	10.3 (0.5%)	$ 95		$ 0.11	$ 32,000
Classroom Teachers -					
Average Time Per Teacher	199.2 (12.2%)	$ 2745			
Number of Teachers	x 30	x 30			
	5976.00	$82,350		$ 99.22	$ 28,873,000
Instructional Specialists -					
Bilingual Coordinator	156.25 (9.2%)	$ 2760			
Bilingual Teacher (assists with testing)	8.08 (0.5%)	112			
	164.33	$ 2872		$ 3.46	$ 1,007,000
Instructional Aides (Paraprofessionals) -					
Aide to Reading Resource Teacher (n = 1)	109.45 (20.6%)	$ 657			
Aide to Instructional Specialist (n = 1)	4.58 (0.9%)	$ 27			
Classroom Aides (per classroom)	39.48 (7.8%) $237				
Number of Classrooms	x 30 x 30				
	1184.50	$ 7110			
	298.5				
Classroom Volunteers	92.2	7794		$ 9.39	$ 2,732,000
Student Time - (Average Time Per Pupil)	76.1 (8.6%)	---			
TOTAL COSTS FOR SCHOOL (1981-82 School Year)		$108,174			$ 37,926,000 (or about $130 per pupil)
AVERAGE COSTS PER CLASSROOM (n=30; avg 27.67 pupils/class)		$ 3,606			
COSTS PER PUPIL		$ 130.33		$ 130.33	
PROPORTION OF DISTRICT ANNUAL EXPENDITURE PER CHILD (= $3300)		3.9%			

Our Cityside School case study and our national survey of schools' testing practices allow us to estimate the types of tests accounting for the more than $130 worth of resources per pupil estimated to be devoted to testing in Metro District's elementary schools. According to our respondents at Cityside, the vast majority of these resources are devoted to tests imbedded in curriculum materials or to tests developed by teachers or the schools themselves. Table 19 shows that more than 80% of testing resources are directed toward these tests (commercial curricular plus teacher developed). The data further show that only about 7% of testing resources are expended to satisfy state requirements for pupil assessment and demonstration of competencies. Table 19 also shows that the reported distribution of testing resources at Cityside School does not depart radically from national patterns of test use at the elementary level.

This analysis also suggests that the basic skills testing (i.e., non-curricular) activities in the Metro District would account for about 20 percent of elementary testing costs. This implies that Metro District spends about $26 per pupil, in comparison to about $11 per pupil for elementary level basic skills testing at Littleton. This relationship is consistent with the greater number of elementary grades annually tested in the Metro District (compare Table 6b with Table 12).

Table 19

Distribution of Total Costs for Testing Per Pupil
in Metro District: Elementary Grades by Type of Test
[Per Cityside Case & Per National Survey Estimates of Distribution]

Type of Test	Distribution Per Case		Distribution Per National Survey	
	%	$	%	$
State Assessment	7.0%	$ 9.09	3.0%	$ 3.91
MCT's			1.5%	1.95
Curriculum Materials Tests	38.1%	49.66	31.5%	41.06
Other, Commercially Published	8.3%	10.82	17.5%	22.81
Locally Developed	3.3%	4.30	10.5%	13.68
School or Teacher Developed	43.3%	56.46	36.0%	46.92
	100.0%	$ 130.33	100.0%	$ 130.33

Junior high and high school testing costs: Our findings of total Metro District costs for testing at the secondary level, as we previously mentioned, are not empirically based but rather offer a sketch of what cost patterns might look like if our Littleton findings held up in the much larger Metro District. This secondary level cost portrayal in Metro is further hampered by the fact that our Littleton District analysis surveyed only basic skills testing and not testing done to satisfy curriculum requirements. The analysis which follows, therefore, is restricted to basic skills testing at the secondary

level, which typically accounts for considerably less than half of all testing activity.

The analytical reasoning we employed to project Metro's junior high and high school testing cost was as follows. If per pupil basic skills testing costs at the Metro junior high and high schools were equivalent to what we observed in Littleton, the total basic skills testing costs in Metro could be obtained simply by multiplying the per pupil cost estimates by actual enrollments. Furthermore, if in both districts these costs were incurred in similar patterns across the different types of resources used in testing, we can base the estimated distribution of Metro District costs on the pattern observed in Littleton. In addition, our national survey of secondary level testing allowed us to suggest just which types of tests these resources might be devoted to. These constructions, then, despite their limited foundations, are anchored to various segments of our three-years of research.

Table 20 shows that, if the $20.93 overall per pupil cost for Littleton's basic skills testing were to characterize Metro's costs for the same activities, Metro would spend a total of about $2.5 million on these tests in its junior high schools. This represents a little less than 1% of the districtwide average per pupil general expenditure. If the distribution of these costs is also similar to that observed in the smaller district, where the costs of teacher time account for about three-fourths of the basic skills testing resources, this $2.5 million would be distributed as shown in the right-hand column of Table 21.

Table 20

Projected Basic Skills Testing Costs in Metro District: Junior High School

[Based on Littleton District Estimates of School Level Costs & Metro District Central Cost Analysis]

	Cost By Category	Total Metro District Costs [120,000 Enrollment]
Central Cost*	$ 2.64	$ 316,800
Administrators/ Counselors	2.71**	325,200
Clerical	0.79**	94,800
Teachers	14.79**	1,774,800
	$ 20.93 per pupil**	$2,511,600
	(< 1% of district jr. high budget per pupil)	(< 1% of district jr. high budget)

* Estimated in Metro District Central Office Analysis. Includes Purchases of Materials/Services.

** Derived from Tables 6 and 7.

Our national survey of testing practices suggested that different types of basic skills tests might occupy differing amounts of time at the junior high school level.* Table 21 incorporates the national distribution of basic skills type tests, and applies this distribution to the $20.93 in per pupil resources we have identified as suggestive of Metro District junior high test costs. As we have previously pointed out, about a third of all basic skills testing at this level is done to satisfy state mandates, and the balance is intended to satisfy local demand for basic skills development information.

* Our 10th grade estimates from the survey are used for these projections. No junior high grades were surveyed.

Table 21

Distribution of Metro District Junior High Basic Skills Testing Costs

[Per Total Cost Estimates (Table 20) and
National Survey of Test Use Distributions.]

Type of Basic Skills Test	% of All Basic Skills Test Time Reported	Per-Pupil Cost Distribution
State Assessment	29%	$ 6.06
MCT	6%	1.29
Local or District Developed	29%	6.06
Other, Commercially Developed	36%	7.52
		$ 20.93 per pupil

Treatments analogous to those presented for junior high school estimates were used to derive estimates for Metro District high school level basic skills testing costs (Tables 22 and 23). Littleton District reported "spending" only $3.82 per pupil for basic skills testing efforts. A similar level of costs in the Metro District as suggested in Table 22, would imply that a total of about half a million dollars would be devoted to basic skills testing for the 127,000 pupils in its high schools. The pattern of costs among resources (also shown in Table 22) is weighted comparatively toward administrators and counselors at the high school level. Littleton reported a predominance of centrally administered basic skills tests, and the distribution shown above reflects the teachers' comparatively low use of these tests.

Table 22

Projected Basic Skills Testing Costs in Metro District: High Schools
[Based on Littleton District Estimates of School Level Costs]

	Cost By Category	Total Metro District Costs [127,000 Enrollment]
Central Cost*	$ 2.64	$ 335,300
Administrators/ Counselors	0.59**	74,900
Clerical	0.31**	39,400
Teachers	0.28**	35,600
	$ 3.82 per pupil**	$ 485,200
	(< 1% of district budget per pupil)	(< 1% of district budget)

* Estimated in Metro District Central Office Analysis. Includes Purchases of Materials/Services.

** Derived from Tables 6 and 7.

Table 23

Distribution of Metro District High Schools Basic Skills Testing Costs

[Per Total Cost Estimates (Table 18) and National Survey of Test Use Distributions.]

Type of Basic Skills Test	% of All Basic Skills Test Time Reported	Cost Distribution
State Assessment	14%	$ 0.53
MCT	14%	0.53
Local or District Developed	29%	1.11
Other, Commercially Developed	43%	1.65
		$ 3.82 per pupil

Table 23 shows how this small level of testing costs at Metro District high schools would be allocated across different types of basic skills tests, if Metro's patterns were similar to those found in our national survey. In comparison to the junior highs, high school costs are somewhat more tied to state assessments and competency testing, but are still dominated by local demands for basic skills testing.

Summary Comments

Our limited efforts to gain a representative view of the more than 500 elementary and secondary schools in the Metro district restricted our ability to provide concrete estimates of what the district actually spent on testing beyond the central office level. In Littleton District, through simple surveys and interviews we were able to produce a relatively complete portrait of district testing practices. In the case of Metro District, given its size and great diversity of schools and pupils, as well as our limited research budget, we were not able to derive reliable total cost estimates. Instead, we have offered a characterization of school level testing costs which is based on a partial view of actual district practice, on inferences drawn from our in-depth study of a smaller district, and on our national survey of testing practices.

MONETARY COSTS OF ALL ACHIEVEMENT TESTING IN TWO ELEMENTARY SCHOOLS

The preceding section offered an accounting of basic-skills testing costs in the Littleton and Metro School Districts. This section analyzes the costs of all achievement testing, in basic skills and other areas, in one elementary school in each district.

As noted earlier, cost accounting information was gathered in extended interviews with the school administrators, classroom teachers, and instructional specialists who described the time and other resources that they and their students expended on achievement testing of all types in all school subjects through the 1981-82 academic year.

Testing Costs in Littleton District's Hillview School

Hillview is the smallest of Littleton's four elementary schools. Its eleven classrooms and learning library serve 191 students: 50% of Asian background, about 45% from White Anglo families, the remaining 5% Hispanic or Black. Specific socioeconomic indices were unavailable, but the neighborhood from which Hillview children come is one of the higher-income areas in generally well-to-do Littleton. Homes within the school's attendance boundaries are valued in the $250,000 - $400,000 range, substantially above the $120,000 average for the county. Students' parents work largely in professional, executive, and scientific-research positions.

Hillview participates in no special educational programs sponsored by the state or federal government. Its program is supported exclusively by Littleton District funds.

The school has a reputation for excellence in the Littleton District, and its students are considered "very high achievers" by the teaching staff. As the principal noted, "A so-called "average" kid (in terms of national norms) is not average here. He's below average."

Hillview educators are experienced, and most have been at the school for some time. The principal has served at Hillview for fifteen of his twenty-six years as a head administrator. The teachers' length of service at Hillview is, on the average, nine years. Most taught elsewhere before joining the Hillview faculty.

It is difficult to present a comprehensive summary of Hillview's testing program because there is considerable variation from classroom to classroom. Table 24, however, shows those measures that are widely and/or consistently administered. In addition to those shown are various tests and quizzes developed or selected by individual faculty members.

Hillview Testing Costs in Overview

Table 25 itemizes the total costs for all achievement testing reported for Hillview during the 1981-82 school year. Most entries in this table are self-explanatory, especially in light of the accounting procedures previously explained. Derivations of the "present work time" and the dollar equivalents for staff time are clarified in footnotes to the table.

The first item, district-office costs, is incurred in the time personnel in Littleton District's central office devote to testing. (See Tables 6-9.) Here, the $1.30 per pupil cost, shown in Tables 8 and 9, is applied to Hillview's 191 students.

TABLE 24

Hillview Elementary School Testing Program

Test	Grade(s)	Required by:	Administered per year
Multi-Subject			
Stanford Achievement Test	K - 6	District	2
Otis-Lennon Intelligence Test	K - 6	District	2
State Assessment Program	1,3,6	State	
Reading			
Ginn 720 Placement Test	1*	District	1
Ginn 720 Criterion (Unit) Test	1 - 6	District	9 - 20**
Ginn 720 Mastery Test	1 - 6	District	1 - 2
Ginn 720 Booster Test	1 - 6	District	as needed
Math			
Scott-Foresman Unit Pre-Test	2 - 6	District	5 - 12**
Scott-Foresman Unit Post-Test	1 - 6	District	5 - 12
District-Developed MATH Operations Test	1 - 5	District	weekly-monthly
Math Proficiency Test	4	District	1
Junior High School Math Placement	6	District	1
Spelling			
Teacher-Developed or Commercial-Curriculum Spelling Test	1 - 6		weekly or bi-weekly
Physical Education			
Physical Performance Test	5	State	1

* The instructional specialist in the Hillview learning laboratory also routinely administers the Ginn placement test to all students new to the district except those not proficient in English.

** Variations noted in the frequency of curricular testing were reported from classroom to classroom. In some instances, variations occurred within classrooms where individualization of instruction permitted learners to progress through the curriculum at different rates.

As with other Littleton elementary schools, Hillview makes no direct purchases in conjunction with testing. The district and state supply various mandated tests. Consumable test booklets that accompany commercial curriculum materials in reading and math are bought by the district. (In the district budget, these costs are included under general outlays for instructional materials. We could not differentiate and prorate them for Hillview. A rough estimate, however, suggests that the cost of these curriculum-embedded testing materials would be under $1,000 for Hillview's 191 students.)

Of course, teachers consume paper, duplicating fluid, ditto masters, and even chalk in the process of producing their own tests. But no one at Hillview could estimate what proportion of these and similar supplies went for testing. In any case, the cost of routine stationery supplies for testing is almost certainly minimal.

Table 25 shows then, that virtually all of Hillview's economic testing costs are indirect: i.e., they are the dollar values of the staff time devoted to testing. As indirect dollar costs they are borne by the district, which pays staff salaries. But the staff time invested in testing can also be construed as an opportunity cost -- that is, as the allocation of a resource to one activity (testing) instead of another (for example, explicit instruction). Seen from this perspective, the cost of testing in staff time is borne by multiple constituencies. These can include the staff members themselves, the students, their parents, and the community, as well as the school district.*

* One can reasonably argue that the value gained by the allocation of staff time to testing -- e.g., in more appropriate instruction; in clearer communication of students' educational status to parents, next year's teacher, and subsequent school, etc. -- is well worth the information that tests yield.

TABLE 25

Total Costs for All Achievement Testing in Hillview School -- Littleton District
[Enrollment = 191]

	Hours/Year (% Work Time)[2]	Dollar Equivalent[3]
District Office Costs[1]:		
$1.30 per pupil X 191 pupils		$ 248
Direct Costs to School:		none reported
Indirect Costs for School (Personnel Time)		
Administrators/Coordinators		
Principal	63.75 (3.75%)	$1125
Teacher Testing Coordinator	36.00 (2.12%)	477
	99.75	$1602
Secretarial/Clerical		none reported
Classroom Teachers		
Average Time Per Teacher	252.96 (15.5%)	$3875
Number of Teachers	X 11	X11
	2782.50	$42,625
Instructional Specialists[4]		
Learning Laboratory/English as a Second Language	197.63 (11.6%)	$2610
Instructional Aides (Paraprofessionals)		none employed
Classroom Volunteers	77.66 (?)	
Student Time[5]		
Average Time Per Pupil	88.04 (9.95%)	
TOTAL COSTS FOR SCHOOL (1981-82 School Year)		$47,085
AVERAGE COSTS PER CLASSROOM (n = 11; avg. 17.36 pupils/class)		$ 4,280.45
COSTS PER PUPIL		246.56
PROPORTION OF DISTRICT ANNUAL EXPENDITURE PER CHILD (= $1670)		14.7%

TABLE 25 (continued)

1 Calculations of district office costs are shown in the preceding section.

2 The "% Work Time" figures are based on respondents' reports of hours worked per week before, during, and after school hours. These reported hours per week were averaged by role category across the two schools studied (Cityside and Hillview). Reported hours were within similar ranges at both schools. Work times used are as follows:

(a) For administrators, coordinators, and instructional specialists: 46 hours per week X 37 weeks per year.

(b) For classroom teachers: 44 hours per week X 37 weeks per year = 1628 hours per year.

(c) No total hours per unit or person could be ascertained for volunteers.

3 Dollar equivalents are based upon the proportion of work time expended at the following salary estimates:

(a) For administrators and coordinators - $30,000 salary and fringe benefits.

(b) For classroom teachers and the instructional specialists - $22,50 salary range and fringe benefits.

These salary estimates are equivalent to those used in the analysis of district costs, but are 20 - 25% lower than those actually in effect in this school.

4 Instructional specialist time reported is devoted to assessing the language competence of incoming students, other placement testing of new students, and recurrent assessment of students enrolled in an English as a Second Language (ESL) course.

5 Student time shown equals the time spent by the typical student in each classroom averaged across the school's regular classrooms. The percentage shown is based on 5 class hours per day (not counting the hour for lunch and recess) for 177 school days per year, which equals 885 classroom hours per school.

The allocation of staff time is by far the most substantial economic cost of testing at Hillview, and so deserves further examination.

Administrators' time was spent in a number of ways. Hillview's principal devoted some of his testing time to district-wide administrators' meetings for "in service" on state- and district- required tests. He expended eight and three quarter hours on these sessions through the year.

