DOCUMENT RESUME

ED 225 834 SE 040 193

AUTHOR Herbert, Martin
TITLE Fifth Grade Evaluation: Volume II, Test Data. Evaluation Report 7-B-2. Extended Pilot Trial of the Comprehensive School Mathematics Program.
INSTITUTION CEMREL, Inc., St. Louis, Mo.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Dec 80
NOTE 116p.; For related documents, see SE 040 181-196 and SE 040 340-348.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Educational Research; *Elementary School Mathematics; Evaluation Methods; *Grade 5; Intermediate Grades; *Mathematics Achievement; Mathematics Curriculum; Mathematics Instruction; *Program Evaluation; *Quantitative Tests
IDENTIFIERS *Comprehensive School Mathematics Program; *Mathematics Education Research

ABSTRACT

The Comprehensive School Mathematics Program (CSMP) is a program of CEMREL, Inc., one of the national educational laboratories, and was funded by the National Institute of Education (NIE). Its major purpose is the development of curriculum materials for grades kindergarten through 6. This volume describes the results of a series of mathematics achievement tests administered in spring 1980 to 31 CSMP and 25 non-CSMP classes. Report sections cover: 1) Setting; 2) The MANS Tests; 3) Methods of Analysis; 4) Comparison of Class Means; 5) MANS scores According to Reading Level of Student; 6) Graphs of District Means; and 7) New Students. There are two appendices: A) Statistical Data for Each MANS Scale; and B) Comparison of Results Using Different Units of Analysis. It is noted that within the 13 scale categories, there was a significnat CSMP advantage in 6 when one looks at either the whole category or at the proportion of individual scales that were significant. In 3, there were significant differences in favor of CSMP for the whole category. There were no significant differences in categories titled computation, most reasonable answer, measurement estimation, and organizing data. (MP)

EXTENDED PILOT TRIALS OF THE COMPREHENSIVE SCHOOL MATHEMATICS PROGRAM:

EVALUATION REPORT SERIES

Evaluation Report 7-B-2

Fifth Grade Evaluation: Volume II, Test Data

Extended Pilot Trial of the Comprehensive School Mathematics Program

Evaluation Report 7-B-2

Fifth Grade Evaluation: Volume II, Test Data

Martin Herbert

December, 1980

Developed by CEMREL, Inc., a private nonprofit corporation supported in part as an educational laboratory by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement should be inferred.

Description of Evaluation Report Series

The Comprehensive School Mathematics Program (CSMP) is a program of CEMREL, Inc., one of the national educational laboratories; and is funded by the National Institute of Education. Its major purpose is the development of curriculum materials for grades K-6.

Beginning in September, 1973, CSMP began an extended pilot trial of its Elementary Program. The pilot trial is longitudinal in nature; students who began using CSMP materials in kindergarten or first grade in 1973-74, were able to use them in first and second grades respectively in 1974-75, and so on in subsequent years. Hence the adjective "extended".

The evaluation of the program in this extended pilot trial is intended to be reasonably comprehensive and to supply information desired by a wide variety of audiences. For that reason the reports in this series are reasonably non-technical and do not attempt to widely explore some of the related issues. The list of reports through year six is given on the next page. The following reports are planned for year 7:

7-B-1 - Fifth Grade Evaluation: Volume I, Summary
7-B-2 - Fifth Grade Evaluation: Volume II, Test Data
7-B-3 - Fifth Grade Evaluation: Volume III, Non-Test Data
7-B-4 - Reevaluation of Second Grade, Revised MANS Tests
7-B-5 - Achievement of Former CSMP Students at Fourth Grade
7-B-6 - Student Achievement, Rapid Implementation Model

Extended Pilot Trials of the
Comprehensive School Mathematics Program

Evaluation Report Series

Evaluation Report 1-A-1	Overview, Design and Instrumentation
Evaluation Report 1-A-2	External Review of CSMP Materials
Evaluation Report 1-A-3	Final Summary Report Year 1
Evaluation Report 1-B-1	Mid-Year Test Data: CSMP First Grade Content
Evaluation Report 1-B-2	End-of-Year Test Data: CSMP First Grade Content
Evaluation Report 1-B-3	End-of-Year Test Data: Standard First Grade Content
Evaluation Report 1-B-4	End-of-Year Test Data: CSMP Kindergarten Content
Evaluation Report 1-B-5	Test Data on Some General Cognitive Skills
Evaluation Report 1-B-6	Summary Test Data: Detroit Schools
Evaluation Report 1-C-1	Teacher Training Report
Evaluation Report 1-C-2	Observations of CSMP First Grade Classes
Evaluation Report 1-C-3	Mid-Year Data from Teacher Questionnaires
Evaluation Report 1-C-4	End-of-Year Data from Teacher Questionnaires
Evaluation Report 1-C-5	Interviews with CSMP Kindergarten Teachers
Evaluation Report 1-C-6	Analysis of Teacher Logs
Evaluation Report 2-A-1	Final Summary Report Year 2
Evaluation Report 2-B-1	Second Grade Test Data
Evaluation Report 2-B-2	Readministration of First Grade Test Items
Evaluation Report 2-B-3	Student Interviews
Evaluation Report 2-C-1	Teacher Questionnaire Data
Evaluation Report 2-C-2	Teacher Interviews, Second Grade
Evaluation Report 2-C-3	Teacher Interviews, First Grade
Evaluation Report 3-B-1	Second and Third Grade Test Data Year 3
Evaluation Report 3-C-1	Teacher Questionnaire Data Year 3
Evaluation Report 4-A-1	Final Summary Report Year 4
Evaluation Report 4-B-1	Standardized Test Data, Third Grade
Evaluation Report 4-B-2	Mathematics Applied to Novel Situations (MANS) Test Data
Evaluation Report 4-B-3	Individually Administered Problems, Third Grade
Evaluation Report 4-C-1	Teacher Questionnaire Data, Third Grade
Evaluation Report 5-B-1	Fourth Grade MANS Test Data
Evaluation Report 5-B-2	Individually Administered Problems, Fourth Grade
Evaluation Report 5-C-1	Teacher Questionnaire and Interview Data, Fourth Grade
Evaluation Report 6-B-1	Comparative Test Data: Fourth Grade
Evaluation Report 6-B-2	Preliminary Test Data: Fifth Grade
Evaluation Report 6-C-1	Teacher Questionnaire Data: Grades 3-5

