

## College Prep Math Assessment in Algebra I and Geometry 1992 Results

This report is a careful look at the results of a study of approximately 2400 students from Algebra and Geometry classes who participated in an assessment during May of 1992. Each student worked on two of the exam questions like the ones shown in Attachment 1 and was required to show his or her work.

Papers from CPM and non-CPM students were coded so that the students could not be identified and mixed together to guarantee uniform grading. Each paper was graded holistically on a scale of 0-4 with 3 representing satisfactory work and 4 representing excellent work.

The overall result is that CPM students have uniformly higher means on these tests than students from non-CPM classes (more than 35% higher on algebra and more than 28% higher in geometry). The results are true for both sexes, all ethnic groups, all grade levels and are significant overall at the .0001 level; that is, there is a probability of .0001 or less that the CPM and non-CPM students performed equally well. (See the “1992 Report by Decile” for additional information.)

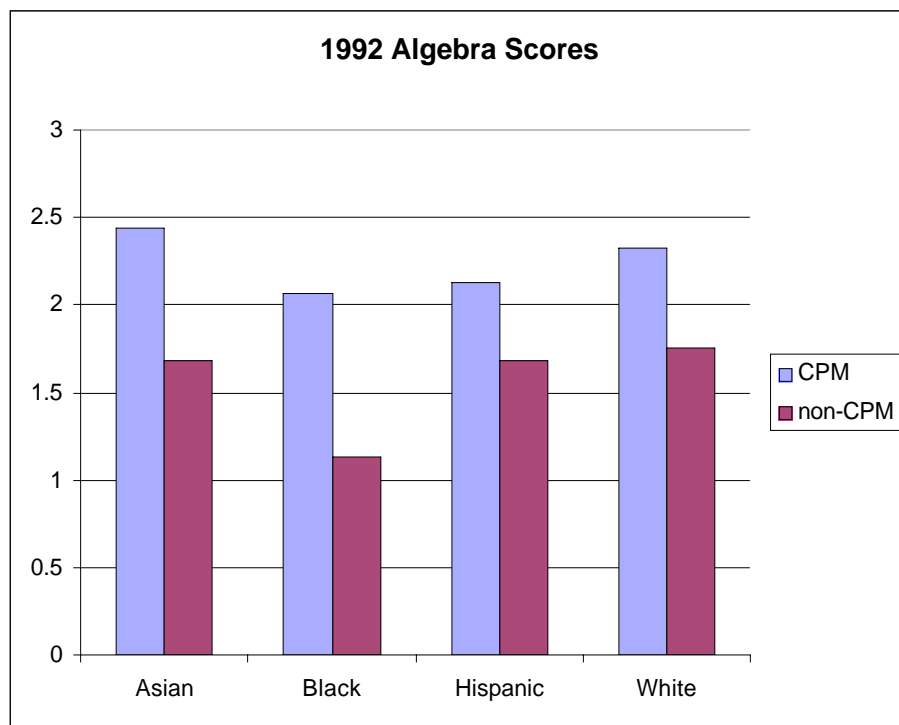


Figure 1

## Methodology

In May, 1992, examinations were sent to 16 schools which had both CPM and non-CPM classes in algebra and geometry. (Here we use the term "non-CPM" to denote a class using any one of the various standard textbook series--e.g., Houghton-Mifflin, Prentice-Hall, etc.) Each CPM teacher was asked to find non-CPM teachers at his or her school who would give the examination to their students as well so that we could be assured of roughly comparable groups of students taking the examination. All of the teachers involved were volunteers.

Each of the students was asked to fill out a brief questionnaire detailing their age, grade level, gender and ethnicity. Then students were given 20 minutes to work on the two questions from one of six forms of the assessment. (Typical questions are shown in Attachment A.) Care was taken to ensure that each of the tested classes used approximately the same number of each of the six different forms of the test. Each teacher also filled out a questionnaire giving their gender, number of years of teaching experience, the text used, and the number of years they had taught from this particular text.

All of the examinations were sent to the CRESS Center at UC Davis, where each student's responses were coded by the same number on their information sheet and each of their two response sheets. Then the papers were separated into piles by question number and the papers within each pile were shuffled so that they would be graded in random order. Approximately 1800 students took the Algebra 1 assessment and 650 students took the Geometry assessment. From these original numbers, we were able to salvage 1740 Algebra and 606 Geometry complete data sets. In each case about 2/3 of the papers were done by CPM students.

During June of 1992, a group of 30 teachers participated in the grading which was done holistically on a 5-point scale. On this scale a 0 represented no useful work (either nothing done or nonsense); 1 represented an understanding of the problem, but little progress toward a solution; 2 a reasonable start toward a solution, but an approach with significant gaps or misconceptions; 3 an almost complete solution which might contain minor errors; and 4 a full and complete solution. For each problem, the group agreed on a rubric for the scoring by looking at several sample papers. After the group had come to an agreement the shuffled papers were each graded by at least two people and, in case of disagreement, by a third person. The consensus score was recorded.