More of his time on testing was given over to processing materials for these extramurally mandated measures. As described by the principal, this work included "receiving the tests, distributing them to the teachers, collecting them again, checking them over, packing them for mailing, and so on." He reported spending four and one quarter hours on these tasks in the fall and again in the spring during the conjoint administration of the Stanford Achievement Test and Otis-Lennon Intelligence Test. Similar handling of the State Assessment tests and fourth-grade proficiency test consumed three hours and a half hour, respectively.

But the greatest proportion of the time the principal gave to testing was spent in review and analysis of test results. He routinely calculated year-to-year comparisons of scores for different classrooms and grade-levels, noted trends, and disseminated these and similar analyses to teachers. In so doing, he extended the information provided in the reports of the state or testing companies. (Note that this time is a cost of obtaining assessment information. The time the principal and teachers spent making use of test results is

not included here or elsewhere in this report.) Some 42 of the principal's work hours were in test-score review and analysis through 1981-82.

A second staff member, the instructional specialist who ran Hillview's learning lab, assisted the principal in coordinating the Stanford Achievement testing. She gave 18 hours of her time to this work in the fall and once again in the spring. Her responsibilities included helping to distribute test forms; answering teachers' questions about administration procedures; assuring that all test forms were returned; and re-checking the students' answer sheets to be sure that stray pencil marks were erased, answer slots were sufficiently "bubbled in," etc.

As Table 26 shows, the principal and learning lab instructor together expended 99.75 hours on testing. For both, testing responsibilities consumed less than 5% of their school-year work time.

Classroom teachers' time on testing was spent in diverse ways. As Table 25 indicates, the average (mean) time Hillview teachers spent on testing in 1981-82 was about 253 hours. Calculating annual work time as described in the footnote to Table 25, this constitutes 15.5% of a Hillview teacher's yearly work effort. Naturally, these averages mask some diversity in the allocation of time to testing. (A simple listing reveals the extent of this variation). Below, teachers' total time on testing per annum are displayed, together with the number of different kinds of tests that they reported giving through the year. (Here, "kind of test" refers broadly to such separate measures as a weekly spelling test, reading unit tests, reading quizzes, the Otis-

TABLE 26

Summary of Administrators' Annual Time

(In hours, showing % of their total time on testing)

District in-service to prepare for testing	8.75 (8.8%)
Processing test form, overseeing administration	49.00 (49%)
Reviewing and analyzing test results	42.00 (42%)
	99.75

Teacher (Grade)		Number of Different Tests	Hours per Year on Testing
Fulsome	(K)	8	210.50
Gardener	(1)	9	215.05
Jameson	(2)	10	163.91
Koviak	(2/3)	11	288.90
Fushima	(3)	13	386.67
LaMarr	(4)	16	250.91
Earle	(4)	16	395.85
Vera	(5)	19	306.05
Hurteby	(5)	18	260.93
Leacock	(6)	8	151.75
Coxe	(6)	8	152.25

Lennon, etc.) Teachers' grade levels are indicated parenthetically (the names used are fictitious).

The number of different kinds of test given increases regularly until the sixth grade, where teachers Jameson and Koviak team teach and employ a variety of assignments and projects, instead of tests, for assessment. Nevertheless, in some instances, the time devoted to testing varies markedly within a grade and between adjacent grades. (Compare the total hours of Jameson, Koviak, and Fushima, or of LaMarr and Earle.)

On the average, Hillview teachers spend only about a third (34.2%) of their testing-related time in actually administering tests. Here, we conceptualized test administration to include all the classroom time from the moments when the teacher begins to give students directions about doing the test until the teacher moves on to the next class activity. Thus, such activities as re-arranging seating, explaining the test format, answering students' questions beforehand, distributing and picking up test papers, and so on are all included in this definition of administration time. So, too, are relaxation periods between and immediately after different portions of a test battery. (Many teachers at Hillview and elsewhere provide their children time to "cool out" or "settle down" after sections of standardized tests.) This, then, is a broad (but appropriate) operational definition of test administration. Nevertheless, the mean time devoted to these "during testing" activities in 1981-82 was about 86.5 hours, or about 34% of the average of 253 testing-related hours per teacher per year.

Put another way, roughly two-thirds (65.8%) of Hillview teachers' average testing time (again, averaged across the school's eleven classroom instructors) was spent before and after classroom testing episodes. Time before testing was invested in constructing and duplicating tests, reviewing the appropriateness of questions in commercial curricular measures, reading administration directions for annual and bi-annual test batteries, and (in some instances) foregoing routine instruction to drill students on information and skills in explicit preparation for a test.* The Hillview faculty spent an average of 27.5 hours in 1981-82 (10.9% of the mean total testing time) on such "before testing" tasks.

Post testing activities -- recording scores, examining and "cleaning up" special answer sheets for machine scoring, grading, and so on -- consumed a mean time of 138.98 hours a year for the Hillview classroom staff. This constitutes 54.9% of the average of 253 testing-related hours per teacher per year.

The time that teachers devote to these before-, during-, and after-testing activities (summarized in Table 27) comprises by far the largest proportion of Hillview's annual testing "budget:" $42,625 (or 90.5%) of the $47,085 total. Bear in mind that this is an indirect cost, one met within the routine payment of teachers' salaries.

---

* Instructional activities such as these were included as testing time costs only when teachers reported that they would not have conducted them were it not for the test. Routine teaching of skills covered by a test was not included in calculating staff time allocated to testing.

Table 27
Summary of Classroom Teachers' Annual Testing Time

Mean time per teacher per year devoted to:		
"before testing" activities	27.5 hours	(10.9% total)
"during testing" activities	86.5 hours	(34.2% total)
"after testing" activities	138.98 hours	(54.9% total)
Mean, all testing-related activities	252.96 hours	(54.9% total)
Proporation of average annual work time testing*	15.5% total	

The Instructional Specialist's testing time, in her capacity as learning lab resource teacher, was spent in three general ways. First, she gave reading and math placement tests to all new students and also elicited a writing sample from them. During the 1981-82 school year, she expended 71.3 hours on these tasks. Second, in accordance with state law, she assessed the English language proficiency of incoming students when English was not the language spoken in their homes. (In some instances, the results of this assessment suggested that the writing sample and/or reading placement should be omitted.) This responsibility consumed 70 hours of her time during the year. And third, she routinely tested students in her daily English-as-a-Second-Language (ESL) class in language arts and spelling. This took up 56.3 hours in 1981-82. In all, then, the Hillview instructional specialist spent 197.6 hours on testing throughout the year. Using the salary rates described in Table 25, the dollar value of this time equals $2610 -- about 5.5% of Hillview's annual testing costs.

Table 25 also shows that the testing efforts of Hillview's paid

* See Table 25 footnotes for calculation of classroom teacher's average annual work time.

professional staff were supplemented by 77.6 volunteer hours throughout 1981-82. While volunteers' time is "free", the allocation of their hours to testing constitutes an opportunity cost of Hillview's assessment program. The use of volunteer time for other tasks was foregone on behalf of testing.

For the most part, parent volunteers at Hillview helped with standardized testing. Some helped proctor; others examined completed answer sheets for stray marks, insufficiently darkened "bubbles" (answer markings), and incomplete or incorrect student identification information. They also helped with such jobs as alphabetizing the forms.

Student time on testing is the last item in the overall itemization of Hillview testing costs presented in Table 25. Note that across Hillview's eleven regular classrooms, mean time per student per year is a fraction over 88 hours. This is roughly equivalent to the mean time per teacher spent in "during testing" administration (86.5 hours). But note also that on the average, nearly three hours of teacher time are required to deliver each hour of testing to the students.

Students at Hillview rarely spend cost-generating time on assessment before or after the test-taking episode. Based upon teachers' reports, the mean "before testing" time per student per year was 2.8 hours. (This of course excludes the routine teacher-learning time that precedes a test.) The mean "after testing" time per student per year was 5.3 hours. Together, these opportunity costs comprise only 9.4% of the 88 hours per student annual average. What is more, most

of this "before" and "after" time can be traced to Hillview's two fifth grade classrooms, where students spent considerable amounts of time in explicit preparation for a State-mandated physical education assessment. From September to April, they devoted a portion of their daily physical education period to practicing exercises included on the test, exercises which would otherwise not have been part of their P.E. program. The fifth grade teachers also routinely engaged their pupils in in-class test correction (defined here as an after-testing activity). Approximately 50% of the "before testing" and "after testing" student time investment reported school-wide occurred in these two classrooms.

Finally, the general testing budget in Table 25 shows that Hillview's annual testing costs of $47.085 (all indirect costs) equal $246.56 per pupil. This may seem a large amount, but it comprises only 14.7% of Littleton District's annual per-pupil expenditure ($1670).

The preceding discussion constitutes a basic accounting of Hillview Elementary School's 1981-82 testing costs. This information can be reconfigured to address a number of interesting and important questions.

Hillview's Costs for Required and Non-Required Testing

Tables 28 and 29 show the proportions of Hillview's yearly testing costs incurred as a result of various testing requirements.

State required testing consisted of: (1) an annual State Assessment at grades 1, 3, and 6; (2) the once-a-year physical performance test at grade 5; and (3) the language assessment of all potentially

TABLE 28

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME PER YEAR
On Required and Non-Required Testing*

Each staff category cell shows:
° No. of staff members involved
° Avg. hours/staff member/year
° % Total testing time for staff by category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS
Required by State	1 15.75 15.8%	9 8.66 2.8%	1 70.0 35.4%		163.6 5.2%	4.46	9
Required by District	2 42.0 84.2%	11 117.66 46.5%	1 71.3 36.1%	3 24.22 93.6%	1522.2 48.2%	40.26	11
Required by School Principal		2 12.91 0.9%			25.8 0.8%	5.08	2
TOTAL REQUIRED (In person hours)	99.75 (100.0%)	1397.9 50.2%	141.3 71.5%	72.66 93.6%	1711.6 54.2%	44.46	11
NOT REQUIRED (In person hours)		1384.6 49.8%	56.33 28.5%	5.0 6.4%	1445.9 45.8%	43.57	11
TOTALS by staff category (In person hours)	99.75 (100.0%)	2782.5 (100.0%)	197.63 (100.0%)	77.66 (100.0%)	3157.5 (100.0%)		

* Required testing includes any testing mandated by someone or some agency in the organizational hierarchy above the classroom teacher.

TABLE 29

Hillview School - Littleton District
Distribution of Testing Costs Per Year

Required & Non-Required Testing

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	TOTAL DOLLAR VALUE (% Total)
Required by State	$ 253	$ 1193	$ 924	$ 2370 (5.0%)
Required by District	$1349	$19821	$ 942	$22112 (47.0%)
Required by School Principal		$ 384		$ 384 (0.8%)
TOTAL Required	$1602	$21398	$1866	$24866 (52.8%)
TOTAL Not Required		$21227	$ 744	$21971 (46.7%)
TOTAL by category (% Total)	$1602 (3.4%)	$42625 (90.5%)	$2610 (5.5%)	$46837
			District Office Testing Costs (0.52%)	+ $248
			TOTAL	$47085

non-English proficient pupils, as mandated in state bilingual education legislation. Collectively, these requirements fell more heavily upon the instructional specialists' and principals' time, but comprised a very small proportion of the overall staff-time investment in testing. As Table 29 indicates, a mere 5% of Hillview's testing costs in 1981-82 were allocated to State-required testing.

As seen in Table 29, the district testing requirements, as they affect Hillview, seem at first glance to have occasioned 47% of all 1981-82 testing costs. Note, however, that among the district-required tests were various measures accompanying the reading and math text series that all teachers used. A substantial proportion of Hillview school's staff-time testing costs were incurred in the use of these measures. In fact, if we exclude the time spent on them from the "required-by-district" total, that total is very nearly cut in half. Some 739 person hours are deleted from the total of 1522 spent on district-required testing, leaving about 783. This would constitute 25% of the total staff person hours devoted to testing, rather than the 48.2% shown. Instead of 52.8% of Hillview's testing costs (Table 29) being devoted to all required testing, only 31% would be.

Of importance in the preceeding discussion is that, although curriculum reading and math tests <u>are</u> required, the issue with regard to testing requirements is usually framed in terms of testing <u>added on top of curriculum-embedded measures, on top of teachers' routine testing.</u> Teachers, for instance, sometimes argue that such testing takes up their time but provides little new information about their students. From the perspective of teachers and their advocates, then,

"required testing" is often of marginal necessity. But the routine tasks associated with teaching -- such as monitoring students' learning progress, grading, and conferencing with parents -- require recurrent assessment. Tests intimately connected with the curriculum-in-use are a practical necessity. If some such measures were not mandated, teachers would probably need to select or devise alternatives. In light of all this, we might wonder how the required/non-required testing picture would look at Hillview where the Ginn 720 reading tests and Scott-Foresman math tests are not mandated.

As matters stood, however, these tests were mandated by Littleton District. District-required testing was responsible for 47% of Hillview's 1981-82 testing costs. And slightly over half these costs resulted from mandates originating outside Hillview School.* The mean time per teacher per year devoted to required testing was about 127 hours; to non-required testing, approximately 126 hours. And notice that the typical student at Hillview spent just slightly more than half of his/her testing time, on the average, on mandated measures.

Hillview's Costs for Different Types of Testing

Tables 30 and 31 display Hillview School's 1981-82 testing costs by test type. The categories we used for typifying tests are eclectic in nature but isomorphoric with practitioners' everyday ways of talking about tests. They were identified as such in our first-year exploratory fieldwork and were employed throughout the project.

---

* Two fifth-grade teachers reported that the principal-required formal penmanship samples five times a year. This was the only school level testing mandate identified.

Several categories deserve brief explication. "Other, miscellaneous" testing at Hillview included: (1) the State-mandated physical performance test; (2) handwriting samples requested by the principal; (3) assessment of language competence as required by State bilingual legislation; and (4) certain commercially available, diagnostic instruments employed in the early grades.

District-continuum testing consisted only of the district-developed mathematics operations tests, which seemed based on a sequence of math objectives.

Minimum competency testing took the form of a locally available "proficiency test" administered in fourth grade.*

The "general intelligence" test category did not fall in our study of achievement testing. Teachers repeatedly mentioned it in interviews, however, and we included it here to provide a more complete picture of Hillview's testing.

With these elaborations, the findings shown in Tables 30 and 31 are self-explanatory. The largest percentage of staff and student time is devoted to tests accompanying commercial curriculum materials. Considerable time was also expended on teacher-constructed tests and quizzes (also closely tied to the curriculum), as well as on the standardized, norm-referenced Stanford Achievement Test.

Hillview's Costs for Testing in Different Subject Areas

The magnitude of Hillview School's testing costs for different subject areas is shown in Tables 32 and 33. The former reveals that

* The Littleton District's list of district tests indicates that proficiency testing occurs at the fourth and sixth grades. Sixth grade teachers at Hillview, however, did not report the test.

TABLE 30

Hillview School - Littleton District
Distribution by Staff & Student Testing Time Per Year
by Type of Test

Each Staff Category Cell Shows:

- No. of Staff/Members involved
- Avg. hours/staff member/year
- % Total testing time for Staff Category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (Hours)	NUMBER OF CLASSROOMS
Standardized, Norm-Referenced	2 42 84.2%	11 34.18 13.5%		3 17.8 68.7%	513.38 16.2%	19.6	11
State Assessment Program	1 11.25 11.3%	5 7.0 1.25%			46.25 1.5%	3.0	5
Minimum Competency	1 4.5 4.5%	2 9.33 .67%			23.16 0.73%	3.5	2
District Continuum		8 36.55 10.5%		1 12.33 15.9%	304.73 9.6%	6.9	8
Commercial, Curriculum-Embedded		11 122.48 48.4%	1 71.3 36.1%	1 5.0 6.4%	1423.58 45.1%	34.5	11
Teacher Constructed		11 55.5 21.9%	1 56.33 28.5%		666.83 21.1%	23.7	11
General Intelligence		7 4.39 1.1%		2 3.5 9.0%	37.73 1.2%	2.9	7
Other, Miscellaneous		5 14.37 2.6%	1 70.0 35.4%		141.85 4.5%	8.18	5
TOTALS by Staff Category (In Person Hours)	99.75 100.0%	2782.5 100.0%	197.63 100.0%	77.73 100.0%	3157.51		

Note that the number of classrooms in which each type of test is administered varies, thus the proportion of time the typical student spends on each type of test varies from classroom to classroom and the average times shown cannot be appropriately added.

TABLE 31

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
By Type of Testing*

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	TOTAL DOLLAR VALUE (% Total)
Standardized, Norm-Referenced (Grades K-6)	$ 1349	$ 5754		$ 7103 (15.1%)
State Assessment Program (Grades 1,3,6)	$ 181	$ 537		$ 718 (1.5%)
Minimum Competency (Grade 4)	$ 72	$ 286		$ 358 (0.76%)
District Continuum (Grades 1-5)		$ 4476		$ 4476 (9.5%)
Commercial, Curriculum-Embedded (Grades 1-6)		$20660	$ 942	$21602 (45.9%)
Teacher Constructed (Grades K-6)		$ 9335	$ 744	$10079 (21.4%)
General Intelligence (Grades K-6)		$ 469		$ 469 (1.0%)
Other, Miscellaneous (Grades)		$ 1108	$ 924	$ 2032 (4.3%)
TOTAL by category (% Total)	$ 1602 (3.4%)	$42625 (90.5%)	$ 2610 (5.5%)	
			District Office Testing Costs† (0.52%)	+ $ 248
				$47085

* Costs of staff time are calculated by multiplying percentage of staff time spent per category or cell (Table 30) by total dollar equivalent for staff category.