Key to Indexing

Evaluation Reports are labelled m-X-n,

where m is the year of the pilot study, with 1973-74 as Year 1.
X is the type of data being reported where A is for overviews and summaries, B is for student outcomes and C is for other data.
n is the number within a given year and type of data.

Table of Contents

Introduction to Volume II

In the spring of 1980, a series of mathematics achievement tests were administered to 31 fifth grade classes using the Comprehensive School Mathematics Program and to 25 comparison classes using more traditional programs. Volume II of this report describes the results of the testing; Volume III provides information dealing with the implementation of the program, and with teacher and student attitudes, and relates these data to test scores. Volume I is a summary report.

Three kinds of tests were administered:

a) The MANS scales, a series of short test scales intended to assess some of the underlying goals of the CSMP curriculum. Various MANS scales have been used in other comparative evaluations, beginning in second grade (see the list of titles from the CSMP Evaluation Report Series, page iii).

b) The items from the Computation Test of the Comprehensive Tests of Basic Skills, Form S, Level 2.

c) The items from the Reading Comprehension Test of the CTBS. These reading scores were used as covariates in the main analysis of class mean scores on the mathematics tests, i.e. they served as a statistical control for differences in the ability level of the various classes.

Setting

Altogether, there were 44 classes studying the 5th grade CSMP curriculum, and 31 of these classes participated in this study. These included classes from all but one of the school districts with more than two 5th grade CSMP classes, though sampling occurred in one site.

Comparison classes were selected from other schools in the respective district which were thought to be similar to the CSMP schools. In two districts it was not feasible to select comparison classes within the district (either because there were no comparable schools or because CSMP was already in use in all schools), but two other districts just beginning CSMP at lower grades agreed to provide comparison classes.

A brief description of the sites is given in Table 1 below.

Table 1

Site Designations[1]	Number of Classes CSMP	Number of Classes non-CSMP	Type of Community	Approximate Socio-Economic Status
1 △	6	0	Suburb of small city	Middle
2 +	0	5	Small city	Middle/Lower middle
3 ◇	5	3	Inner-city of large city	Low
4 ○[2]	9	6	Inner suburb of large city	Middle/Lower middle
5 ×	0	3	Medium City	Middle
6 ▽	2	2	Exurban	Middle/Lower middle
7 □	6	6	Suburb of large city	Upper Middle
8 ⬭	3	0	Suburb of large city	Upper Middle
Total	31	25		

[1]There are two designations for each site. The numerals are used in the graphs of district means (page 27) and the geometric symbols are used in the graphs of class means (Appendix A).

[2]Site 4 is made of classes from 4 school districts which were similar, located near one another and treated as a single site in this study. Otherwise, site = school district.

Except for districts 3 and 8, all CSMP classes had studied CSMP since first grade. In these two districts however, these students began the program in fourth grade with a special "entry" program and consequently were about one-half a semester behind the other classes.

Based on scores from the reading test administered to all classes, the 30 classes were reasonably similar in ability to the 25 non-CSMP classes. Furthermore, the classes tended to be rather above average in ability, with the mean reading scores corresponding to percentile ranks of about 61 and 60 for CSMP and non-CSMP respectively.

The MANS Tests

The MANS Tests (Mathematics Applied to Novel Situations) are short test scales developed especially to assess what are thought to be some of the underlying thinking skills of CSMP. MANS scales of various kinds have been used in the evaluation of CSMP in second through fifth grade.

The scales are administered by trained testers, who follow a standardized script including sample problems for each scale. Then the students do the test items in that scale and the process is repeated for the next scale. The scales do not contain any of the special vocabulary or techniques of the CSMP program and most of them are built around mathematical situations that are unfamiliar to both CSMP and non-CSMP students.

An intensive pilot test and review procedure is used in developing MANS Scales; Evaluation Report 4-B-3 contains a detailed description of this process in an earlier study. Previous scales are often reused and new ones continually added. In the present study there were a total of 32 MANS scales, containing an average of about 8 items and requiring an average of about 5 minutes each, though the tests were essentially untimed except for those dealing with estimation. Three testing periods of 50-60 minutes each were required.

Method of Analysis

Although various analyses were carried out at the student, school and district levels, the main analysis was done on class means. For each test scale, a mean score was calculated across all the students in class who took the test[1] and who also took the reading test. The corresponding mean reading test score was also calculated. In both cases raw scores were used. Appendix A gives the graphs of these class means so that one can compare visually test score versus reading score for the set of 56 classes.

An analysis of covariance procedure (1 and 48 degrees of freedom with reading as covariate) was then used to compare the mean score for the 31 CSMP classes versus the mean score for the 25 non-CSMP classes on each individual scale. Analysis of class mean data is presented in the next section, "Comparison of Class Means".

A later section, "Graphs of District Means", page 27, presents graphs of various test scores aggregated by district.