## **Analysis of the data.**

All of the data was analyzed by the SPSS statistical package with the assistance of Neil Willett of the UC Davis Statistical Consulting Laboratory. In the first pass of the data it was determined that there was no significant difference in the performance of males and females on the same problem, but there were differences attributable to differences between ethnic groups, grade levels, teachers, schools and program. Because of differences between individual teachers and schools, the data was analyzed by a *random effects* analysis, where we assume that differences between schools and teachers exist as well as by a *fixed effects* analysis, where we assume that all teachers and schools are substantially identical.

**In each case of program comparison, the difference favored CPM students, generally by a substantial and statistically significant margin.**

When ethnic differences existed within a program, the white and Asian (and sometimes Hispanic) students scored higher than Blacks. However, CPM students of all ethnic groups outscored traditional students of all ethnic groups. That is, Black CPM students scored higher than students from any ethnic group in traditional classes. (This is illustrated in the graph of Figure 1 under the Executive Summary.)

In addition to comparing overall means of the two groups, we compared the proportion of those who scored a 3 or a 4 on their responses. These were the responses that were considered essentially correct. Again, the proportion of CPM students who scored a 3 or a 4 was significantly higher than the proportion for students in traditional classes.

Attachment B contains details about numbers of students in each category together with a variety of graphs which illustrate the statements above while a summary report of the statistics is given in Attachment C.

## Attachment A          Sample problems

Two typical problems from each assessment are given below. The two numbers after each problem give the mean scores of the CPM students and the non-CPM students, respectively. A single asterisk denotes a significant difference at the 5% level for this particular problem and a double asterisk denotes a significant difference at the 1% level.

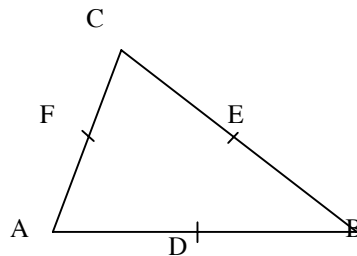
### Algebra Problems

- C1. The city of Nampa currently has 45,000 people and is growing at 13,000 people per year. Caldwell has 66,000 people and is growing at 9,000 people per year. In how many years will the two towns have the same population? (3.00, 2.14)\*\*
- D1. You are in charge of buying potato salad for the annual Central High picnic and are told that 50 pounds of salad will feed about 140 people. About how many pounds of salad should you buy if you want to feed 300 people? (3.11, 2.46)\*\*

### Geometry Problems

- C1. Begin with a regular 7-sided polygon ABCDEFG. Extend the sides AB and DE until they meet at point P. Find the measures of each of the angles of the pentagon APEFG. (2.76, 1.21)\*\*

- D2. In  $\triangle ABC$ ,  $AB = 30$  cm,  $BC = 22$  cm, and  $AC = 18$  cm. Point D is the midpoint of AB, E is the midpoint of BC and F is the midpoint of AC. Prove that  $\triangle BDE$  is congruent to  $\triangle EFC$ . (2.28, 1.59)\*



## Attachment B.

1. The tables below show that the two groups of students who took the algebra examination were essentially the same. Percentages in each group are given. There were 1001 CPM students and 739 non-CPM students in the algebra sample. The tables are similar for the 389 CPM students and 217 non-CPM students in the geometry sample.