† District Office Costs pro-rated for Hillview School ($1.30 per pupil x 191 pupils = $248). These costs cannot be apportioned exactly by test type for Hillview Elementary, but see the section of this report dealing with district costs for a description of how Littleton District resources are allocated across different parts of the district-wide assessment program.

Hillview educators concentrate their formal assessment efforts mainly in the basic-skills subjects. Except for administrators, all categories of Hillview's assessment participants spend the greatest amount of their time on testing in math. Reading and spelling also receive larger commitments of staff and student time.

Testing in social studies, science, and subjects categorized under "other" (such as art and music) occurs in comparatively few Hillview classrooms.* And in those where teachers and learners do give time to testing in these subjects, it is usually less time per year than in the basic skills.**

In the next section of this report we will examine the costs of testing at Metro District's Cityside School and then discuss implications of the test-accountings for both schools.

Testing Costs in Metro District's Cityside School

Cityside is one of more than a hundred elementary schools in the large Metro School District. Of Cityside's 830 students, approximately 70% are Black; 28% are Hispanic; the remaining 2% is comprised of Asian, Pacific Island, and White Anglo children. Once an affluent Black neighborhood, the Cityside attendance area now ranks socioeconomically in Metro District's lowest quartile.***

---

* Teachers who do not test in science, social studies, art, etc. report evaluating students' progress in other ways -- through special projects, assigned reports, and routine classwork, for example.

** This may be explained by the fact that many teachers report spending less instructional time in "the content areas" than in the basic skills. If less material is covered per year, it may not be necessary for tests to occur as frequently or to last as long.

*** Metro District's socioeconomic rankings are based on the proportion of students' families receiving Aid to Families with Dependent Children (AFDC) and the percentage of enrollment qualifying for free school lunches under federal guidelines.

TABLE 32

Hillview School - Littleton District
Distribution of Staff & Student Testing Time by Subject

Each Staff Category Cell Shows:
- No. of staff members involved
- Avg. hours/staff member/year
- % Total testing time for staff category

SUBJECT AREAS	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (Hours)	NUMBER OF CLASSROOMS
Reading		11 52.47 20.7%	1 17.4 8.8%	1 5.0 6.4%	599.57 19.0%	12.12	11
Mathematics		11 77.11 30.5%	1 53.9 27.3%	3 15.44 59.7%	948.43 30.0%	25.11	11
Language Arts		8 24.30 7.0%	1 34.75 17.6%		229.15 7.3%	7.81	8
Spelling		8 51.42 14.8%	1 21.58 10.9%		432.94 13.7%	19.34	8
Social Studies		5 19.55 3.5%			97.75 3.1%	4.53	5
Science		5 28.0 5.0%			140.0 4.4%	5.8	5
Health - Phys. Ed.		3 8.33 0.9%			24.99 0.8%	7.19	3
Other, Miscellaneous		3 8.61 1.0%	1 70.0 35.4%		95.83 3.0%	3.39	3
Multi-Subject*	2 49.87 100.0%	11 42.06 16.6%		3 8.78 33.9%	588.74 18.6%	23.93	11
TOTALS by Staff Category (In person hours)	99.75 100.0%	2782.37 100.0%	197.63 100.0%	77.66 100.0%	3157.41 100.0%		

*The Multi-subject category includes standardized tests which assess performance in several subject areas. Also included in this category is the General Intelligence Test given twice a year at the same time as (i.e., on a day contiguous with) the standardized test. Some respondents reported time devoted to the intelligence test as separate from that given to the standardized test; others did not. Thus, time devoted to both is collapsed here.

TABLE 33

Hillview School - Littleton District
Distribution of Testing Costs Per Year
By Subject

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	TOTAL DOLLAR VALUE (% Total)
Reading		$ 8823	$ 230	$ 9053 (19.2%)
Mathematics		$13001	$ 713	$13714 (29.1%)
Language Arts		$ 2984	$ 459	$ 3443 (7.3%)
Spelling		$ 6308	$ 284	$ 6592 (14.0%)
Social Studies		$ 1492		$ 1492 (3.2%)
Science		$ 2131		$ 2131 (4.5%)
Health - Phys. Ed		$ 384		$ 384 (0.8%)
Other, Miscellaneous		$ 426	$ 924	$ 1350 (2.9%)
Multi-Subject	$1602	$ 7076		$ 8678 (18.4%)
TOTAL by category (% Total)	$1602 (3.4%)	$42625 (90.5%)	$2610 (5.5%)	$46837
			District Office Testing Costs (0.52%)	+ $248
			TOTAL	$47085

Urban schools with low-income students are often portrayed as trouble environments. Cityside, however, is among the many Metro elementary schools that belie this stereotype.

The mean length of professional staff employment at the school was just under six years. Overall, the faculty averaged fourteen-and-a-half years in the field of education. The core of urban veteran teachers who managed Cityside's programs cited the "strong, experienced" faculty as a strength of the school. The Cityside principal concurred. (Although new to the school in 1980-81, he had many years leadership in other Metro District schools.)

The staff found their students capable and easy to work with. As one program coordinator put it, "we have a fairly good student body; it's not a rough school." Another with experience in other Metro schools described her Cityside position as "a plum."

The average income level of students' families qualifies Cityside for compensatory-education and other special funding under a variety of federal, state, and district categorical programs. Chief among these are the federally sponsored Chapter I (formerly Title I) program and various supports for bilingual education. These and others provide support for additional personnel who assist Cityside's thirty classroom teachers. Three-hour-a-day aides (or paraprofessionals) are available. Special program funds also support a reading resource teacher and her aide, Chapter I and bilingual program coordinators, and specialists who respond to children with special learning needs.

Among the many Metro District elementary schools with compensatory education funding, Cityside ranked in 1979-80 among the top 2% in

reading achievement. Its sixth-grade median on the Comprehensive Tests of Basic Skills (CTBS) was then at the 56th percentile, compared to a median of the 31st percentile for all Metro District's compensatory education. Its scores declined to the 38th percentile in 1980-81, but they remained above the district-wide median for schools with compensatory programs (32nd percentile, based on schools' sixth-grade medians).

Cityside's testing program shows somewhat more classroom variation than Hillview's. This occurs largely because Cityside's teachers have greater discretion over curricular testing in reading and math. Table 34 displays the tests routinely given at Cityside Elementary.

Cityside's Testing Costs in Overview

Table 35 provides a comprehensive look at Cityside's yearly testing costs. The distribution of costs is quite similar to Hillview's. The chief differences are: (1) unlike Hillview, Cityside made some direct, testing-related purchases; (2) Cityside's indirect costs in administrative time were higher; and (3) their costs in personnel time were distributed across a greater number of kinds of staff.

As in the Hillview overall cost accounting (Table 25), the first item in Table 35 carries district-office costs forward for Cityside's 830 pupils.

Direct dollar outlays come next in the itemization of Cityside's testing costs. At the principal's request, Metropolitan Achievement tests were given annually, at a cost of $1200 per year. A basal reading series was supplemented at Cityside with the Metro District's

TABLE 34

Cityside Elementary School Testing Program

Test	Grade(s)	Required by:	Administrations Per Year
Multi-Subject			
Metropolitan Achievement Test†	1 - 6	Principal	1
Comprehensive Test of Basic Skills (CTBS)	3, 5	District	1
CTBS-Español	1, 2	District	1
District Continuum Basic Skills Survey*	1 - 6	District	1
State Assessment Program	3, 6	State	1
Reading			
District Reading Program†	K - 6		3 - 10
San Diego Quick Assessment†	1 - 5		1
Math			
Teacher-constructed math tests or those included in "Math for Individual Achievement" text	1 - 6		variable
Spelling			
Teacher-constructed spelling tests; some use of commercially available word lists†	1 - 6		weekly
Language Competence			
Basic Inventory of Natural Language (BINL)	K	District	2
Moreno (Assessment of Second Language Acquisition)	K	State	1
Physical Education			
Physical Performance Test	5	State	1

† Test widely administered but not in every classroom

* The District Continuum-Based Skills Survey is required by the district at every grade. Items vary from grade to grade, covering district-defined "essential skills." The tests at grades 3 and 6 function to fulfill state requirements for minimum competency testing (and are counted as such in the following cost itemizations), although they are no different in design than those given at grades 1, 2, 4, and 5.

TABLE 35

Total Costs For All Achievement Testing In Cityside School - Metro District

Enrollment = 830

	Hours/Year (% Work Time)[2]	Dollar Equivalents[3]
District-Office Costs[1]:		
$264 Per Pupil x 830 Pupils		$ 2,191
Direct Costs to School:		
Purchase of Metropolitan Achievement Test		1,200
Purchase of Curricular Reading Tests		5,000
Purchase of Scantron Scoring Machine Forms		200
		$ 6,400
Direct Costs for School (Personnel Time):		
Administrators/Coordinators -		
Reading Resource Teacher	328.5 (19.3%)	$ 5,790
Title I Program Coordinator	11.5 (0.7%)	210
Teacher Testing Coordinator	35.0 (2.1%)	472
	375.0	$ 6,472
Clerical/Secretarial	10.3 (0.5%)	$ 95
Classroom Teachers -		
Average Time Per Teacher	199.2 (12.2%)	$ 2,745
	x 30	x 30
	5976.0	$ 82,350
Instructional Specialists[4] -		
Bilingual Coordinator	156.25 (9.2%)	$ 2,760
Bilingual Teacher (assists with testing)	8.08 (0.5%)	112
	164.33	$ 2,872
Instructional Aides (Paraprofessionals) -		
Aide to Reading Resource Teacher (n=1)	109.45 (20.6%)	$ 657
Aide to Instructional Specialist (n=1)	4.58 (0.9%)	27
Classroom Aides (per classroom)	39.48 (7.8%)	237
Number of Classrooms	x 30	x 30
	1184.40	$ 7,110
TOTAL AIDES	298.5	$ 7,794
Classroom Volunteers	92.2 (??)	-----
Student Time[5] (Average Time Per Pupil)	76.1 (8.6%)	-----
TOTAL COSTS FOR SCHOOL (1981-1982 School Year)		$108,174
AVERAGE COSTS PER CLASSROOM (n=30; avg. 27.67 pupils/class)		$ 3,606
COSTS PER PUPIL		$ 130.33
PROPORTION OF DISTRICT ANNUAL EXPENDITURE PER CHILD (= $3300)		3.9%

TABLE 35 (continued)

[1] Calculations of district office costs were previously shown.

[2] The "% Work Time" figures are based on respondents' report of hours worked per week before, during, and after school hours. These reported hours per week were averaged by role category across the two schools studied (Cityside and Hillview). Reported hours were within similar ranges at both schools. Work times used are as follows:

(a) For administrators, coordinators, and instructional specialists: 46 hours per week x 37 weeks per year.

(b) For clerical/secretarial personnel: 40 hours a week (roughly 22.5 work days or 180 work hours per month) x 11 months per year.

(c) For classroom teachers: 44 hours per week x 37 weeks per year = 1628 hours per year.

(d) For instructional aides: 3 hours per day per classroom x 177 school days per year = 531 hours per year per classroom.

(e) No total hours per unit or person could be ascertained for volunteers.

[3] Dollar equivalents are based upon the proportion of work time expended at the following salary estimates:

(a) For administrators and coordinators - $30,000 salary and fringe benefits.

(b) For clerical/secretarial - $20,000 salary and fringe benefits.

(c) For classroom teachers and instructional specialists (except coordinators) - $22,500 salary and fringe benefits.

(d) For instructional aides - $6.00 per hour.

Salaries listed under (a) are somewhat lower than the actual compensation afforded at this school, but are equivalent to estimates used in the analysis of district costs.

[4] Instructional specialist time reported is devoted to coordinating and conducting achievement testing for bilingual students.

[5] Student time shown equals the time spent by the typical student in each classroom averaged across the school's regular classrooms. The percentage shown is based on 5 class hours per day (not counting the hour for lunch and recess) for 177 school days per year, which equals 885 classroom hours per school.

skills-oriented reading program. It was accompanied by consumable tests, costing $5000 annually. The school also had a Scantron scoring machine, but teachers used it infrequently.

Administrators/coordinators of Cityside's school-wide testing spent 375 test related hours in 1981-82. They performed many of the same testing-related tasks as Hillview's administrators, but Cityside's greater enrollment meant that certain tasks took longer. Furthermore, special-program funding allowed Cityside coordinators to support classroom teachers' assessment efforts in a wider range of ways.

The reading resource teacher's work illustrates the latter point. She managed a "retrieval room" from which classroom teachers could obtain the supplementary District Reading Program materials. She ordered the tests accompanying this program, periodically inventoried them, and conducted staff development sessions in how to use the tests and associated record-keeping forms. When class teachers needed a specific test, the reading resource teacher located it and signed it out. During 1981-82, these activities consumed 279 of the 328.5 hours that the reading resource teacher spent on testing.

Another of her responsibilities was to help proctor classroom testing. She spent 10 hours proctoring when the District Continuum-Based Skills Survey was given and another 10 hours during CTBS testing in grades 3 and 5. Before the skills survey was administered, their reading resource teacher gave a one-hour in-service session reviewing test administration procedures with teachers and aides.

Finally, the resource teacher saw to the purchase and distribution of the Metropolitan Achievement Test. She also answered faculty questions on how to administer and score it. These tasks required 18.5 hours of her time at the outset of the school year.

The Cityside Chapter I Program Coordinator assumed primary responsibility for the District Continuum-Based Skills Survey. He obtained the requisite test forms from the district's testing office (three hours), secured extras when shortage appeared (fifteen minutes), "oriented" new teachers to Skills Survey administration procedures (one hour), and planned the school-wide schedule for Skills Survey testing with the teacher testing coordinator (two hours). He gave another two hours to "scheduling the set up and orientation" for teachers, and another half hour to arranging for supervision of half of the teachers' classes while the other half was being tested.* Helping check over students' answer sheets, and alphabetizing and packaging them for scoring took another 70 minutes of the coordinator's time, for a total of almost 10 hours on Skills Survey testing.

The chapter I coordinator also devoted an hour-and-a-half annually to consulting with the reading resource teacher about her orders for test materials and passing those orders on to be typed. Finally, he gave about twenty minutes to answering teachers' questions about the State Assessment measures.

A first-grade teacher at Cityside oversaw school-wide testing.

---

* Metro District recommended that teachers test one-half of their class at a time, in order to assure an environment more conducive to concentration.

This involved distributing appropriate numbers of tests and answer sheets to teachers, collecting test materials after administration, checking over answer sheets for correct identification information, etc. She also responded to teachers' procedural questions. Altogether, the teacher testing coordinator invested 35 hours in these tasks during the year of inquiry. In all, coordination of testing consumed 375 hours of administrators' working time in 1981-82. In addition, the reading resource teacher's aide assisted her with all of her testing-related responsibilities, adding an extra 109.45 hours to the staff's investment in test coordination. (See the item headed "Instructional Aides" in Table 35.) The total, 484.45 hours per year, far exceeded the time (99.75 hours) Hillview administrators spent coordinating and facilitating school-wide testing. On a per pupil basis, however, the difference appears less great: .58 hours per pupil at Cityside; .52 hours per pupil at Hillview. Significantly, the administrators'/coordinators' time spent at Cityside did not include extending the analyses of scores that were returned to the school. (Recall that Hillview's principal spent his time developing year-to-year comparisons for grade levels and individual classrooms.) Instead Cityside administrators and coordinators spent more time facilitating the test-administration process. Conducting assessment in the supplementary District Reading Program, together with the more complex testing logistics in the larger school, made this necessary.

Clerical time was also a cost of testing at Cityside. Over the course of the year, secretarial staff spent 10.3 hours preparing the orders for the tests that the school purchased.

Teacher time at Cityside was mostly given over to the same type of activities found at Hillview. And again at Cityside, there was substantial variation in the time per teacher per year allocated to testing. Seventeen of Cityside's thirty classroom teachers were interviewed during the study.* The total time each spent on testing is displayed in Table 36 below (teachers names are again fictitous).

Teachers' annual testing hours spanned a greater range at Cityside than at Hillview (55.0-501.5 at Cityside; 151.7-395.8 at Hillview). Moreover, the within-grade variation is much larger at Cityside, because teachers had greater latitude in deciding how to assess student progress in reading and math. Unlike Hillview, there were no required curriculum-embedded tests in these subjects at Cityside. Further, though Cityside teachers used common reading materials, they tended to use them differently--from daily to once or twice a week. Greater use of the materials meant students' passed through program "steps" or "levels" more rapidly, and so were tested more often with program instruments.