Analysis of results across students rather than across classes is included in Appendix A. In fact, Appendix A, by itself serves as a fairly complete report of the results.

Finally, Appendix B compares the results obtained when an analysis of covariance procedure was used with classes, schools, and districts as the various units of analysis.

[1] New students who joined their class after the end of September were not included in the analysis; see page 35.

Comparison of Class Means

On the following pages, summary data are presented for each scale. The scales have been grouped into categories according to the kind of task involved.

For each scale, a brief abstract and sample item are given. Then the mean scores across CSMP classes and across non-CSMP classes are compared.[1] Finally, the p-value[2] of this comparison is given, i.e. the probability that a difference that large between the two groups could have occurred by chance if the two groups were "really equal". A p-value of .05 or less is often designated as "significant".

Since class means were the major unit of analysis, it was possible to optimize the time available for testing by having random halves of each class take different scales. This was possible in the cases where the two scales had identical directions and those scales in which this occurred will be indicated on the following pages. It was also possible on occasion to do this at the item level: on the CTBS Reading Comprehension Test, half the students took one set of 25 items while the other half took another set of 25 items, there being 5 items in common. Thus for some of the scales, the class means were based on a random of half of the students in the class. This is one of the reasons that the corresponding mean reading score for a class varied slightly from test to test.

A summary of the results for the various scale categories is given on page 19.

[1] What are given are actually adjusted mean scores, i.e. mean scores adjusted for differences in reading ability between the two groups. Since such differences were small, these adjustments amounted to about 0.1 on most of the scales.

[2] Using an F-test with 1 and 48 degrees of freedom.

C1 CTBS Computation, Level 2, Form S

48 multiple-choice items, 12 for each operation
Roughly half the items involved whole number algorithms, a quarter of them involved fractions, and a quarter decimals.
A random half of each class took a set of 24 items, the other half of the class took the other, 24-item set.
Class means are based on the entire set of 48 items.

	Adjusted Means CSMP	Adjusted Means non-CSMP	p-value
a) Addition	9.6	9.3	.07
b) Subtraction	9.6	9.0	.53
c) Multiplication	8.5	8.4	.68
d) Division	7.6	7.6	.95
Total, Computation	34.9	34.3	.42

MENTAL ARITHMETIC Scales

		Adjusted Means CSMP	non-CSMP	p-value
C3-C6	Mental Arithmetic			
	Each item was an open number sentence to be done mentally, i.e. without "scratch" work. The answer box could appear on either side of the equals sign. Many of the items required more than merely calculation skills (see C6). Half the students did C3,C6; the other half did C4,C5.			
C3	Addition (5 items)	3.4	2.9	.01
	Sample: 9,001 + ☐ = 9,100			
C4	Subtraction (5 items)	3.1	2.5	.01
	Sample: 700 - 401 = ☐			
C5	Multiplication (11 items)	6.4	5.1	.01
	Sample: 12 x 500 = ☐			
C6	Division (11 items)	6.7	5.4	.01
	Sample: 1,200 divided by 30 = 40 1,200 divided by 15 = ☐			
	Total, Mental Arithmetic	19.7	15.9	.01

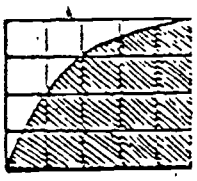
		Adjusted Means CSMP	non-CSMP	p-value
E2-E4	Estimation Intervals			
	Determine which of several given intervals contains the answer to a computation problem. There was a time limit of 1½ minutes for each of E2, E3, E4.			
E2	Addition (8 items)	6.6	6.3	.14
	Sample:			
	279 + 165 0 10 50 100 500 1000			
E3	Multiplication (7 items)	5.2	4.7	.01
	Sample:			
	11 x 50 0 10 50 100 500 1000			
E4	Division (7 items)	3.9	3.4	.01
	Sample:			
	133 divided by 50 0 1 10 20 100			
	Total, Estimating Intervals	15.6	14.5	.01
E6-E9	Most Reasonable Answer			
	For a given computation problem, determine which of 3 answers (all of which are wrong) is most reasonable. Half the students took E6 and E9; the others took E7, E8. There was a time limit of $1\frac{1}{2}$ minutes for each of E6-E9.			
	Example: 21 x 123 = 257 / 2,557 / 25,557			
E6	Addition (6 items)	3.9	3.7	.09
E7	Subtraction (6 items)	3.3	3.4	.38
E8	Multiplication (6 items)	3.1	3.0	.86
E9	Division (6 items)	2.7	2.7	.69
	Total, Most Reasonable Answer	12.9	12.9	.82

MEASUREMENT ESTIMATION SCALE

	Adjusted Means CSMP	Adjusted Means non-CSMP	p-value
M1 Measurement Estimation (6 items)	2.1	2.1	.81

Estimate the answer to a visually presented problem in area, volume, height, etc.
A range of answers was accepted.

Sample:



This playground is divided into 20 sections.
It takes one gallon of paint to cover one section.

About how many gallons of paint would it take to cover the shaded part of the playground? _____

NEGATIVE NUMBERS Scale[1]

	Adjusted Means CSMP	Adjusted Means non-CSMP	p-value
N2 Negative Hits and Misses (10 items)	6.4	5.8	.05

Given two rules: each hit means a gain of 5 points
each miss means a loss of 1 point

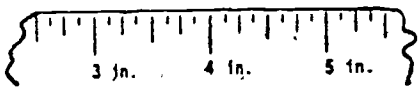
Determine the missing piece of information.
Half the students took one set of 5 items, the others took 5 other items of a similar format.