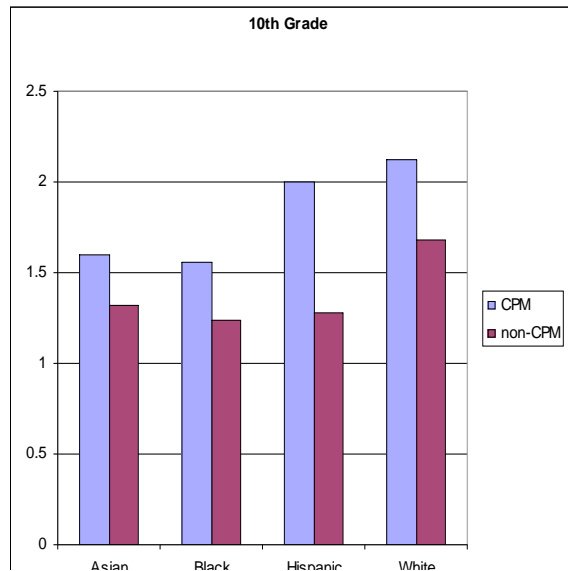
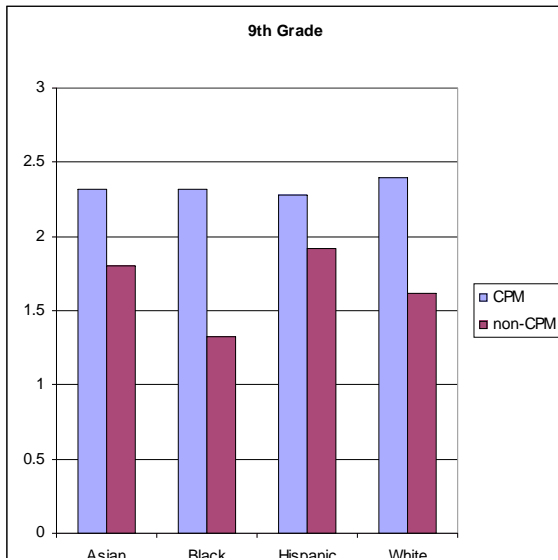
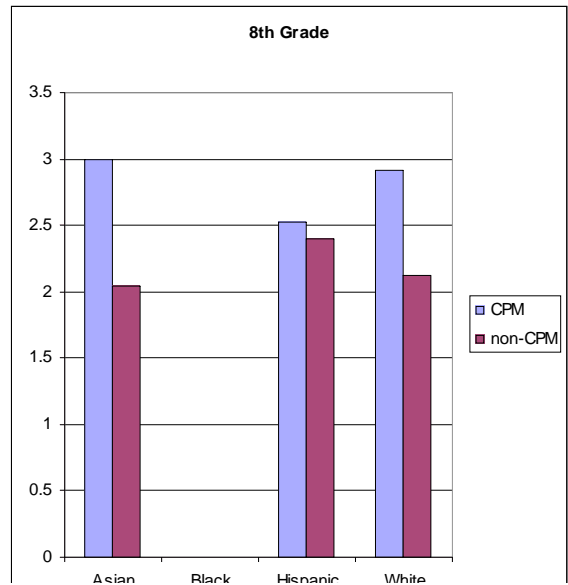
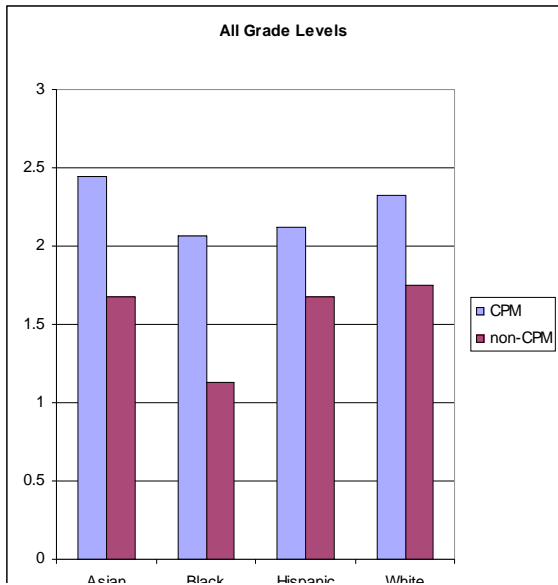
### Summary Statistics for Algebra Students