* Although informed consent for participation in the study was gained from Metro District and Cityside School, eight Cityside teachers declined to be interviewed. Six others professed willingness to assist in the research and scheduled interviews, but their other responsibilities recurrently kept them from keeping these appointments. As a consequence, the cost accountings that follow are based upon data reported by the seventeen teachers, supplemented by estimates for those teachers who were not interviewed. In each case, the estimates were made by ascribing the mean number of hours reported by teachers at each grade level to the teachers at that grade level who were not interviewed. Further, this estimated time was divided for each non-interviewee by test type, subject matter, and mandate based on the mean proportions of time allocated to each test type, subject matter, and mandate by teachers at the non-interviewees grade level.

TABLE 36

Total Time Spent on Testing by Cityside Teachers: 1981-82

Teacher (Grade)			Hours Per Year
Gonsalves	(K)		377.16
Lehrman	(K)		55.00
White	(1)		167.95
Jackson	(1)		56.38
Irvine	(1)		87.00
Prickett**	(2)	153.08	
Prickett	(1)	161.83	314.90
Moy	(2)		331.81
Hillsen	(2)		198.10
Washington**	(2)	100.46	
Washington	(3)	146.00	246.46
Benson	(3)		262.70
Krupp	(4)		299.41
Belendez	(4)		113.41
Faschinna	(5)		107.11
Ewing	(5)		248.63
Leiderman	(5)		85.91
Berriman	(6)		105.90
Smith**	(4)	155.96	
Smith	(5)	185.23	
Smith	(6)	160.40	501.59

* Teaches multi-grade class. Time spent on testing shown for each grade.

At Hillview, on the other hand, team teaching tended to reduce the amount of within-grade variation. In the fifth grade at Hillview, for example, one teacher did all the teaching and testing for both classes in math and science; the other, in reading and social studies. Teachers in other grades conjointly planned so that instructional schedules and rates of progress were similar. This was not true at Cityside.

Finally, some of Cityside's within-grade variation in testing time per teacher per year is ascribable to differences in both the instructional and assessment programs for limited-English-proficient and fluent-English-proficient students. Students who spoke primarily Spanish, for example, worked in a Spanish-language version of the District Reading Program through their early grades, and they were tested on a different schedule than students using the program's English-language version. Limited-English-proficient kindergarten children were given individually administered oral measures that fluent English-speakers were not required to take. As the number of limited-English-proficient youngsters in a class increased, so did the teacher time spent administering tests.

The distribution of Cityside teachers' annual testing time was quite similar overall to that at Hillview. "After testing" activities consumed the greatest proportion of Cityside teacher's time across the year (mean percentage = 53.5). But the mean proportion of time spent by Cityside teachers "during testing" (27.8%) was less than at Hillview (34.2%). And by roughly the same proportion, Cityside teachers' "before testing" time was greater (mean percentage = 18.7%

as compared to 10.9% at Hillview). Cityside's teachers spent more time, on the whole, preparing for classroom test administration. Several factors underly these differences.

First, Cityside teachers collectively devoted a larger proportion of their total testing time to teacher-constructed tests. Time taken to design and duplicate these measures was counted here in the "before testing" category. Second, pre-administration logistics -- in-service training or orientation, obtaining appropriate numbers of test forms, etc. -- consumed more time at Cityside than at Hillview. Third, more Cityside teachers spent time before testing, reviewing with students the skills to be tested and practicing test-taking skills. A summary of the main findings of Cityside teachers' 1981-82 testing time allocation appears in Table 37.

Instructional aides (or paraprofessionals') time on testing provided a substantial supplement to that of teachers' at Cityside. As Table 37 shows, Cityside teachers allocated a mean of 199.2 hours per year to obtaining test results. This compares to a mean of about 253 hours across the Hillview faculty. But as Table 35 indicated, Cityside's classroom aides supplied (on the average) another 39.4 hours a year of staff testing time to each Cityside class. Combining their time with teacher time increases the average to 238.7 hours per year of staff assessment time in each classroom.* Thus, the classroom staff testing time difference between Cityside and Hillview is not as great as initially appeared.

* Note, too, that Cityside students (again, on the average) receive fewer hours of testing per year than Hillview students. Using means, the ratio of staff to student hours of testing is 3.13:1 at Cityside; it is 2.87:1 at Hillview.

Aides' time is less costly than teachers': use of aides leads to savings in indirect testing costs. The Cityside aides' mean time of 39.4 hours per class per year cost only $237 at aides' hourly rates. In teachers' salary, the same amount of time per class per year would have had a dollar value of about $546.

A good deal of the classroom aides' time was devoted to tasks before and after the test-administration episode. Altogether, Cityside aides spent a mean of 26.5% (or about 10.5 hours) of their annual time on "before testing" activities -- including duplicating teacher-constructed tests, assisting in instruction explicitly undertaken for test preparation, procuring appropriate test forms for the class, etc. On the average they gave another 32.2% (12.7 hours per class) over the year to such "after testing" tasks as grading tests and quizzes, recording scores, returning tests to students, and checking over answer sheets prior to machine scoring. In all, then, about 58.7% of aides' testing-related time was allocated to tasks outside the test-administration episode. Still, Cityside aides spent a substantial proportion of their time on testing in the "during" phase. (Mean for classroom aides = 16.2 hours, or about 41.3% of their mean total time.)* During test administration they might supervise or instruct sub-groups of students not being tested at the moment, and/or proctor the test-taking group.

Aides also spent time on such routine activities as distributing and collecting test booklets and answer sheets, answering students'

* On the average aides spent a higher proportion of their testing-related time in the "during" phase of testing than did teacher (mean proportion = 27.8% of teachers' mean total testing time).

TABLE 37

Summary of Cityside Classroom Teachers' Time on Testing

Mean number of hours given to:

"before testing" activities	37.2 (18.7% of total)
"during testing" activities	55.3 (27.8% of total)
"after testing" activities	106.5 (53.5% of total)

Total: Mean Number of Hours per Teacher per Year: 199.2

Proportion of Average Total Annual Work Time* = 12.2%

Range: 55.0-501.5 hours

*Calculation of average total annual work time is explained in a footnote to Table 35 above.

procedural questions, and helping to re-arrange student seating at the outset and the conclusion of the administration period.*

Classroom volunteers' testing time was consumed by the same types of responsibilities often assigned to aides at Cityside. In at least two cases, volunteers shared testing tasks with both the classroom teacher and an aide.

The testing time of the instructional specialist** at Cityside was allocated exclusively to assessment of non-English-proficient and limited-English-proficient learners. The bilingual coordinator conducted CTBS-Espanol testing for grades three through six students whose English-Language competence was insufficient for them to take other school-wide multi-subject measures. She also adminstered the Basic Inventory of Natural Language (BINL) throughout the year as new students who qualified for language assessment arrived at Cityside. In addition, she taught Spanish readers in a daily class, assessing their oral and written language skills on a weekly basis. A bilingual first-grade teacher also contributed a small amount of her annual time to administration of CTBS-Espanol. In all, instructional specialists spent 164.3 hours annually on these activities.

Student time on testing average 76.1 hours per student per year across Cityside's thirty classrooms. Calculating annual class time at 885 hours (see Table 35 footnotes), this equals 8.6% of the yearly time available for classroom learning.

* Recall that by the definition in use here, these activities are all part of the test administration episode.

** The testing time of instructional specialists who taught learning disabled youngsters is omitted here as outside the domain of inquiry.

Cityside students spent most of their assessment-related time in test-administration episodes. Mean hours per student per year in the "during" phase of testing equaled 41.7. This constituted 54.9% of the mean annual total of 76.1 -- substantially less than for Hillview students, where "during testing" activities consumed neary 91% of students' average annual testing time. Conversely, Cityside students spent more of their time on testing before and after test administration. The typical Cityside pupil devoted 10.8 hours per year (14.3% of the mean total) getting ready to take tests and 23.4 hours yearly (30.8% of the mean total) on such "after testing" activities as in-class grading and "going over" the results of teacher-scored tests. Hillview children, in contrast, spent only 9.4% of their assessment-related time in the before and after administration phases.

Overall, Cityside's yearly costs for testing totaled $108,174. Of this total, all but $6,400 were incurred indirectly, i.e., in the dollar values of paid staff members' time. Therefore, a little over 94% of Cityside's annual testing costs were indirect, personal-time items. The magnitude of the total is put in perspective by considering it on a per-pupil basis. Cityside's assessment cost per child came to $130.33 in 1981-82. The Metro School District expended $3300 per student in that school year; Cityside's per pupil testing costs come to 3.9% of this figure.

The per-pupil costs of testing at Cityside were substantially less than those at Littleton District's Hillview School ($246.56 per student). On the one hand, Cityside's testing "expenses" were higher in several areas: District-office costs per pupil; administrators,

coordinators, and clerical time; direct purchases. (Hillview had no costs in the last two categories.) But in view of the entire testing "budget," these costs were only fractionally higher at Cityside.

On the other hand, Cityside teachers spent less of their annual work time on testing than Hillview teachers. And Cityside's use of aides resulted in savings. The factor most relevant to the per-pupil cost differential between the two schools, however, was the number of students per classroom. The number at Hillview averaged about 18; while at Cityside the average was about 28. But the ratio of classroom staff to student hours on testing was similar at both schools: 3.13:1 at Cityside, 2.87:1 at Hillview. To provide an hour of testing to a class, the instructional staff at both schools spent roughly the same time, but that hour of testing at Cityside was delivered each time to an average of about 10 more students than at Hillview. Therefore, Cityside's per-pupil annual testing costs, because more pupils were tested on each occasion, were lower than at Hillview's. Using aides and devoting fewer hours of testing per pupil per year also contributed to Cityside's lower per pupil costs.

Cityside's Costs for Required and Non-Required Testing

This section of the report discusses Cityside's cost distribution for achievement testing for mandated and discretionary testing, by test type and subject area.

Table 38 itemizes Cityside's staff-time assessment costs by source of mandate. Table 39 converts these to dollar values and incorporates costs of other kinds. (Table 34 shows which tests are required by each source.)

Cityside's staff-time required testing costs were rather low, and markedly lower than Hillview's. At Hillview, 54.2% of staff testing time (and 50.2% of teachers' alone) was given over to mandated testing. Even excluding Hillview's District-mandated reading and math testing, 31% of staff testing time was invested in required measures. At Cityside, by contrast, the proportion of staff time on required assessment was a little under 15% and about 12% for classroom teachers.

The distribution of testing dollars in Table 39 reflects the staff-time allocation; the addition of Cityside's costs for testing purchases does little to change the overall picture. Some 83.3% of Cityside's annual testing costs were allocated to measures given at teachers' discretion.

Cityside's Costs for Different Types of Testing.

Tables 40 and 41 show the distribution of Cityside's costs for testing of different types. (The test-type categorization system is identical with that used in discussing Hillview's costs, and each category is described in that discussion.)

Recall that the same series of tests (the District Continuum-Based Skills Survey) falls under two categories in Tables 40 and 41. At grades 3 and 6 the Skills Survey functioned to meet state requirements for minimum competency testing. At grades 1, 2, 4, and 5, the Skills Survey is counted as a District Continuum test. (At all grades, the Skills Survey assessed students' learning of skills on district reading, math, and language arts continua that have been designated as "essential".)

TABLE 38

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME PER YEAR
On Required and Non-Required Testing*

Each staff category cell shows:
° No. of staff members involved
° Avg. hours/staff member/year
° % Total testing time for staff by category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Paraprofessionals) TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS
Required by State	2 3.14 7.1%		17 22.1 6.3%	1 74.0 45.0%	11 9.39 8.0%		559.20 7.1%	15.0	17
Required by District	3 19.97 22.6%		20 11.73 3.9%	2 8.2 10.0%	22 3.5 6.0%	1 5.2 5.6%	393.90 5.0%	8.6	20
Required by School Principal	3 9.83 7.8%	1 .50 4.9%	23 4.9 1.9%		23 2.61 4.6%		202.2 2.5%	2.4	23
TOTAL REQUIRED (In person hours)	95.7 25.5%	.50 4.9%	722.4 12.1%	90.33 55.0%	241.16 18.6%	5.2 5.6%	1155.30 14.6%	15.0	30
NOT REQUIRED (In person hours)	279.33 74.5%	9.8 95.1%	5252.9 87.9%	74.0 45.0%	1057.29 81.4%	87.0 94.3%	6760.32 85.4%	61.1	30
TOTALS by staff category (In person hours)	375.00 100.0%	10.3 100.0%	5975.32 100.0%	164.33 100.0%	1298.5 100.0%	92.2 100.0%	7915.6		

* Required testing includes any testing mandated by someone or some agency in the organizational hierarchy above the classroom teacher. Testing required exclusively to meet federal education program requirements has been waived for Metro District.

TABLE 39

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
Required & Non-Required Testing

TYPES OF TESTING	DIRECT DOLLAR COSTS	ADMINIS-TRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	TOTAL DOLLAR VALUE (% Total)
Required by State		$ 110		$ 5188	$ 1292	$ 624	$ 7214 (6.7%)
Required by District		$ 1036		$ 3212	$ 287	$ 468	$ 5003 (4.6%)
Required by School Principal	$ 1200	$ 505	$ 5	$ 1565		$ 358	$ 3633 (3.4%)
TOTAL Required	$ 1200	$ 1651	$ 5	$ 9965	$ 1579	$ 1450	$ 15850 (14.7%)
TOTAL Not Required	$ 5200	$ 4821	$ 90	$72385	$ 1293	$ 6344	$ 90133 (83.3%)
TOTAL by category	$ 6400	$ 6472	$ 95	$82350	$ 2872	$ 7794	$105983
(% Total)	(5.9%)	(6.0%)	(0.09%)	(76.1%)	(2.6%)	(7.2%)	(2.0%)
						Plus District Office Costs	2191 (2.0%)
						TOTAL	$108174

Overall, Cityside staff gave the largest proportion of their assessment time (46.5%) to teacher-constructed measures. Over half of classroom teachers' time on testing occurred in conjunction with these. Another 38.3% of the staff's time was allocated to testing with commercial, curriculum-embedded measures. (Most of the aides' time was spent on these.) The average time spent on testing per student per year was also highest for these two types of measures.

As Table 41 indicates, 82.5% of Cityside's direct and indirect costs were incurred for teacher-constructed and commercial, curriculum-embedded testing. This was higher than at Hillview, but there commercial and teacher-made curricular measures still consumed a substantial 67.3% of the annual testing resources. (As Table 30 showed, the Hillview staff-time commitment was larger for commercial curricular testing and lower for teacher-constructed-tests -- just the reverse of Cityside's.)

Cityside's Cost for Testing in Different Subject Areas

The distribution of Cityside's staff-time on assessment in different subjects is displayed in Table 42. Table 43 converts these to dollar values and adds direct-purchase testing cost.

As at Hillview, Cityside's staff-time testing costs were concentrated in the basic skills subjects of reading, math, and spelling. Also, as at Hillview, Cityside invested a substantially lower proportion of staff total testing time to the basic skills of language arts (grammar, writing, oral communication -- but excluding spelling here)

TABLE 40

Cityside School - Metro District
Distribution of Staff & Student Testing Time Per Year
By Type of Test

Each staff category cell shows:
- No. of staff members involved
- Avg. hours/staff member/year
- % Total testing time for staff category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Paraprofessionals) TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT#(hours)	NUMBER OF CLASSROOMS
Standardized, Norm- Referenced (Grades 1-6)	3 15.83 12.7 %	1 0.50 4.9 %	20 11.62 4.0 %	2 8.16 9.9 %	22* 4.49 7.6 %	2 2.6 5.6 %	400.7 5.1 %	5.54	20
State Assessment Program (Grades 3,6)	2 3.14 1.7 %		8 3.32 0.4 %		2 0.89 0.14%		34.62 0.43	2.4	8
Minimum Competency (Grades 3,6)			8 6.10 0.8%		2 5.98 0.9 %		60.79 0.76%	5.5	8
District Continuum (Grades 1,2,4,5)	3 13.97 11.2 %		20 5.76 1.9 %		9† 4.29 3.0 %		195.71 2.5 %	4.7	20
Commercial, Curriculum-Embedded (Grades K-6)	2 139.67 74.4 %	1 9.8 95.1 %	30 69.80 35.0 %		31§ 18.25 43.6 %	3 26.95 87.7 %	3029.89 38.3 %	21.7	30
Teacher Constructed (Grades 1-6)			26 119.9 52.2%	74.0 45.0 %	26 18.8 37.7 %	1 6.16 6.7 %	3685.33 46.5 %	48.1	26
Other, Miscellaneous (Grades K-6)			20 17.14 5.7 %	2 37.0 45.0 %	9 10.28 7.1 %		509.22 6.4 %	10.3	20
TOTALS By staff category (In person hours)	375.0 100.0 %	10.3 100.0 %	5975.32 100.0 %	164.33 99.9 %	1298.5 100.0 %	92.22 100.0 %	7915.7		

* Aide time includes 18 hours spent annually by reading resource teacher's aide in coordinating and proctoring, and 4.58 hours spent in similar duties by a bilingual specialist's aide. Omitting these times, aides in 20 classrooms spend an average of 3.8 hours on standardized, norm-referenced testing.