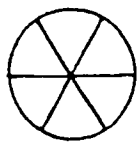
Sample:

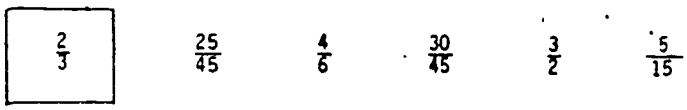
Peter	Started with a score of	Number of Hits	Number of Misses	Ended with a score of
	10 below zero	1		12 below zero

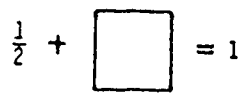
[1] Negative Numbers, Decimals and Fractions (next pages) were all labelled "N" for Number Systems.

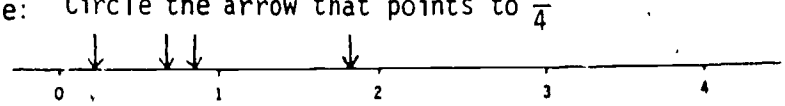
		Adjusted Means CSMP	non-CSMP	p-value
	(Half the students took N3, N6, and N9; the other half took N5, N7, N8, and N10.)			
N3	Measuring Fractional Inches (3 items)	1.4	1.6	.17
	Sample: Put an arrow at $3\frac{3}{4}$ inches			
N5	Fractional Areas (8 items)	4.1	3.9	.25
	Sample: Shade $\frac{2}{3}$ of the figure			
N6	Equivalent Fractions (4 items x 5 per item)	14.0	13.4	.19
	Sample: Circle the fractions that are equal to the one in the box.			
N7	Fractional Open Sentences (6 items)	3.2	2.6	.01
	Sample: $\frac{1}{2} + \square = 1$			
N8	Which Fraction is Larger (5 items)	3.4	3.1	.13
	Sample: $\frac{3}{4}$ or $\frac{5}{10}$			
N9	Fractional Word Problems (5 items)	3.0	2.6	.01
	Sample: $\frac{1}{4}$ of a 200-page book is _____ pages.			
N10	Other Representations of Fractions (6 items)	3.9	3.9	.85
	Sample: Circle the arrow that points to $\frac{1}{4}$			
	Total, Fractions	33.0	31.1	.03











15

		Adjusted Means CSMP	non-CSMP	p-value
N1	Decimal Gas (7 items)	4.5	3.5	.01

A series of simply worded word-problems about gasoline involving decimal numbers.

Sample:

Tom has 6.5 gallons.

He buys 3.5 more gallons.

How much gas will he have then? _____

N4	Decimal Magnitudes (10 items)	6.6	5.0	.01

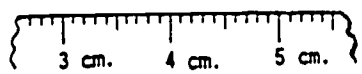
A composite of two kinds of items:

Sample 1:

Which is larger? 4.999 or 5.1

Sample 2:

Put an arrow at 3.4 cm.



Total, Decimals	11.1	8.5	.01

ORGANIZING & INTERPRETING DATA Scale

		Adjusted Means CSMP	non-CSMP	p-value
O1	Weight Graph (10 items)	6.5	6.5	.81

Given a graph in which weight (axis labelled at 10 pound increments for each 5 units) is plotted against age (axis labelled at 2 year increments for each 2 units), determine age per given weights and vice versa.

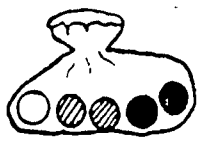
PROBABILITY Scales

P1	100 Outcomes (24 items)	13.5	12.3	.02

Various random devices are given.
In 100 trials give the best estimate for how often each outcome will occur?

Sample:

Joe plays the game with marbles and a bag.
He closes his eyes and takes a marble out.
Then he puts it back.



SUPPOSE JOE PLAYED THE GAME 100 TIMES

About how many times would he get a black marble? _____

About how many times would he get a white marble? _____

About how many times would he get a shaded marble? _____

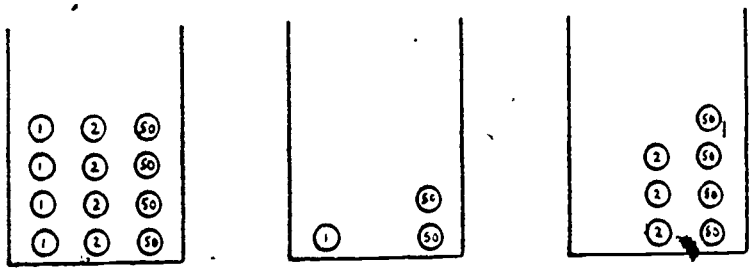
About how many times would he get a marble that is not white? _____

P2	Which Box? (6 items)	3.7	3.5	.40

Given three boxes containing various 1, 2 and 50-cent "balls", determine from which box it would be best to make a blind draw.

Sample:

WHICH BOX WOULD YOU CHOOSE?



20

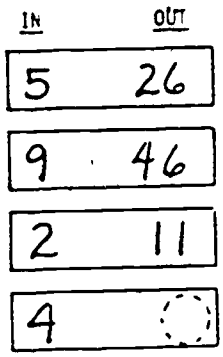
Total Probability	17.2	15.8	.02

NUMBER RELATIONS Scales

		Adjusted Means CSMP	non-CSMP	p-value
R1	Solving Functions (8 items)	5.5	4.8	.01

Given 3 pairs of numbers produced by a "number machine", deduce the missing number from the 4th pair.

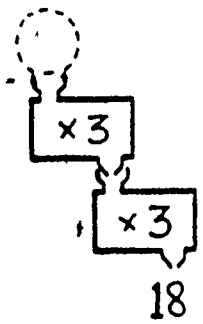
Sample:



R2	Using Number Machines (10 items)	6.7	5.5	.01

Given a set of labelled number machines in sequence, find the original input or the final output.