	% CPM	% non-CPM
Asian	11	11
Black	7	8
Hispanic	21	16
White	56	62
Other	5	3

	% CPM	% non-CPM
8th	18	17
9th	48	47
10th	20	24
11th	9	10
12th	3	1

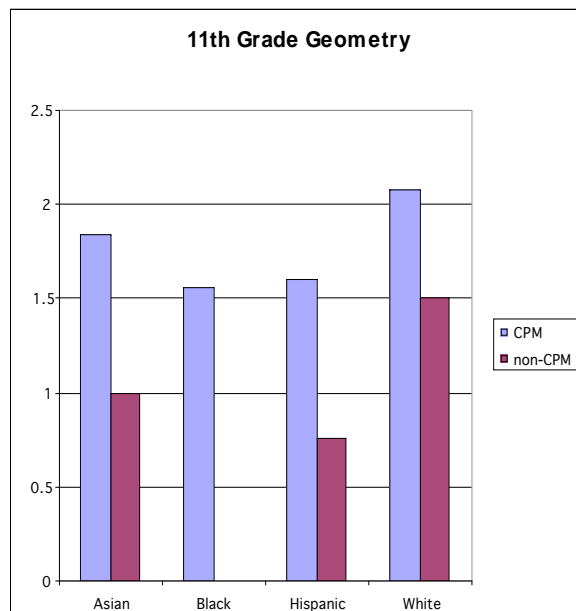
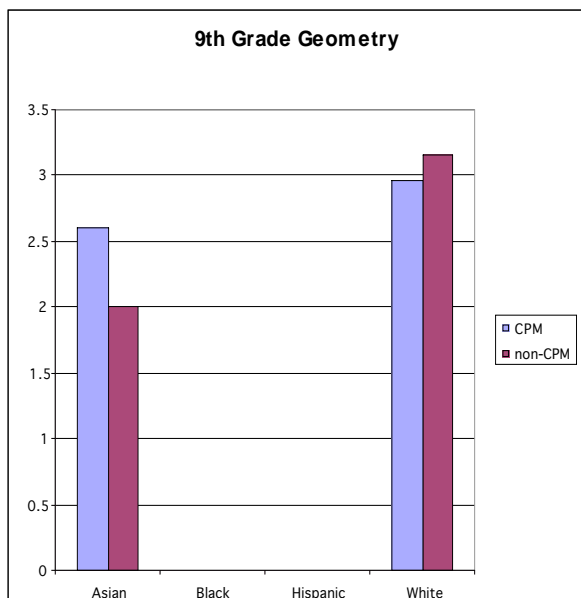
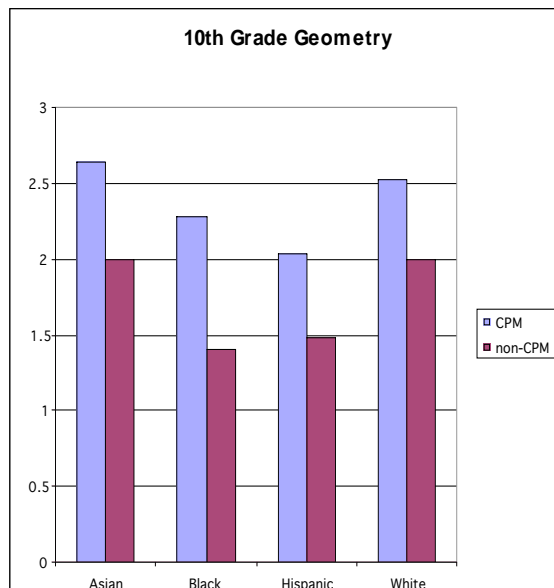
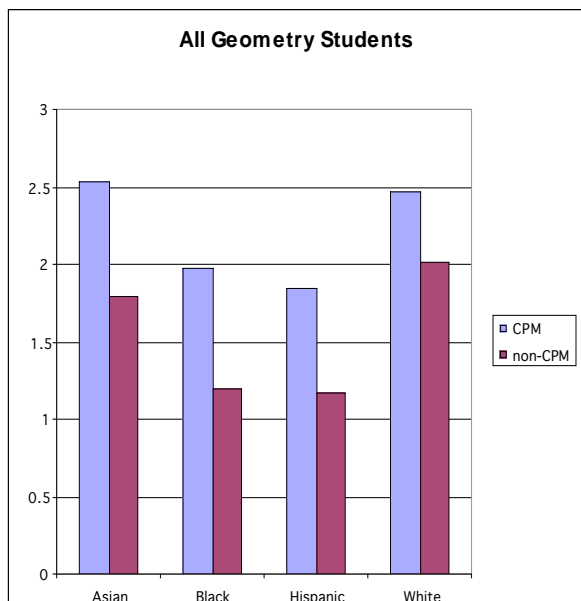
## Algebra Scores By Ethnic Group

In each graph, the bars give the percentage by each major ethnic group with the light bars representing CPM students and the dark bars representing the scores of non-CPM students. Missing bars reflect insufficient data for that group.



## Geometry Scores By Ethnic Group

The graphs for the geometry students are shown below in the same format as for the algebra students. Missing bars reflect insufficient data for that group.



## Attachment C

The data that follows all comes from reports done by Neil Willits of the University of California, Davis, Statistical Consulting Laboratory. The original reports ran to over 80 pages, but only the summary data is shown.

**Significant Differences.** The data shows the F and p values for the scores of all students on all forms of the test. Not surprisingly, grade effects are very large (eighth graders do better than ninth or tenth). There was some effect due to ethnic differences, but the most important result is that program effects—that is the difference in success between CPM and non-CPM students—are very large and are significant at the .0001 level.

### 1992 CPM diagnostic test results, Algebra students

<u>form</u>	<u>variable</u>	<u>effect/analysis</u>	<u>F</u>	<u>p</u>
(all)	scores 1+2	form	12.17	.0001
		program	40.96	.0001
		program (random)	12.18	.0001
		grade	26.12	.0001
		ethnicity	5.74	.0001

### 1992 CPM diagnostic test results, Geometry students

<u>form</u>	<u>variable</u>	<u>effect/analysis</u>	<u>F</u>	<u>p</u>
(all)	scores 1+2	form	9.72	.0001
		program	30.21	.0001
		program (random)	9.72	.0001
		grade	41.58	.0001
		ethnicity	2.90	.0137