† Aide time includes 10 hours spent annually by reading resource teacher's aide in proctoring test administration. Omitting this time, aides in eight classrooms spend an average of 3.6 hours annually on testing associated with district continuum testing.

§ Aide time includes 81.45 hours spent annually by reading resource teacher's aide in distributing, organizing, inventorying and re-ordering reading test materials. Excluding this time, aides in 30 classrooms spend an average of 16.1 hours annually on testing that is embedded with commercially available curriculum materials.

\# Note that the number of classrooms in which each type of test is administered varies; thus, the proportion of time the typical student spends on each type of test carries from classroom to classroom and the average times shown cannot be appropriately added.

TABLE 41

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
By Type of Testing*

TYPES OF TESTING	DIRECT DOLLAR COSTS	ADMINISTRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Paraprofessionals) TIME	TOTAL DOLLAR VALUE (% Total)
Standardized Norm-Referenced (Grades 1-6)	$ 1200	$ 822	$ 5	$ 3294	$ 287	$ 592	$ 6200 (5.7%)
State Assessment Program (Grades 3,6)		$ 110		$ 329		$ 11	$ 450 (0.4%)
Minimum Competency (Grades 3,6)				$ 659		$ 71	$ 730 (0.7%)
District Continuum (Grades 1,2,4,5)		$ 725		$ 1565		$ 234	$ 2524 (2.3%)
Commercial, Curriculum-Embedded (Grades K-6)	$ 5000	$ 4815	$ 90	$28822		$ 3398	$42125 (38.9%)
Teacher Constructed (Grades 1-6)				$42987	$ 1292.50	$ 2935	$47214.5 (43.6%)
Other, Miscellaneous (Grades K-6)	$ 200			$ 4694	$ 1292.50	$ 553	$6739.50 (6.2%)
TOTAL by category (% Total)	$ 6400 (5.9%)	$ 6472 (6.0%)	$ 95 (0.09%)	$82350 (76.1%)	$ 2872 (2.6%)	$ 7794 (7.2%)	$105,983
						District-Office Costs† (2%)	$ 2,191 (2.0%)
							$ 108174

* Costs of staff time are calculated by multiplying percentage of staff time spent per category or cell by total dollar equivalent for staff category.

† District Office Costs pro-rated for Cityside School ($2.64 per pupil x 830 pupils = $2191). These costs cannot be apportioned exactly by test type for Cityside Elementary, but see previous discussion of how Metro District resources are allocated across different parts of the district-wide assessment program.

than the other basic skills.* Another similarity between the two schools -- a corollary to the basic-skills testing emphasis -- was evident in the comparatively low allocation of staff time to science and social studies testing.

Cityside's staff-time commitment in multi-subject testing was about half Hillview's (9.4% as compared to 18.6% of total annual staff assessment time).**

Summary and Discussion

Formal interviews and supplemental fieldwork at two elementary schools provided a comprehensive picture of their annual monetary costs for achievement testing. Findings of principal interest follow.

Overall Costs

- At a large, urban elementary school (Cityside) serving a low-income enrollment of 839, annual costs for achievement testing of all types in all subjects were $108,174, or $130.33 per pupil.
- At a small, suburban elementary school (Hillview) serving a relatively high-income enrollment of 191, annual costs for achievement testing of all types in all subjects were $47,085, or $246.52 per pupil.
- Nearly all of these costs were incurred indirectly as a result of staff time spent on testing.

---

* Many teachers interviewed at both schools expressed a preference for non-test assessment strategies in language arts, but interviewers were asked to include regular, formal writing assignments among language arts testing.

** Multi-subject tests at Cityside included the Metropolitan Achievement Test, the Comprehensive Tests of Basic Skills, the District Skills Survey, and State Assessment measures. The last two of these cover exclusively basic-skills subjects; the first two concentrate heavily upon them.

TABLE 42

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME
By Subject

Each staff category cell shows:
° No. of staff members involved
° Avg. hours/staff member/year
° % Total testing time for staff category

SUBJECT AREAS	ADMINISTRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS Total = 30
Reading	2 139.66 74.5%	1 10.3 100.0%	28 54.61 25.6%	1 74.0 45.0%	26 15.31 30.7%	1 11.67 12.6%	2302.42 28.8%	9.43	28
Mathematics			27 67.58 30.5%		25 15.51 29.9%	2 33.06 71.8%	2278.38 28.6%	21.01	27
Language Arts			16 25.42 6.8%		10 3.63 2.8%		443.0 5.5%	18.71	16
Spelling			22 54.25 20.0%		18 11.17 15.5%	1 9.17 10.0%	1403.67 17.6%	25.83	22
Social Studies			10 17.65 2.9%		6 4.12 1.9%		201.20 2.6%	10.33	10
Science			5 16.4 1.4%		2 0.63 0.09%		83.25 1.0%	4.33	5
Health - Phys. Ed			6 16.55 1.7%		6 9.52 4.4%		156.47 2.0%	30.28	6
Other, Miscellaneous			6 40.27 4.0%	1 74.0 45.0%	4 10.34 3.2%		356.96 4.5%	0.39	6
Multi-Subject	3 31.90 25.5%		26 16.24 7.1%	2 8.16 10.0%	28 5.39 11.6%	2 2.6 5.6%	690.45 9.4%	9.62	26
TOTALS By staff category (In person hours)	375.0 100.0%	10.3 100.0%	5975.32 100.0%	164.33 100.0%	1298.5 100.09%	92.22 100.0%	7915.8		

TABLE 43

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
By Subject

TYPES OF TESTING	DIRECT DOLLAR COSTS	ADMINIS-TRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Par-aprofession-als) TIME	TOTAL DOLLAR VALUE (% Total)
Reading	$ 5000	$ 4822	$ 95	$ 21081	$ 1292.50	$ 2393	$34683.50 (32.1%)
Mathematics				$ 25117		$ 2330	$27447 (25.4%)
Language Arts				$ 5600		$ 218	$ 5818 (5.4%)
Spelling				$ 16470		$ 1208	$17678. (16.3%)
Social Studies				$ 2388		$ 148	$ 2536 (2.3%)
Science				$ 1153		$ 7	$ 1160 (1.1%)
Health-Phys. Ed.				$ 1400		$ 343	$ 1743 (1.6%)
Other, Miscellaneous	$ 200†			$ 3294	$ 1292.50	$ 249	$ 5035.50 (4.6%)
Multi-Subject	$ 1200	$ 2749		$ 5847	$ 287	$ 904	$ 9888 (9.1%)
TOTAL by Category (% Total)	$ 6400 (5.9%)	$ 6472 (6.0%)	$ 95 (0.09%)	$ 82350 (76.1%)	$ 2872 (2.6%)	$ 7800* (7.2%)	$105989
						Plus District Office Costs	$ 2191 (2.0%)
							$108180

† Expenses for scantron scoring forms are ascribed to "other miscellaneous" category.

* Total is slightly larger for this category than in previous tables as a result of rounding off percentages in Table 38. (Dollar amounts here are based upon those time allocation percentages.)

- The single largest item in each school's annual testing "budget" was the time that classroom teachers gave to assessment, an indirect cost of testing borne by the school districts.

  (Teacher time on assessment as a proportion of total annual testing costs: Hillview = 90.5%; Cityside = 76.1%.)

Staff Time

- Total administrator/coordinator time per year on testing:

  Hillview = 99.7 hours/year
  .5 hours/year/pupil

  Cityside = 375 hours/year
  .5 hours/year/pupil

- Mean annual time per teacher per year on testing:

  Hillview = 252.9 hours (15.5% annual mean work time)
  Cityside = 199.2 hours (12.2% annual mean work time)

- Paid para-professional (aide) time per classroom per year:

  Hillview = none present
  Cityside = 39.4 hours

- Volunteered time (both schools) and clerical time (Cityside) were incidental in magnitude.

- Classroom teachers at both schools spent more than two-thirds of their testing-related time in activities preceding and following the testing episode.

Distribution of Teacher Time

- Proportion of total teacher time per year on testing required by supraordinate individuals and agencies:

  Hillview = 5.2%
  Cityside = 12.1%

- Types of testing consuming greatest proportions of teachers' testing time:

	Hillview	Cityside
Teacher-constructed	21.9%	48.4%
Commercial curriculum	45.1%	35.0%
Norm-referenced, standardized batteries	13.5%	4.0%

- School subjects receiving largest proportions of teachers' annual testing time:

	Hillview	Cityside
Reading	20.7%	25.6%
Math	30.5%	30.5%
Spelling	14.8%	20.0%
Multi-subject test batteries	16.6%	7.1%

Student Time

- Average time per student per year spent on all achievement testing in all subjects (and percent total annual classroom instructional time of 885 hours):

  Hillview = 88.0 (9.9%)
  Cityside = 76.1 (8.6%)

- Average time per student per year on testing required by individuals and agencies supraordinate to the classroom teacher (and percent of mean total):

  Hillview = 44.4 hours (50.5%)
  Cityside = 15.0 hours (19.7%)

- Average student time per testing per year on subjects in which typical student spends most testing time (shown in hours per year):

	Hillview	Cityside
Reading	12.1	9.4
Math	25.1	21.0
Spelling	19.3	25.8
Multi-subject test batteries	23.9	9.6

The findings reported have provided a first, comprehensive look at the magnitude of elementary schools' testing costs. And they yield a detailed portrait of how much time teachers and students spend on testing of different types.

The findings become more useful, however, when one has some sense of whether the magnitude and distribution of these particular two schools' testing costs are typical or unique. We can address this issue in a general way by juxtaposing the findings from our two case-study elementary schools with our national survey data.

Upper elementary grade teachers in the national survey were asked to "compile a list of tests given to assess or evaluate your students" in reading and math. Teachers were directed to report the number of times per year a "typical student" took each test listed and the "approximate time for (the) typical student to complete one." Teacher responses offer a national view of students' annual testing time in reading and math.

Table 44 juxtaposes the nation-wide and the Cityside and Hillview findings. Cityside students appear to be a fraction below the national average for reading testing. Otherwise, Hillview and Cityside (at least in math) appear to be "high testing" schools. Of course, teachers in the national survey were not asked to report student testing-related time spent before or after test adminstration. They were only directed to report test-taking time. Therefore, we adjusted their national averages to incorporate an estimate of student time before and after testing, so as to be able to compare the national with the case-study picture.

In Table 45, the survey averages for hours per student per year in reading and math testings have been adjusted upward. We made this adjustment by averaging the proportions of the mean testing time that students at Hillview and Cityside spent during test administration. This amounted to 91% at Hillview and 55% at Cityside, for an average of 73%. We then took the mean times reported on the survey as reflecting an estimated 73% of the total time actually spent on testing. We then increased this figure by the remaining 27% estimated to be given over to test-related activities occuring before and after

TABLE 44

Average Hours Per Student Per Year Spent in Reading and Math Testing:

Comparison of Hillview and Cityside to National Survey Data

	Nation-Wide	Hillview	Cityside
Reading	9.93	12.12	9.43
Math	12.47	25.11	21.01
TOTAL	22.40	37.23	30.44

TABLE 45

Adjusted Comparison

Average Hours Per Student Per Year Spent In Reading and Math Testing

	Nation-Wide	Hillview	Cityside
Reading	13.6	12.12	9.43
Math	17.08	25.11	21.01
TOTAL	30.68	37.23	30.44

the testing episode. With this "best guess" adjustment, Hillview and Cityside students appear to spend a bit less than the estimated national average on math testing. Cityside's total is quite near the adjusted national average; Hillview's is seven hours higher.

Although this comparison is admittedly rather crude, it suggest that the amount of testing in the two case-study schools (especially in the basic skills) may not diverge dramatically from the amount of testing conducted in many other elementary schools in the nation.

Additional support for this cautious claim is seen in national survey findings on the allocation of student testing time by test type. The survey showed that in the upper-elementary grades, the greatest proportions of students' annual testing time were devoted to school or teacher developed measure (35% - 37%). These figures are consonant with the findings in the two case-study schools (compare Tables 1, 30, and 40.).

We make no claim here for the generalizability of our findings. We simply suggest that until further research indicates otherwise, the levels and costs of testing reported here are probably "in the same ballpark" as the costs of testing in a good many other American elementary schools. On that basis we offer the following extrapolations.

Testing does not appear to infringe greatly on students' instructional time. Students in the two case-study schools spent about 9% - 10% of their annual classroom time on testing of all types in all subject areas. (This comes to an average of two or two-and-a-half hours per week.) Furthermore, some 60%-70% of this time was spent on

testing closely linked (in intent at least) with the content and process of teaching-learning, i.e., with teacher-constructed and commercial curricular testing. Assuming that regular assessment is an important part of good teaching, the scope of student time on testing certainly seems reasonable.

The costs of assessment in teacher time do not seem especially high. Assuming that a typical elementary teacher spends 44 hours a week on job-related activities over 37 weeks a year (as teachers in the two case-study schools reported), then about 12%-15% of their yearly work time is spent on testing. This amounts to some five-to-seven hours a week, a good bit of which is spent outside of school hours on grading tests and recording test scores. This is not a paltry amount of time. But in our case-study schools, much of this time was invested in curricular testing (about 87% at Cityside; about 70% at Hillview). And this testing was undertaken either at the teachers' consent (in the case of Hillview's commercial curricular measure in reading and math). Testing divorced from the curriculum and required by teachers' supraordinates consumed about 15%-30% of their total testing time -- or about 2% of their work time at Cityside and 5% at Hillview. As the next section of this report indicates, many teachers voiced frustrations and aggravations in conjunction with non-curricular, required types of assessment such as annual or biannual standardized testing. For teachers, such tests may entail subjective costs disproportionate to the amount of time they consume. This is certainly an important consideration. But in a literal, objective sense,

the time-costs of testing which is both required and divorced from routine teaching-learning are not large.

The direct costs of testing do not appear to be great. Even if districts and schools were cut back sharply on the amount of testing they conduct, they would not find themselves with a vast sum of re-allocatable dollars. A far greater proportion of districts' and schools' "expenses" for testing are incurred indirectly through the time staff members devote to assessment.

Eliminating mandated testing would probably save only very modest amounts of school-level educators' time. State-mandated testing at the two schools studied consumed only 5.2% (at Hillview) and 7.3% (at Cityside) of the total yearly staff hours devoted to testing. And these hours themselves constituted a small proportion of staff members' work time across the school year. District requirements comprised only another 5.6% (at Cityside) and 25% (at Hillview, excluding curricular testing requirements) of this already small proportion. The four preceding proportions suggest two issues germaine to educational testing policy.

The greatest testing cost districts and schools appear to bear is the opportunity cost of teacher time. Teachers, in turn, spend the greatest proportion of their time in curricular testing.

Districts and schools interested in economizing on assessment, therefore, should probably try to find ways to reduce the time teachers spend constructing their own tests and scoring these and other curricular measures. Item-banking and computerized test scoring and score analysis should be considered. These and similar procedures

may have large start-up costs, but over the years they could free substantial proportions of teacher time for classroom instruction.

More broadly, the issue of test quality emerges from our findings. The questions of how much testing is going on and how much it costs seem less important than the question of the quality of the tests being used. Teachers spend substantial proportions of their assessment time constructing and administering their own tests and administering curriculum-embedded measures. Teachers also report considering these tests heavily in making instructional decisions. We know very little about the quality of these types of tests. But we do know that most teachers receive little pre- or inservice training in test construction or test selection. While the monetary costs of testing seem modest, the impact of curricular test results certainly is not. The quality of curricular testing, then, merits futher attention.

PSYCHOLOGICAL COSTS OF TESTING: TEACHER ATTITUDES

In the two-case study schools, we closed out our interviews on staff members' testing time by asking a series of questions probing concerns and anxieties associated with testing. The questions were asked so as to elicit responses showing whether these anxieties were borne by teachers, students, administrators, or others. Any relevant, unsolicited commentary offered during other stages of the interview were also recorded. At Hillside we interviewed all of the school's eleven teachers. At Cityside we interviewed 16 teachers and 2 instructional students, better than half of the total instructional staff. Interviewees covered all six elementary school grades. School principals and other administrators were also interviewed.

The findings indicate that -- at least in the two schools we studied -- testing and the use of test results do not cause deep worry or distress; aggravation, perhaps some anxiety, appears to be the principal psychological cost of testing. This aggravation is reflected in teacher concerns about test utility, appropriateness of tests and their uses, testing effects, and impact on instructional time.

Test Utility

Virtually every teacher interviewed at Cityside commented, explicitly or implicitly, on the utility of some of the tests in use at their school. Fourteen teachers -- about half the instructional staff and representing all grades -- made very explicit comments suggesting that having to administer tests of little direct use to teachers is a

fairly prevalent concern at Cityside. Many of the negative comments reflected problems with tests that teachers are required to administer, usually norm-referenced or minimum competency tests, or tests associated with external reporting requirements. They range from simple statements asserting a general lack of test relevance to comments suggesting differential value of specific parts of a particular testing program.