Sample:



Total, Number Relations	12.2	10.2	.01

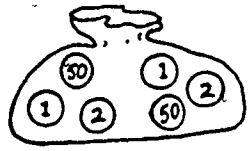
ELUCIDATION Scale

		Adjusted Means CSMP	non-CSMP	p-value
U1	Elucidation (4 problems, 25 possible correct answers)	16.2	13.2	.01

Find as many solutions as possible to a given problem.

Sample:

Close your eyes.
Pick out three balls.
Add to get a total score.



What are the possible total scores? 52, _____

WORD PROBLEMS Scales[1]

W2	Two-Stage Word Problems (7 items)	4.5	4.2	.11

Sample:

Jim has $10 in his bank now.
Each week he will add $5 to his bank.
In how many weeks will he have $30 in his bank? _____

W3	Three-Stage Word Problems (5 items)	2.2	1.9	.12

Sample:

Joe puts boxes into piles.
Each box is $\frac{1}{2}$ foot high.
Each pile is 5 feet high.
How many boxes does he need to make 3 piles? _____

Total, Word Problems	6.6	6.1	.03

[1] Half the students took W2, the other half took W3.

Table 2, below, summarizes the class mean data by categories. Every instance of significant differences favored CSMP classes.

Table 2

Summary of Class Mean Data
by Scale Category

Category	Scales	ns/tn[1]	Adjusted Means CSMP	Adjusted Means non-CSMP	p-value
Computation	C1: a,b,c,d	0/4	34.9	34.3	.42
Mental Arithmetic	C3-C6	4/4	19.7	15.9	.01
Estimating Intervals	E2-E4	2/3	15.6	14.5	.01
Most Reasonable Answer	E6-E9	0/3	12.9	12.9	.82
Measurement Estimation	M1	0/1	2.1	2.1	.81
Negative Numbers	N2	1/1	6.4	5.8	.05
Decimals	N1,N4	2/2	11.1	8.5	.01
Fractions	N3, N5-10	2/7	33.0	31.1	.03
Organizing Data	O1 -	0/0	6.5	6.5	.81
Probability	P1,P3	1/2	17.2	15.8	.02
Number Relations	R1,R2	2/2	12.2	10.2	.01
Elucidation	U1	1/1	16.2	13.2	.01
Word Problems	W2,W3	0/2	6.6	6.1	.03
All Scales		14/31	193.9	177.2	.01

[1] ns/tn = number of scales in category which produced a significant difference ($p < .05$) divided by the total number of scales in category.

It can be seen that in six categories the CSMP advantage was decisive whether one looks at p-value for the whole category (almost always $<.01$) or at proportion of individual scales significant (12/13 times across the six categories). These categories were Mental Arithmetic, Estimating Intervals, Negative Numbers, Decimals, Number Relations, and Elucidation.

In three categories there was a significant difference in favor of CSMP on the total for the category, though most of the individual scales did not produce differences large enough to be significant. These categories were Fractions, Probability, and Word Problems.

In four categories there were no significant differences, either in the category total or in any of the individual scales. These categories were Computation, Most Reasonable Answer, Measurement Estimation, and Organizing Data.

MANS Scores According to Reading Level of Student

Students were assigned to one of four groups according to their reading scores; then the mean scores on each MANS scale were computed separately for CSMP and non-CSMP students in each of these four groups.

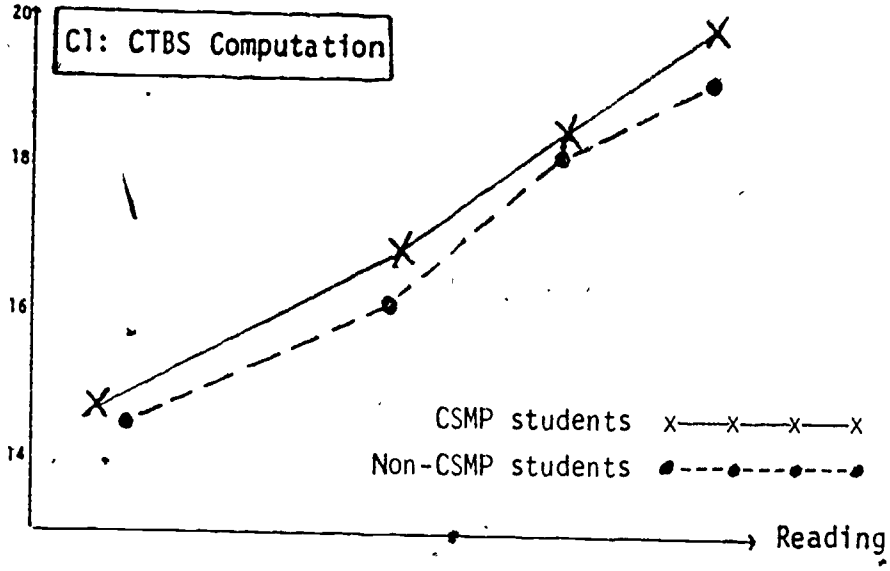
The number and percent of students in each group is shown in the table below.