In contrast to Cityside, teachers at Hillview made few direct comments about test utility. In fact, only two teachers at Hillview mentioned this concern.

The concerns about test utility expressed by Cityside teachers, categorized by theme, are detailed below.

Lateness of test score reports: Five staff members commented on the lateness or non-receipt of test results. Of the test required for assessing limited-English-proficient students' language dominance, the bilingual coordinator noted:

> (it has) rather dubious value. There is a delay in getting the scoring back. You wait four to six weeks to get a return (and) by the time you get the results back, you've forgotten the individual child.

Similarly, one of the first-grade teachers noted that she never sees the results of the Comprehensive Tests of Basic Skills (CTBS) Español, which is required for students in the school's bilingual classes, "nor are they ever given to the students or their teachers in the next grade." This teacher generally felt that she has to give a lot of tests but "gets nothing back." One of the grade two teachers commented that she can get the CTBS Español results if she asks for them, but "the results come back too late" to have any instructional

use. The bilingual coordinator also emphasized this problem in her comment that "the kind of test we give at the end of the school year (e.g., CTBS Español), the teachers never see the results." A third-grade teacher preferred her own tests over more formal measures because of their immediate feedback potential.

Discussing the Continuum-Based Skills Survey (CBSS), a minimum competency test administered across all grades at Cityside, one of the fifth-grade teachers noted that:

> the results come back too late. I don't know who they will benefit. (I) can't wait (for the scores) to do (student) grouping. I don't really use the test scores.

Lack of relevance or test redundancy: Six teachers at Cityside commented on the problem of test relevance or actual redundancy. For example, one of the first-grade teachers noted that the school-required Metropolitan Achievement Test (MAT) does not help her with the kinds of instructional or classroom management decisions she has to make early in the school year. However, it may later "back up what I've (already) done" in terms of decisions about student diagnosis and grouping in reading and math made on the basis of less formal measures. One other colleague in the first grade amplified this issue by asserting that there are too many tests that "basically tell me the same thing."

Concerns with lack of test relevance appeared to be a problem for some of Cityside upper grade teachers as well. Discussing the MAT, a fourth-grade teacher observed that she used this test because she:

> didn't have a choice. (I) didn't find it helpful. It was a good idea to have an achievement test, but (on) this one (the student scores were) so low. They (the students) function so much better than (the scores would indicate).

Also commenting on the MAT, a fifth-grade teacher noted that the "results aren't worth the time it takes," and went on to describe the results of the CBSS and CTBS in similar terms. According to this teacher:

> One year-end test is enough. (We) need one formalized test that is useful. Two tests (are) redundant and take time away from the program.

This concern was shared by a second fifth-grade teacher, who felt that the MAT "took too much time and I didn't agree with the results." A sixth-grade teacher also observed that the MAT "was a waste (and) I didn't agree with the results."

Two teachers at Hillview School who chose to comment on test utility offered similar remarks regarding certain tests that they were required to administer.

Differential value of parts of a testing program: Five teachers at Cityside referred to the tests associated with the Developmental Reading Program (DRP), which is used by many teachers in the school. All of these comments indicated that the teachers saw no value in administering unit pretests. Most of these teachers simply admitted that they use only the unit posttests. One of the first-grade teachers went on to justify this practice:

> I don't waste time on the pretest...I only give them the posttest (and) if they pass I move them on to the next step. If they don't pass, they go over the things they miss,...then go on to the next step. It's great for diagnostics.

It would be inaccurate to say that the pre-tests associated with the DRP create a psychological cost for Cityside teachers; teachers can simply omit them. However, that several regularly do so suggests that dollars invested in pre-tests may not be a wise investment for all teachers.

Appropriateness of Tests and Their Uses

As was the case with test utility, virtually every teacher interviewed at Cityside had something to say about the appropriateness of tests and/or the the uses to which they are put. About a dozen teachers, covering most grade levels, had concerns about test/test use appropriateness. Teacher commentary in this category, while a great deal of it was negative, also tended to show that Cityside teachers are not bothered by all forms of testing. Nor did Cityside teachers tend to single out tests as inappropriate on the basis of their generic features (e.g., norm- versus criterion-referenced).

With the teachers in Hillview, a different kind of picture emerged. Here only about half of the eleven teachers commented directly on test appropriateness. And in each case the comment reflected a concern about the manner in which a test score was used and the effect of its use on students and teachers.

Most of the Cityside comments on appropriateness fell into the following categories.

Ease/difficulty of tests Seven teachers at Cityside made statements about the ease or difficulty of a test or kind of test. In terms of minimum competency testing, for instance, the school's bilingual coordinator noted that there is a "need for a test like the CBSS,

(though) it should be more of a challenge (for the students)." One of the first grade teachers amplified this attitude toward minimum competency testing as follows:

> The CBSS, I think, should be harder...I wouldn't eliminate the CBSS, but I'd revamp (it) to where, instead of having minimal (skills), it would have maximum (competencies).

Three of the second-grade teachers agreed. One commented that the "CBSS (is) not useful. There is no worthwhile feedback." For another the test "is too easy, not valuable," while the third felt that "the Survey could be better...it doesn't tell me how far the student can go."

Similar comments were made about some of the norm-referenced tests administered at Cityside. The bilingual coordinator observed that the "CTBS Español is far more difficult (than the CBSS), which is very minimal." This specialist was very concerned about the disparity of difficulty levels between the two tests.

The first-grade teacher quoted above believed that tests <u>like</u> the Skills Survey and CTBS (i.e., minimum competency and norm-referenced) served justifiable purposes, but felt that the purposes were not adequately fulfilled by these two particular tests. Discussing the CTBS, which was once (but no longer) required on a school-wide basis, this teacher commented:

> That's one thing the CTBS had that was good; it went far beyond what (the students) should know. But I didn't like the CTBS because it didn't start at a low enough level; it was too hard.
>
> So you need (a test) that starts at a very minimal level and goes up beyond what (students') capabilities are, so you really get a true picture of what the potential is of the best and of the slowest.

A second-grade teacher similarly criticized the CTBS and the Skills Survey. The CTBS, she felt, is:

> too hard for most (students). They are frustrated. The Skills Survey is silly. It is costly and doesn't give a true picture.

One of the fourth-grade teachers agreed in somewhat stronger terms:

> The Skills Survey is not timed. All but three students finished. One girl got them all wrong. All she did was mark it; she wasn't even trying. It's the same when we give the CTBS. (A certain student) got the highest score, and he couldn't read. He is now in EH. I know he can't do it. He guessed.

This kind of problem was also recognized by the Cityside principal, who is concerned about the CBSS because it has "no norming data (and has) low-level expectancy." Further, because the CTBS is no longer required school-wide, and because the principal sees some value in generating school-wide norm-referenced data, "that's why I spend $1200.00 for the MAT."

While some teachers at Cityside do see a need for minimum competency and group-administered, norm-referenced tests, they are not particularly pleased with the tests currently being used for these purposes. Comments amplifying their frustration appear below.

Technical problems: Three Cityside teachers and the Chapter I coordinator commented on this issue. One of the second-grade teachers criticized the CTBS Español because "some of the words don't translate into Spanish...(and) the print is too small...(the test) is not testing Spanish skills." A fifth-grade teacher noted similar problems with the English-language version of this test:

> (The test) vocabulary is a problem for (the students). Some of the explanations are (written in language) for

> adults. The test is a contradiction. (It) makes criminals of us all. It's unrealistic. It makes us all cheat.

Discussing another kind of technical problem, that of score reporting format, this same fifth-grade teacher observed that:

> There has to be a better way of reporting the scores to the teachers so they can be used...I would like to get a print out on a sheet at the beginning of the year which shows all the Skills Survey and CTBS results...so I can see it all together at a glance. To have to go to everyone's cumulative file is very tedious;...someone in the school, whether coordinator, principal, or whoever is in charge, should get it all together.

That no one in Cityside, "gets it all together" was corroborated by the vice principal. Describing what was a frustrating experience for him as an administrator and for his teachers as well, he commented that:

> Some teachers want to know how students did, because the printouts aren't going to come back until school is out. If they want to know, we have a hand-scoring key if they want to do this. No one interprets school-wide.

The fifth-grade teacher above who cited the concern with the CTBS pointed out another problem with some of the tests administered at Cityside. Teachers are very concerned because they need much more information on what the various tests mean, their "validity and correlation with other tests." Another fifth-grade teacher commented that "testing is not as controlled as it was twenty-five years ago. We would have inservice training to make sure you knew what you were doing." This was also a concern for the vice principal, who commented that teachers at Cityside, in general, need more explanation from the Metro District's research and evaluation office about what the various test scores mean.

Tests viewed favorably: Four teachers at Cityside spoke of the kinds of tests that are viewed more favorably. The bilingual coordinator, for instance, discussing a Spanish reading test she developed herself, noted that this kind of testing

> is not time-consuming. It is something I can get feed-back on immediately. It isn't disruptive; it's a very satisfactory, necessary instrument.

In terms of information on students' reading ability, one of the first-grade teachers described the diagnostic value of the San Diego Quick Assessment as follows:

> I give the San Diego (Quick Assessment), which takes about thirty seconds per child (and) it's pretty accurate...one of the most accurate I've ever seen. It's something I do at the beginning of the year. You can do the whole class in fifteen or twenty minutes.

One of her colleagues strongly agreed. "I don't mind giving (the San Diego) because it doesn't take much time and it's useful." A fifth-grade teacher concurred that the San Diego Quick Assessment "is useful when you want to place a new student."

Many teachers at Cityside viewed the unit posttests of the Developmental Reading Program positively. Some teachers also saw the value of the information they felt they could obtain from a good minimum competency test or a good norm-referenced test, though they were concerned about problems with the two tests actually used in the school for these purposes -- the CBSS and the CTBS.

Effects of Testing

Most of the Cityside teachers commented on problems arising from the effects of testing on students or teachers. At Hillview, nine of the eleven teachers interviewed spoke about the effect that testing

has in fostering student anxiety. Half the Hillview interviewees also expressed concerns with pressures that testing can generate for teachers.

Student anxiety: A majority of the teachers at Cityside were concerned about tests causing students either to become very wound up and/or to become tired and enervated. In this regard, some of the teachers described efforts to provide student "wind-down" time after a testing period by scheduling the test immediately before recess. When this was not possible, they said they generally gave their classes about fifteen minutes (taken out of instructional time) to relax and get over the effects of testing.

About a half-dozen Cityside teachers cited testing as a generally frustrating experience for their students. One first-grade teacher specifically refered to the MAT as "too tiring and frustrating," a view for which she found evidence in students "breaking their pencils" to try to avoid taking the test. One of the third-grade teachers mentioned that her "third-grade students get too many tests, often several at about the same time." This teacher saw her students becoming restless as the Spring testing period wore on; "testing time and its effects take a long time to wear off," she said.

One of the second-grade teachers described certain kinds of tests and their effects on her students as follows:

> The ongoing tests like the District Reading Program...aren't identified as tests by a lot of students. Those that use special pencils (and) answer sheets...are stressful; standardized tests are stressful. In (the lower grades) the students use the restroom during the test even though I take them before. To some kids, they get anxious not being able to sit through it. All of us feel 'tight' after the testing and try to make it an easier, less stressful activity.

Another second-grade teacher agreed. Her students, at a testing period, "cry, sigh, tap feet...(and) show relief when it's over." And one of the fifth-grade teachers was even more forceful in her description of negative test effects:

> The CTBS makes students act high for the rest of the day. Behavior is terrible afterwards. Even on local tests they will act up...They are louder, more uncontrollable, (they) fight sometimes in the play ground (and find it) hard to sit still in a lot of situations if (the test) is too hard for them, like most tests are.

Another fifth-grade teacher agreed, though less vociferously, by describing her test-taking students as "drumming on the desk with pencils, fidgeting, and causing minor disturbances."

Other Cityside teachers indicated that they were less concerned about testing's effects on themselves and their students. These teachers believed that the more positive approach they took to testing made a difference. For example, one of the kindergarten teachers described the situation in these terms:

> Testing is a tool for me and not viewed as a burden. I just keep recycling. Tests that I give don't bother (the students) at all because I enjoy giving them and they're fun. I make (the students) absolutely aware that we're trying to find out something and that I need some information. I don't allow the students to get uptight.

This approach to tests and testing was alluded to by several other teachers at Cityside. For example, a sixth-grade teacher mentioned that "test preparation is fundamental with our children."

That teacher attitude toward tests and testing varied within Cityside, and that this teacher attitude may have a bearing on the amount of stress felt by the students, was corroborated by the school's Chapter I Coordinator. According to this administrator, some

teachers don't understand what a test is for or what the scores mean. Therefore:

> they'd complain and some wouldn't put forth the effort to make sure (they understand the test purpose). They'd give (the test) to the children and tell them to do the best they could.

He then went on to describe the ideal situation and practice which some of the teachers at Cityside try to follow. That is:

> ...to prepare (students) with the (testing) mechanics; not the test, but the mechanics

so that students understand how to take the test. This, the coordinator said, can lead to improved student attitude and higher expectations for themselves.

At Hillview, all of the teachers referred in some manner to the cost that test-anxiety incurs for students. Taken jointly, these teacher comments suggested that testing does not impose a uniformly high psychological stress for all students at Hillview. Nevertheless, comments reveal, some students do occasionally become over-anxious. For example, as explained by the kindergarten teacher at Hillview, "some kids feel pressured in the beginning (but) most kids are okay by May."

However, a first-grade teacher explained that:

> This is a highly competitive group of children. They know what group everyone's in and who's high and who's low--and we never mention it. And when a mastery test is given and we can't let some children go on to the next group, it's devastating to them.

Comments by other Hillview teachers, especially in the upper grades, suggest that test anxiety does not apply to all students. Their remarks indicate that anxiety which does occur is usually

manifested during curriculum or placement tests, which affect student standing in the classroom or placement in a subsequent grade or school. Less anxiety, in these teachers' view, appears during standardized tests which are not used for placement or promotion purposes at Hillview.

Pressure on Hillview students is also increased to some extent, staff members believed, because of parental influence. As a fifth-grade teacher put it:

> There's considerable parent pressure, particularly among Asian parents--a drive for students to get ahead. Parents will drop in and check how their child is doing. They will sign their children up for all different kinds of lessons. In many cases the children don't play with others.

Beyond the question of the anxiety instilled in students because of test or test-related pressures, the teachers at Cityside (but not at Hillview) made comments on other more positive effects of testing.

Student motivation: Three or four Cityside teachers cited testing as a reinforcer or motivator. According to a first-grade teacher:

> Testing is anxiety; that's a built in. That's part of life because you're being tested all the time. Actually that's probably good for (the students)...Once you overcome it and do it, next time you may be anxious but you know you can do it.

The sixth-grade teacher who had commented that test preparation is, or should be, fundamental at Cityside, agreed:

> I feel comfortable about tests. Kids need a certain amount of anxiety. There are no particular tests that cause my students anxiety.

This teacher then described her students' enjoyment and motivation from some kinds of tests:

> They get their (teacher-made spelling tests) back the same day. They love that. They always want to see how they did. They'll come to the aide or me and ask: 'Did you score the papers? Are they ready, yet?'

Obstacles to motivation: Even Cityside teachers who would like to use tests as instructional motivators, however, found that there were obstacles to doing so. Describing the MAT, for instance, one of the fourth-grade teachers was disturbed that "students come out particularly low." Further, for formal tests in general, teachers may not agree with the accuracy of the results, because:

> Many times (the students) don't do well on paper-and-pencil tests. A lot is a guess. If they don't look, they make a mistake...Students may not be motivated. Most of the class has lots of family problems, and other things make it difficult for them. (This leads to) two extremes of (of test behavior); 'I can't do it' or 'I won't do it.' Then they give up.

The problem of students "giving up" was reiterated by the Chapter I coordinator in terms that hark back to an earlier concern with test validity in general. That is:

> There are things in the CTBS that (some) children never come in contact with (and so) it's a waste of time. I think it's better if (the test) includes most of the things they come in contact with. And I think they are frustrated. They don't know the answers.

On the other hand, some teachers believe a student can get a false sense of accomplishment on the basis of scores on tests like the CBSS. Because the ceiling on this test is so low, remarked one of the second-grade teachers, the student "can have a good score and know nothing." The Chapter I Coordinator agreed: "(the Skills Survey) only has the minimum. Children can't be challenged if your expectations are the minimum."

The failure, or in some cases, inability, to use tests as instructional motivators was aptly decribed by the bilingual coordinator. According to this specialist, some students viewed the CTBS as a

> pass or fail situation, and therefore take that quite seriously. This is too bad. Student motivation is wasted because the test is used only for external (reporting) requirements.

Pressure growing from public reporting of scores: The four teachers at Hillview commenting on this issue suggested that they are concerned that school administrators and the public believe that state- and district-mandated tests reflect teachers' competence. As a fourth-grade teacher put it: "Handing in test results to the principal adds pressure." As explained by a fifth-grade colleague, "turning in test scores exerts a psychological pressure on the teacher because each spring the principal posts the standardized test scores by classroom," and "I think there's some pressure on teachers as a result of that." Further, according to this teacher, the principal had been stressing that "he wants to know why" there has been a decline in primary-grade test scores, "and I think this creates some (teacher) anxiety."

How this kind of teacher anxiety in Hillview can grow was explained by a first-grade teacher:

> I think that any time a test is given, a national type test, you don't lose sleep over it or anything, but you're concerned because it is your children being tested. Therefore it's what you have taught them and it is published and it is reflected back onto you if the students are below where they should be.

A fifth-grade colleague agreed:

> ...I would say there's a certain amount of pressure, not on the weekly or unit tests, but (on the) mandated tests at the end of the year...What our principal does is post a list of how the various classes have done. He makes it anonymous but we can figure it out ... it would be very upsetting knowing that it's not always the teaching that produces that ... kind of score (a low growth score) ... and sometimes you look at that kind of list and you know that other people are saying 'here's the good teacher and here's the bad teacher.' It's ludicous. I don't like that kind of comparison.

Loss of Instructional Time

While only one or two teachers at Cityside explicitly stated a concern with test intrusion on instructional time, about half of the teachers at Hillview expressed this concern. As a first-grade teacher put it, "testing cuts in on instructional time; for example, students don't get reading instruction for two weeks." Her team-teaching colleague agreed that "tests add more work" and "cut instructional time."

Many teachers also indicated that some tests create behavior problems with students; hence (as described above) teachers routinely give over at least fifteen minutes of potential instructional time to allow students to wind down before resuming teaching-learning activities.

Summary

Teachers' commentary on psychological and other costs associated with testing generally reflected concerns with test utility or usefulness, the appropriateness of tests for students and/or the appropriateness of how their results are used, the effects of testing, and loss of instructional time caused by testing.

While these concerns were evident to some degree in both schools, the pattern of responses and emphasis varied. The Cityside teachers were annoyed and somewhat frustrated with the imposition of tests that have limited utility and/or are of questionable worth and suitability. However, while they were a bit concerned about the anxiety that tests may cause students, tests were not viewed as a serious source of personal stress. Testing, in other words, may entail noteworthy opportunity costs in terms of time spent in useless or invalid pursuits, but significant psychological costs do not seem to accrue.

In contrast, teachers at Hillview were more vocal about direct psychological costs of testing. All noted test-related anxiety in their students, and over half felt personally (albeit minimally) stressed and pressured by testing. These anxieties may result because test scores have both credibility and utility at Hillview -- within an accountability context -- for everyone in the setting. They carry personal consequences for both students and teachers.

PSYCHOLOGICAL COSTS OF TESTING: STUDENT ATTITUDES

Introduction

This section of the report discusses the results of our interviews with students exploring the psychological costs of testing from their standpoint. One of the reasons behind this aspect of our work is that relatively little is known about students' attitudes and feelings toward assessment. In a 1979 study, Stetz and Beck found that at kindergarten through fourth grades, a majority of students felt somewhat positively toward tests, although 56 percent indicated that they were nervous about taking them. In grades five through twelve, however, only 26 percent of the students felt positively about tests, while 27 percent reported feeling negatively about them. In addition, 30 percent reported getting nervous before taking tests made by the teacher.

Kirkland's (1971) study investigated whether test scores affect student's self-concept. Kirkland found that the effect of receiving information about one's abilities depends on a variety of factors, including the legitimacy of the information source, the perceived accuracy of the test, the degree to which the information confirms one's own self-estimate, and the extent to which it is threatening or rewarding. Test scores have potentially great impact where an individual's self-concept is at considerable variance with the record of performance on the test, where rationalizations of poor performance are unavailable, or where the test score is substantially higher than one's own estimate. Under such conditions, we can expect disagreement

to affect the individual's aspiration level, motivation to achieve, and personal decisions about the future.

However, Kirkland (1971) also found that test scores are of relatively minor importance in shaping one's self-estimate of ability in comparison with school grades, comments made by peers and parents, and relationship with his/her teachers.

In a study by Sharp (1966) of 25 elementary and secondary teachers in Florida, there was an evenly mixed reaction to the question of whether emphasis on testing caused competitiveness in the classroom.

In light of these few and certainly non-definitive findings, the student interviews we conducted explored the effect that different forms of assessment have on students. Do students find testing a positive or negative experience? How worrisome do they find more and less formal means of assessment? How does the assessment experience seem to influence their feelings about their own intelligence, and how others view them? How does the assessment experience affect students' views about "what's important" in their academic career?

Student Interviews

We drew a random sample of 60 students from our two case-study schools, Hillview and Cityside. Twenty students were selected from the fourth, fifth, and sixth grades at each school (10 each grade from the two schools). The overall ethnic composition of the group (using categories applied by the schools) was as follows: 26 Black; 13 White/Anglo; 6 Hispanic; 14 Asian; and 1 Pacific Islander. Thirty-seven were males and twenty- three were females.

An interview schedule was developed in a game-like format involving three tasks. The first activity consisted of a sorting task. Here we asked the student to sort 10 common school activities, including six achievement-assessment activities, into three piles: "Activities I like": "Activities I dislike": and "Activities in the middle/no opinion". After this initial sort, we asked the subject to rank the activities in the "like" and "dislike" piles, putting the most liked (or most disliked) activity on top, followed by the next most liked (disliked), and so forth.

The second task involved a semantic differential exercise with four pairs of descriptors (fun/not fun; important/unimportant; smart/dumb; calm/worried) on a 7 point scale. We asked the student to place each of the ten school activities first used in task one along the 7 point scale on each of the four descriptor dimensions.

In the third task, we asked students to estimate which of five school assessment activities their parents, teachers, students themselves, and their classmates thought that it was "most important to do well on." The assessment activities were: homework; teacher's questions in class; standardized tests; chapter tests; teacher made tests.

The interview was administered individually in a quiet place away from the classroom. It embedded various forms of assessment (standardized tests; chapter tests; teacher-made quizzes; homework, answering teachers' classroom questions; and story writing) among other forms of school activities: physical education games; assemblies; nutrition or snack time; talking with friends. We wished simply to see whether

students did differentiate assessment from non-assessment activities, as well as to see if students differentiated among various forms of assessment. Further, the interview method we adopted allowed us to measure student attitudes toward testing and other school activities in three different ways. This not only provided a measure of the instruments' inherent construct validity, but also measured consistency of students' opinions across different elicitation contexts.

In this section of the report we will first discuss student ratings on the importance of testing activities on tasks two and three (semantic differential and important-to-do-well-on). These findings indicate the importance to students of different types of testing, their feelings about testing as compared to non-testing activities, and the relationship between assessment activities and significant others in the eyes of the student. Second we will present students' global affective responses to different types of assessment activities based on the task one, like/dislike activity. Third, we will provide a more differentiated look at student feelings about assessment compared with other school activities.

Students' Views of the Relative Importance of Different Types of Assessment

Six commonly used forms of student assessment were included in all three tasks on the instrument. These were chapter tests, standardized tests, teacher made quizzes, homework, writing a story, and answering teacher's questions in class. The first three assessment types are more formal, less frequent, and more clearly "marked" as instances of assessment. The other three usually occur more frequently

as part of the regular school routine and/or as more or less formal ways of evaluating students' achievement. In addition to the six assessment modes, four other school activities were included in two of the tasks on the measure. These were recess, talking to friends, p.e./games, and assemblies.

Table 46 shows that students regard assessment activities as more important than non-assessment activities. Clearly, standardized tests and chapter tests were rated (7 point scale) as the most important activities. Assemblies (a non-assessment activity) were viewed as slightly more important than writing a story, which many teachers use to assess language arts skills. (Students may associate assemblies with instruction; assemblies in the two study schools are often used to convey information about school rules and regulations and to show educational films.)

Student ratings on the "important to do well on" task generally supported these findings.

Table 46

Overall Sample: Ordered Mean Ratings for 10 School Activities
Important/ Unimportant (n = 60)

Standard-ized Test	Chapter Test	Home-work	Answering Teacher's Questions	Teacher Quiz	Assemblies	Writing A Story	P.E. Games	Recess/ Nutrition	Talking With Friends
6.63	6.15	6.08	5.80	5.68	5.43	5.33	5.28	4.71	4.41

Table 47 shows students' responses to the matter of which assessment activities are most important from the standpoints of their teachers, parents, themselves, and classroom peers.

Table 47

Overall Sample: Frequency of Ratings on "Most Important to Do Well On" Task (n = 60)

	Homework	Answer Teacher's Questions	Standardized Test	Chapter Test	Teacher Made Quiz
My Teacher	20%	5%	52%	17%	5%
My Folks	40%	7%	33%	10%	8%
Me	17%	12%	43%	20%	7%
Kids in My Class	13%	18%	22%	22%	22%

Over half the students (52%) responded that teachers feel it is most important for students to do well on standardized tests. About 43% of the students also chose the standardized test as the assessment type that they themselves believed it was most important to do well on. The sample was closely divided with regard to parental views: 40% said parents would rate homework as the most important and 33% indicated that standardized tests would be the parents' choice.

Although students in both schools gave standardized tests a similarly high rating across all "significant others," Table 48 shows that there were some differences with respect to other activities. Cityside students indicated that they and their teachers would consider homework to be the next most important activity. Hillview

Table 48

Frequency of Rating for "Most Important to Do Well On" Task by School
[Cityside, n = 30; Hillview, n = 30]

	Homework		Answering Teacher's Questions		Standardized Test		Chapter Test		Teacher Made Quiz	
	City-side	Hill-view	City-side	Hill-view	City-side	Hill-view	City-side	Hill-view	City-side	Hill-view
My Teacher	8	4	2	1	16	15	2	8	1	2
My Folks	12	12	2	2	10	10	2	4	3	2
Me	7	3	5	2	12	14	5	7	--	4
Kids in My Class	3	5	7	4	7	6	6	7	5	8

Table 49

Mean Rating for Assessment Activities by School: Important/Unimportant
[Cityside, n = 30; Hillview, n = 30]

	Standard-ized Test	Home-work	Chapter Test	Answering Teacher's Questions	Teacher Made Quiz	Writing A Story
Cityside	6.73	6.43	6.23	6.03	5.86	5.86

	Standard-ized Test	Chapter Test	Home-work	Answering Teacher's Questions	Teacher Made Quiz	Writing A Story
Hillview	6.53	6.06	5.73	5.56	5.50	4.80

students, on the other hand, rated chapter tests as the next most important. This pattern is also repeated in Table 49, which shows between-school differences in student ranking of assessment activities. Note also that Hillview students rated writing a story as much less important than did students at Cityside.

Table 50 displays students' mean ratings on the "importance" semantic scale by grade level. Across all three grades, students rated standardized tests as the most important activity. Students continue to view chapter tests and homework as among the important forms of assessment, but the priority placed on each differs across grade level.

Further, mean ratings for all six assessment forms tend to decrease in the upper elementary grades. The small sample size (n = 20 per grade level) and degree of these differences, however, suggest care in interpretation. Perhaps the differences reflect that students find the assessment experience--whatever its form--more routine and less awe-inspiring as they continue through school.

Table 50

Mean Rating for Assessment Activities by Grade: Important/Unimportant

[Grade 4, n = 20; Grade 5, n = 20; Grade 6, n = 20]

	Home-work	Writing A Story	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz
Grade 4	6.30	5.60	6.65	6.05	6.50	6.15
Grade 5	6.20	5.30	6.80	5.70	6.15	5.75
Grade 6	5.75	5.10	6.45	5.65	5.80	5.15

In summary, the sixty students we interviewed rated all six assessment modes on the "important" side of the semantic scale. On the whole, they also saw the two more formal and (usually) more comprehensive modes--standardized tests and chapter tests--as more important than the others. Homework (which many students believed their parents emphasized) also received a comparatively high importance rating. Routine oral evaluation (answering classroom questions) and quizzes followed in close succession. Students' mean ratings of importance therefore, seem in a general way to reflect the following principle: measures that occur less frequently and "cover" more content tend to be more important.

By the upper elementary grades, further, pupils can and do seem to differentiate among the relative importance of different forms of assessment. Broadly speaking, their views seem consonant with actual practice. Each instance of a standardized test or a chapter test usually has the potential of making more difference in students' educational careers than each instance of a quiz, homework, or oral classroom performance.

Students' General Affective Demeanor Toward Different Forms of Assessment

The sorting task described previously investigated whether students' general feelings vary with different types of assessment techniques. In this task, students sorted the same ten activities discussed above into three piles: "things I like," things I dislike," and "things in the middle." They then rank ordered the activities placed in the "like" and dislike" piles.

As might be expected, Table 51 shows that students consistently preferred the non academic to the assessment activities. The next most liked activities, overall, were the more routine, less marked forms of assessment. Direct testing activities were less frequently mentioned as liked. Conversely, the most disliked activities were usually the direct forms of testing, followed by indirect assessment activities and social school activities. A significant percentage of the sixty students took a "neutral" position and placed various modes "in the middle."

Table 51

Percentage of Students Who Labeled Each School Activity as "Like", "In the Middle", or "Dislike": Total for Both Schools

	LIKE	IN THE MIDDLE	DISLIKE	TOTAL
Standardized Tests	32%	27%	41%	100
Chapter Tests	17%	40%	43%	100
Teacher Made Quiz	38%	42%	20%	100
Homework	32%	38%	30%	100
Writing a Story	57%	23%	20%	100
Answering Questions	45%	38%	17%	100
Assemblies	53%	38%	9%	100
P.E.	87%	5%	8%	100
Recess	82%	15%	3%	100
Talking with Friends	93%	2%	5%	100

Three observations emerge from our data: First, the types of assessment that students on the whole like less often and dislike more often are those that they also rated as more important. These are assessments that tend to be less frequently administered and more comprehensive in content (standardized and chapter tests), along with homework (which makes a regular claim on children's out-of-school time). Second, a majority of the students viewed even these performance modes positively or neutrally. And only small proportions of students disliked quizzes and answering teacher's questions, while more than half said they enjoyed writing a story. Third, the students who disliked the less frequent, more formal and comprehensive forms of testing constituted a substantial minority.

In Table 52, certain differences in student's attitudes are evident between schools. The most notable of these lies in students' preferences toward standardized tests: 53% of the students at Cityside said they liked standardized tests as opposed to only 10% of the students at Hillview. At the same time, 53% of the students at Hillview said they disliked these tests, compared to 30% at Cityside. The same pattern holds for chapter tests. At Hillview, further, the frequency of like responses is generally lower for each academic assessment activity; Hillview students tend to be more affectively neutral on most.

In general (and especially at Hillview) the more formal and comprehensive tests--standardized and chapter--were viewed most negatively. But only two-fifths of the interviewees found these unappealing, and a majority of responses to each assessment mode were positive and neutral.

On the whole, students at both schools did offer differentiated responses on the sorting task. This is especially evident when their reactions to the academic school activities are compared to their reactions toward the non-academic ones.

TABLE 52

Percentage of Students Who Labeled Each School Activity as "Liked", "In the Middle", or "Disliked": Total by Schools

	CITYSIDE			HILLVIEW		
	LIKE	MIDDLE	DISLIKE	LIKE	MIDDLE	DISLIKE
Standardized Tests	53%	17%	30%	10%	37%	53%
Chapter Tests	30	33	37	3	47	50
Teacher Made Quizzes	50	30	20	27	53	20
Homework	50	20	30	13	57	30
Writing a Story	60	7	33	53	40	7
Answering Teacher's Questions	60	23	17	30	53	17
Assemblies	43	44	13	64	33	3
P.E.	90	7	3	86	3	13
Recess	83	10	7	80	20	-
Talking with Friends	90	3	7	97	-	3

A Finer-Grained View of Students' Feelings About Testing

In the semantic differential task previously described, we asked students to place each of the six assessment and four non-academic activities on the dimensional scales of fun/not fun; calm/worried; smart/dumb; and important/unimportant. The important/unimportant

dimension has already been discussed. The three remaining dimensions combine to offer a more finely-grained view of students' feelings about testing.

1. Students' Experience of Different Assessment Forms as Fun or Not Fun

The fun/not fun scale probably taps an affective dimension similar to the like-neutral-dislike sorting task.* It goes beyond that task, however, in revealing the magnitude of individual students' general feelings about the different assessment modes.

As Table 53 shows, non-academic activities received higher mean ratings (7 point scale) than the assessment activities. Once again, standardized tests, homework, and chapter tests were the most negatively rated.

Table 53

Overall Sample: Mean Ratings for 10 School Activities
Fun/Not Fun (n = 60)

Standardized Test	Homework	Chapter Test	Answering Teacher's Questions	Teacher Made Quiz	Assemblies	Writing A Story	P.E. Games	Talking With Friends	Recess/ Nutrition
3.50	4.06	4.08	4.88	4.96	5.00	5.16	6.30	6.31	6.43

* A cross tabulation shows that, overall, individual students' responses on the sorting task were consonant with their ratings for the same items on the fun/not fun scale for 79% of the interviewees. A consonant response is defined broadly here as (1) a "like" placement on the sorting task with a rating of 7,6, or 5 on the seven-point fun/not fun scale; or (2) an "in the middle" placement with a 5,4, or 3 rating; or (3) a "dislike" placement with a 3, 2, or 1 rating. This definition slightly broadens the "middle" range of the semantic differential scale, which is of course constituted only by the rating "4".