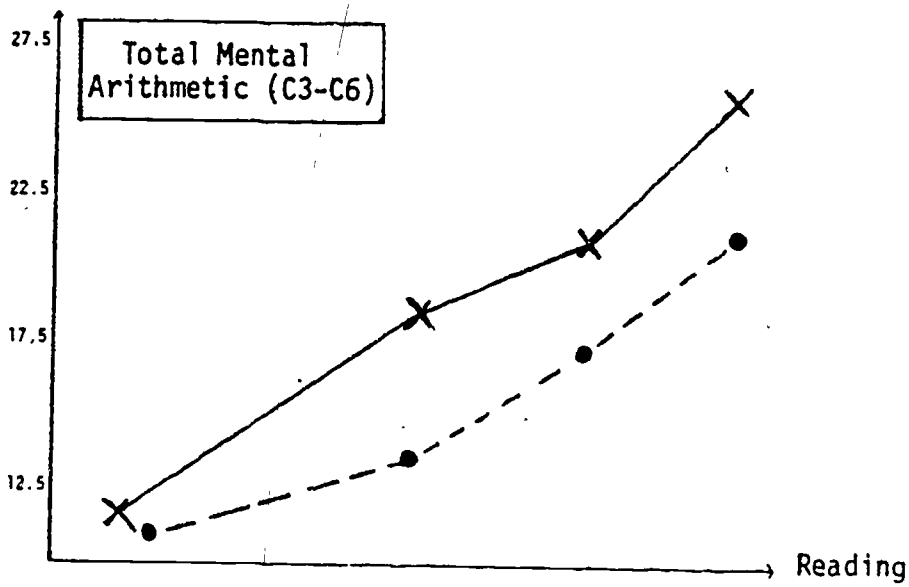
Number and Percent of Students
in Each Reading Group

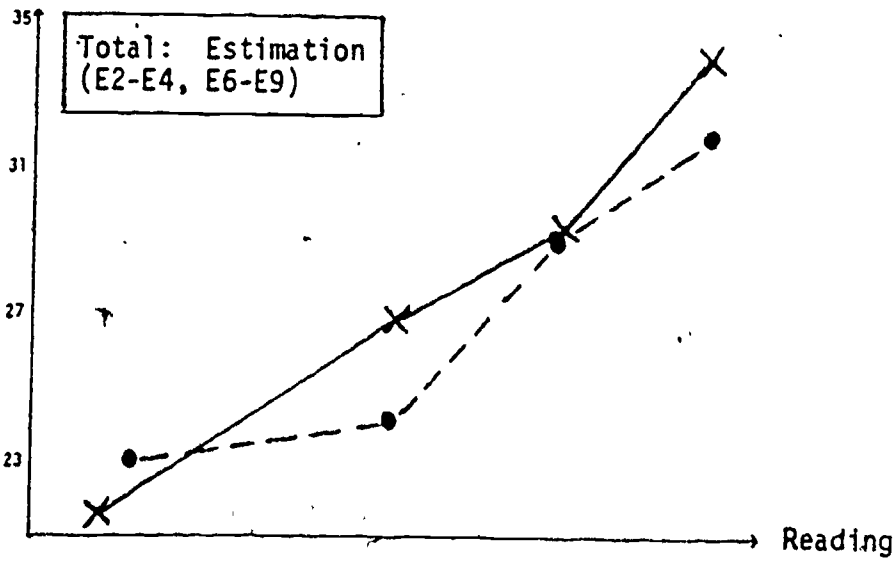
	Mean Reading Score		Number (and Percent)	
	CSMP	non-CSMP	CSMP	non-CSMP
Q4 (lowest quarter)	10.4	11.0	125(21)	114(23)
Q3	16.6	16.4	147(24)	114(23)
Q2	20.1	20.0	147(24)	124(25)
Q1 (highest quarter)	23.2	23.2	186(31)	141(29)
			605	493

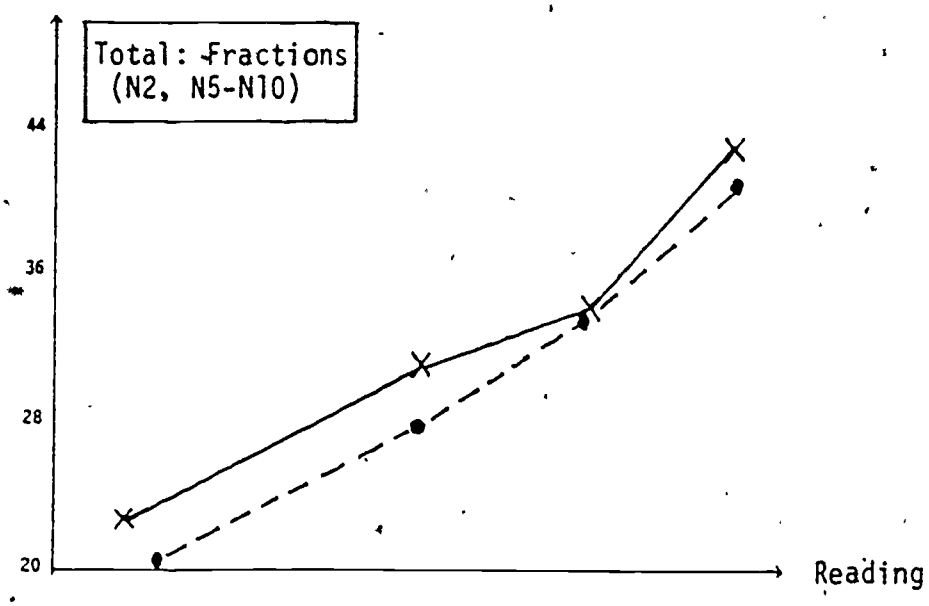
It can be seen there are about equal numbers of students in each group except the highest group which has somewhat more. The points of division between the groups corresponds to approximately the 40th, 60th and 77th percentiles according to the norms of the CTBS Reading Test, though this is only an estimate since individual students only took some of the items of the test. In any case it is clear that the whole group is somewhat higher than average in reading ability with the lowest readers under-represented. Overall, the mean reading score corresponds to a percentile rank of about 60.

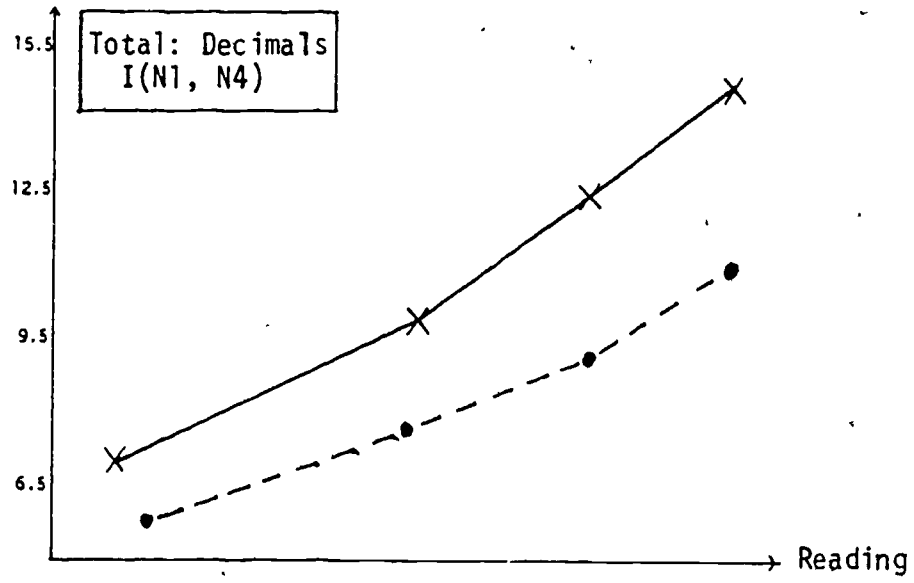
On the pages which follow, graphs are presented for each MANS category. The graphs show, for each reading group, average MANS score versus average Reading score. For CSMP students, x's joined by a solid line represent the four groups (x——x——x——x). For non-CSMP students, dots with a dashed line are used (●---●---●---●).