However, Table 54 below, which describes the frequency of ratings for the six assessment items, shows that the sample was almost evenly divided on their ratings for some of the testing items.

Table 54

Overall Sample: Frequency of Ratings for 6 Assessment Activities Fun/Not Fun (n = 60)

	Fun						Not Fun
	7	6	5	4	3	2	1
Homework	20%	7%	20%	15%	10%	8%	20%
Writing a Story	37%	17%	17%	8%	7%	8%	7%
Standardized Test	15%	10%	8%	15%	15%	7%	30%
Answering Teacher's Questions	22%	18%	13%	32%	7%	5%	3%
Chapter Test	15%	13%	15%	17%	17%	8%	15%
Teacher-Made Quiz	30%	15%	15%	20%	7%	8%	5%

Only one activity, standardized tests, was negatively ranked (rating less than 4) by 50% or more of the students. Although chapter tests and homework were negatively rated by 38 to 40% of the students, they received positive (5 or higher) ratings by 43 to 47% of the students. These items also received distinctly higher percentages of ratings of "1," at the extreme negative end of the scale. Other assessment activities received more positive than negative ratings. Writing a story was rated fun (5-7) by 71%; teacher-made quizzes by 60%; and answering teacher's questions in class by 53%.

The between school comparison of ratings seen below in Table 55 confirms patterns already described. That is, standardized tests, homework, and chapter tests are the most negatively rated activities by students in both schools. A significant means difference was found only for the teacher-made quiz, where Hillview students assigned a more negative rating ($p \leq .01$).

Table 55
Mean Ratings for 6 Assessment Activities by School
Fun/Not Fun

	Standard-ized Test	Chapter Test	Home-work	Writing A Story	Teacher Made Quiz	Answering Teacher's Questions
Cityside	4.06	4.33	4.53	5.53	5.66	5.23

	Standard-ized Test	Home-work	Chapter Test	Teacher Made Quiz	Answering Teacher's Questions	Writing A Story
Hillview	3.03	3.60	3.83	4.26	4.53	4.80

Similar findings were found in grade level comparisons. As Table 56 indicates, homework and standardized tests usually received negative ratings whereas writing a story, answering teacher's questions, and doing teacher-made quizzes receive positive or neutral ratings (values right at 4).

Table 56

Mean Rating of 6 Assessment Activities at Three Grade Levels: Fun/Not Fun
[Grade 4, n = 20; Grade 5, n = 20; Grade 6, n = 20]

	Home-work	Writing A Story	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz
Grade 4	4.85	5.40	3.20	5.10	4.50	5.35
Grade 5	3.85	4.95	4.20	4.85	3.75	4.70
Grade 6	3.50	5.15	3.25	4.70	4.00	4.85

In summary, a majority of the students interviewed found three less-formal, more-routine forms of assessment to be fun. And the sample's mean responses confirm that for most pupils standardized tests, chapter tests, and homework are the least appealing forms of assessment. Finally, roughly a quarter to a third of the students interviewed experience these activities more-or-less aversely: about this proportion rates each of these activities with either a "1" or "2," at the negative end of the fun/not fun scale.

2. Students' Views of Different Forms of Assessment as Worrisome

The mean ratings for the overall sample (Table 57) shows that students feel calm in all non assessment items and in one assessment item, writing a story. Their ratings (7 point scale) of other assessment items were neutral.

Table 57

Overall Sample -- Mean Rating for 10 School Activities
Calm/Worried (n = 60)

Standard-ized Test	Home-work	Answering Teacher's Questions	Chapter Test	Teacher-Made Quiz	Assemblies	Writing A Story	P.E. Games	Recess/ Nutrition	Talking With Friends
4.08	4.33	4.63	4.46	4.71	5.00	5.33	5.85	5.95	6.10

However, when we look at the frequency of ratings for the six assessment activities in Table 58, we find that a small though significant proportion of students, 26 to 38%, worry about some forms of assessment: standardized tests (38%); homework (34%); chapter tests (27%); and answering teacher's questions (26%). The greater proportion of students feel calm across all activities, particularly in writing a story (68%), taking a teacher-made quiz (59%), doing a chapter test (51%), and answering teacher's questions (50%).

Between-school ratings (Table 59) show only that students in both schools rated themselves as calm in writing a story. The only school-to-school difference was that Hillview students, unlike Cityside's, gave homework a negative (worry) rating. All other ratings were neutral.

A display of mean responses on the calm/worried scale shows no general trends. Viewed in juxtaposition with Table 60, however, one minor point emerges. While students' mean ratings of the importance of all assessment forms declines across grade levels, there is no accompanying decline in how much worry students associate with them.

Table 58

Overall Sample: Frequency of Ratings for 6 Assessment Activities
Calm/Worried (n = 60)

	Calm						Worried
	7	6	5	4	3	2	1
Homework	17%	10%	17%	23%	20%	7%	7%
Writing a Story	33%	23%	12%	17%	7%	7%	2%
Standardized Test	15%	17%	7%	23%	13%	12%	13%
Answering Teacher's Questions	20%	12%	18%	23%	18%	5%	3%
Chapter Test	22%	17%	12%	23%	7%	3%	17%
Teacher-Made Quiz	17%	22%	20%	18%	13%	2%	8%

Table 59
Mean Ratings for 6 Assessment Activities by School: Calm/Worried
[School 1, n = 30; School 2, n = 30]

	Standard-ized Test	Chapter Test	Teacher Made Quiz	Home-work	Answering Teacher's Questions	Writing A Story
Cityside	4.13	4.43	4.60	4.76	4.96	5.56

	Home-work	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz	Writing A Story
Hillview	3.90	4.03	4.30	4.50	4.83	5.10

Table 60

Mean Rating of 6 Assessment Activities at Three Grade Levels: Calm/Worried
[Grade 4, n = 19; Grade 5, n = 20; Grade 6, n = 20]

	Home-work	Writing A Story	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz
Grade 4	4.35	5.35	3.85	4.45	4.70	4.65
Grade 5	4.55	5.40	4.90	5.00	4.30	4.60
Grade 6	4.10	5.25	3.50	4.45	4.40	4.90

3. Students' Association of Forms of Assessment with Their Intellectual Self-Esteem

Assessment activities provide occasions for students to do well or poorly, to succeed or fail. Presumably, then, they can influence students' perceptions of their own intellectual competence. What kind of influence assessment has probably depends upon how well students perform when assessed. Nevertheless, it seemed worthwhile to explore the extent to which students associated generic forms of assessment with feelings of intellectual capability or incapability. The smart/dumb semantic scale examined this issue in a general way.

Overall, students did not differentiate the six assessment activities along the smart/dumb semantic scale. As Table 61 illustrates, the testing activities received ratings ranging from a low of 5.36 to a high of 5.65 for the total sample (n = 60). These differences are significant neither intuitively nor statistically.

Table 61

Overall Sample: Ranked Mean Ratings for 6 School Assessment Activities Smart/Dumb (n = 60)

Standard-ized Test	Writing A Story	Teacher Made Quiz	Answering Teacher's Questions	Chapter Test	Home-work
5.36	5.55	5.55	5.60	5.65	5.70

The overall frequency of ratings for the six assessment items (Table 62) shows that 68 to 83 percent of the responses were within the 7 to 5 range (smart") for all items; 12 to 23 percent were in the exact middle of the scale; and only 2 to 8 percent on the negative ("dumb") side of the scale. (Also see mean ratings for each schools' students in Table 63.)

These findings may reflect students' reluctance to admit feeling "dumb," especially to a stranger. It may be, also, that the structure of this question was confusing. Students may not have been able to associate a general view of themselves as feeling "smart" or "dumb" with a generic assessment activity. However, pilot interviews employing this same item "worked" to elicit a substantially wider range of responses. It may simply be, then, that students at Hillview and Cityside -- whatever their individual performance -- rarely felt very "dumb" in the mere presence of assessment activities.

Ethnographic work in the two schools suggests that teachers believe strongly that their students are capable. They appear to routinely communicate this belief to the children. Hillview is often spoken of in Littleton District as the school with the highest

achievers. Cityside was recently cited as outstanding among the Metro District schools with compensatory education programs. Word of their schools' relative standings probably makes its way to students. And within each setting, most students progress through their subjects with rates of achievement that permit them to feel competent. Few are likely to receive consistent evidence that they are incapable academically. Their responses on the "smart/dumb" scale may very well reflect this.

Table 62

Overall Sample: Frequency of Ratings for Six Assessment Activities
Smart/Dumb (n = 60)

	Smart						Dumb
	7	6	5	4	3	2	1
Homework	40%	22%	13%	20%	3%	2%	--
Writing a Story	38%	18%	12%	23%	8%	--	--
Standardized Test	37%	15%	20%	17%	3%	5%	3%
Answering Teacher's Questions	37%	22%	17%	18%	3%	3%	--
Chapter Test	33%	20%	28%	12%	2%	3%	--
Teacher-Made Quiz	25%	32%	22%	18%	2%	2%	--

Table 63

Mean Ratings for 6 Assessment Activities by School: Smart/Dumb
[Hillview, n = 30; Cityside, n = 30]

	Teacher-Made Quiz	Writing A Story	Standardized Test	Chapter Test	Answering Teacher's Questions	Homework
Cityside	5.76	5.93	6.00	6.00	5.93	6.36

	Standardized Test	Homework	Writing A Story	Answering Teacher's Questions	Chapter Test	Teacher-Made Quiz
Hillview	4.73	5.03	5.16	5.26	5.30	5.33

Summary

The data show that students distinguish assessment from non assessment activities across all tasks, and within assessment items on some. Students rated standardized tests as the most important and worrisome activity as well as among the least liked and least fun. Chapter tests and homework competed for second place as the most important, least liked and least fun activity. Their second place rating varied according to whether responses were examined for the total sample, by school, or across grade levels. Teacher made quizzes and answering teacher's questions in class also vied for third place in importance. However, students usually rated them likeable and fun activities. The most popular assessment activity was writing a story. It was given the highest fun and like ratings of the six assessment activities. It was also rated to be the least important one.

The general between-school pattern is that Cityside students gave slightly to moderately higher (positive) ratings than Hillview students did on the "like/dislike" tasks and "fun/not fun" scale.

Across-grade-level variations showed a slight trend: attitudes toward standardized testing, chapter tests, and homework seemed to be more negative in higher grade levels. These activities were experienced as less liked, less fun, and more worrisome by the sixth graders than by the fourth graders. Further, these as well as other assessment activities were viewed as less important from the fourth to the sixth grade.

Student ratings on the dimensions of affect (fun/not fun, calm/worried, smart/dumb) support their teachers' comments on the psychological costs of testing. Teachers indicated that although the majority of their students did not find most assessment activities to be a particularly worrisome or negative experience, a minority of students did manifest some anxiety. Most students indicated that they felt calm and smart during all testing activities even though they did not rate them as fun activities. This includes those activities rated as very important. However, about one third or more of the students (38 to 40%) expressed feelings of anxiety or distaste for standardized tests and chapter tests.

Because of the small sample size (n = 60) and the paucity of research in this topic, these findings suggests potential avenues for research as much as they provide information. For example, Cityside students had generally more positive attitudes toward testing than did Hillview students. Recall that Cityside is an inner city moderate to

low income school. This finding contradicts the stereotypical notion that inner city students are less self-confident and receptive toward testing than their middle class fellow students in the suburbs, such as Hillview.

Students in both schools seemed to find teacher-oriented activities (i.e. quizzes, class questions, story writing) much more positive than the more formal and less frequent standardized tests and chapter tests. It would be interesting and useful (for instructional purposes) to ascertain whether the frequency and source of a test, as well as its potential effect on a student's career, influence student motivation and attitude toward assessment.

Ratings of writing a story are also worth exploring. This assessment technique was thought to be the least important though the most fun and best liked activity. Did students consider this to be an assessment activity or an instructional technique? Had they been asked for their ratings on writing an essay in science or history, would their ratings have changed?

CONCLUSIONS

We began this report by suggesting that federal, state, and local agencies need to be aware of the costs and implications of different kinds of assessment practices. What does our research offer in the way of estimating the magnitude of testing costs and for elaborating their implications?

Monetary Costs of Achievement Testing

In our large, urban elementary school, annual costs for achievement testing of all types in all subjects was about $130.00 per pupil (enrollment = 830). In our small, suburban elementary school, annual costs for achievement testing of all types in all subjects was about $245.00 per pupil (enrollment = 191). Nearly all of these costs were incurred as a result of staff time devoted to testing.

In our two study schools, teachers devoted about 12 to 15 percent of their annual time to testing, and students gave up 9 to 10 percent of their annual classroom time to testing.

Are our two elementary schools in any way representative of other elementary schools across the country? We offer a tentative "we think they might be." That is, based on our national test use survey findings, we make a crude estimate that the kinds of figures reported above may be similar to costs incurred in other elementary schools.

In our two study schools, testing does not appear to infringe greatly on students' instructional time, nor do the costs of assessment in teacher time seem especially high. Further, the direct costs of testing do not appear to be great. Therefore, if other elementary

schools display patterns similar to those in our study, then even if they or their districts cut back sharply on the amount of testing they do, including mandated testing, they would not generate a vast sum of re-allocatable dollars.

Psychological Costs of Testing -- Teachers

In our two study schools, it does not seem that testing or the uses of tests results create deep-seated worry or distress. There does, however, appear to be a degree of concern, perhaps even anxiety, among some teachers. And among these teachers, our findings indicate that there are certain psychological costs of testing, sometimes for teachers, sometimes for their students, and sometimes for both.

Concerns expressed to us generally fell into the areas of test utility -- lateness of test score reports, lack of test relevance, redundancy and/or differential value of parts of a testing program; test appropriateness -- differential ease/difficulty of parts of the testing program, technical concern with test validity for some students; effects of testing -- student worry over formal tests, difficulty in motivating students to do well, adverse publicity resulting from reporting of test scores; and finally, but to a lesserr extent, loss of instructional time.

While these findings (which are generally corroborated in the student data) suggest we face no insurmountable problems in the psychological costs of testing, improvement in district and school testing policy might be considered. We have suggested elsewhere (Burry et al., 1981) the kinds of things a district might consider in the interest of establishing a coherent, economical, multiple-purpose

assessment program that might decrease teacher concerns about testing and its effects.

Psychological Costs of Testing -- Students

Students in our two elementary schools did appear to view assessment activities as more important than non-assessment activities, such as recess or P.E. games. Among the assessment activities, standardized and chapter tests were rated as most important by students themselves; more than half of them also thought that their teachers would agree with this perception, and about one-third also ascribed this point of view to their parents.

While we have no parent data which speak to this question, many of our teacher respondents would probably offer a different picture, one in which chapter tests might maintain prominence while standardized tests would diminish in stature. Again, the question of district and school testing policy comes to the forefront. In particular, what is the school's perception of the importance of standardized tests? What is the teacher's perception? Is this perception similar across all teachers in a grade level? All teachers in a school? How do students come to acquire their sense of the importance of standardized tests? Other forms of assessment?

Related to the matter of the importance of tests, while some students dislike those tests that they rated as important -- standardized and chapter tests -- a majority of the students viewed these activities either positively or neutrally. Further, only about one third of the students reported that they actively worry about these forms of testing; most reported feeling somewhat calm across all assessment modes.

Our findings indicate that the resource and psychological costs of testing are relatively modest for most teachers and students. However, in times of scarce resources even small marginal gains may be significant. Are the costs of testing too high, are they about right? These judgments must be made in relation to the benefits of testing. Enlightened testing policy requires attention to both sides of the equation.

REFERENCES

Burry, J., Dorr-Bremme, D., Herman, J.L., Lazar-Morrison, C., Lehman, J., & Yeh, J.P. Teaching and testing: Allies or adversaries? CSE Report No. 165. Los Angeles,: UCLA Center for the Study of Evaluation, 1981.

Burry, J., Catterall, J., Choppin, B., & Dorr-Bremme, D. Testing in the nation's schools and districts: How much? What kinds? To what ends? At what costs? CSE Report No. 194. Los Angeles: UCLA Center for the Study of Evaluation, 1982.

Dorr-Bremme, D. Assessing students: Teachers' routine practices and reasoning. Evaluation Comment, 1983, 6(4), 1-12.

Kirkland, M. The effects of testing on students and schools. Review of Educational Research, 1971, 41(4), 303-350.

Lazar-Morrison, C., Polin, L., Moyt, R., & Burry, J. A review of the literature on test use. CSE Report No. 144. Los Angeles: UCLA Center for the Study of Evaluation, 1980.

Sharp, E. How teachers see testing. Educational Leadership, 1966, 24(2), 141-143.

Stetz, E., & Beck, M. Comments from the classroom: Teachers' and students' opinions of achievement tests. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

Yeh, J.P. Test use in schools. Los Angeles: UCLA Center for the Study of Evaluation, 1978.