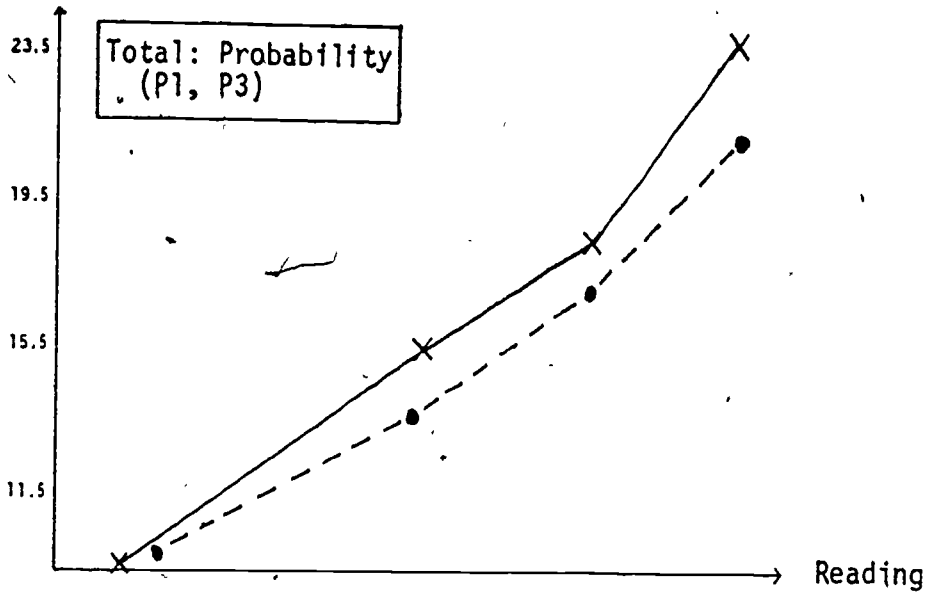

C1: CTBS Computation
20
18
16
14
CSMP students
Non-CSMP students
Reading

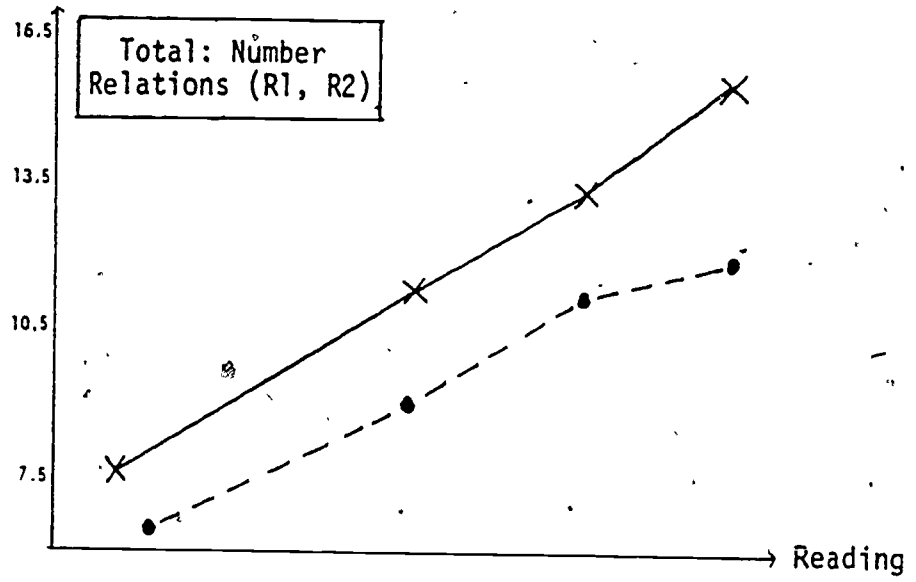

Total Mental Arithmetic (C3-C6)
27.5
22.5
17.5
12.5
Reading

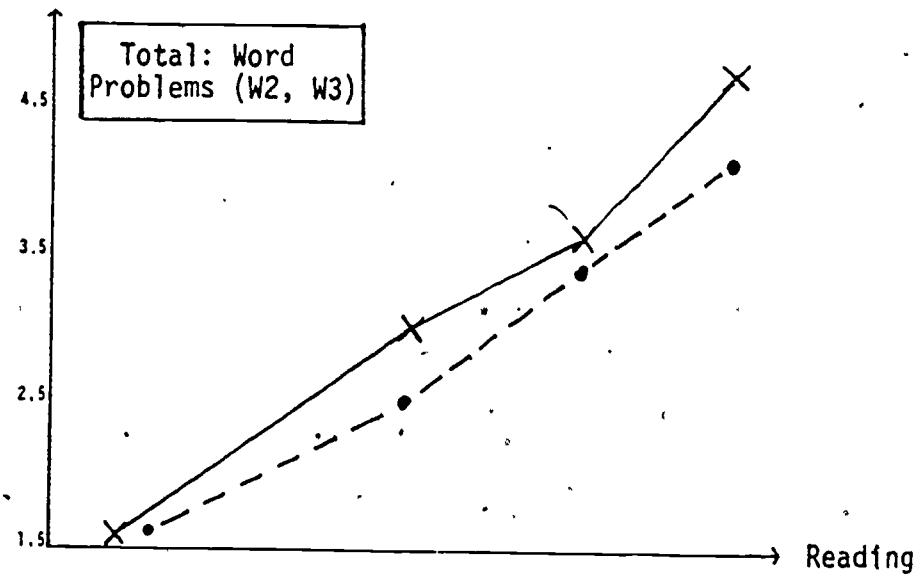

Total: Estimation (E2-E4, E6-E9)
35
31
27
23
Reading

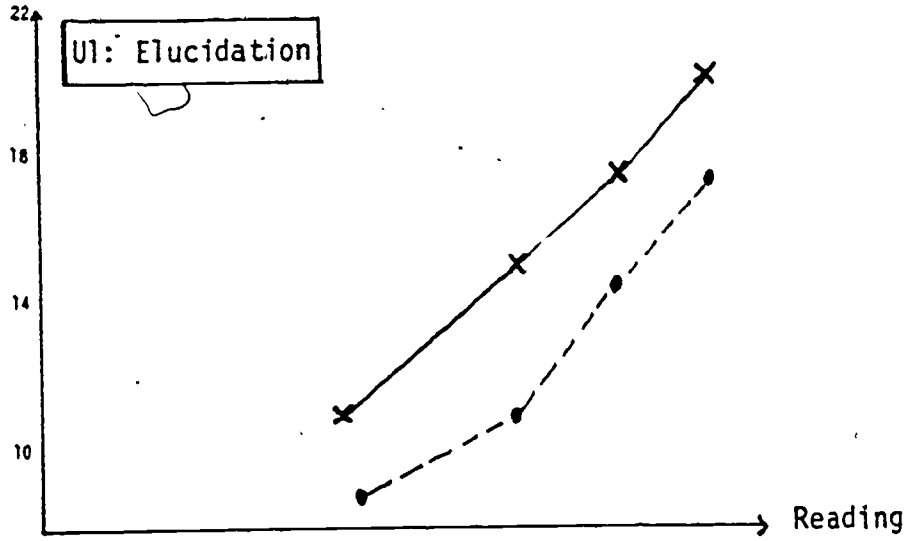

Total: Fractions
(N2, N5-N10)
44
36
28
20
Reading

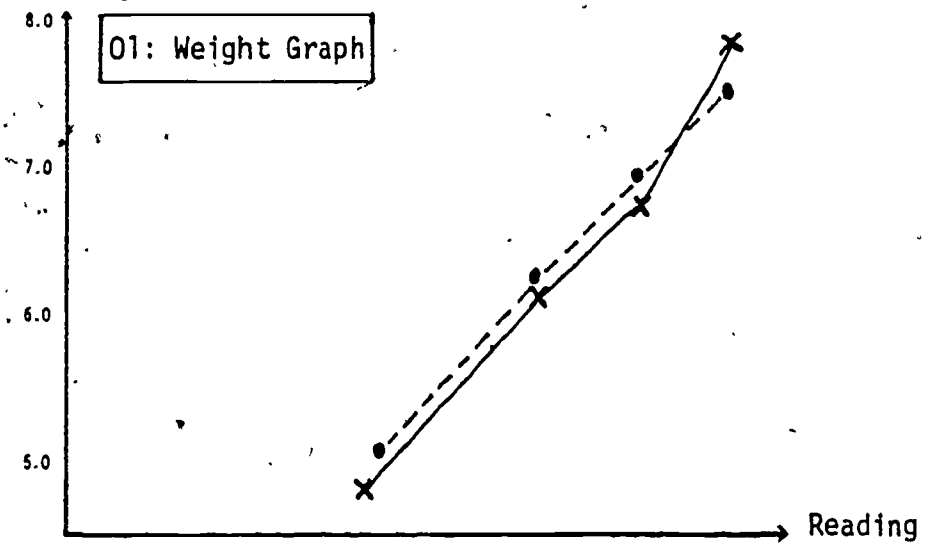

Total: Decimals
I(N1, N4)
15.5
12.5
9.5
6.5
Reading

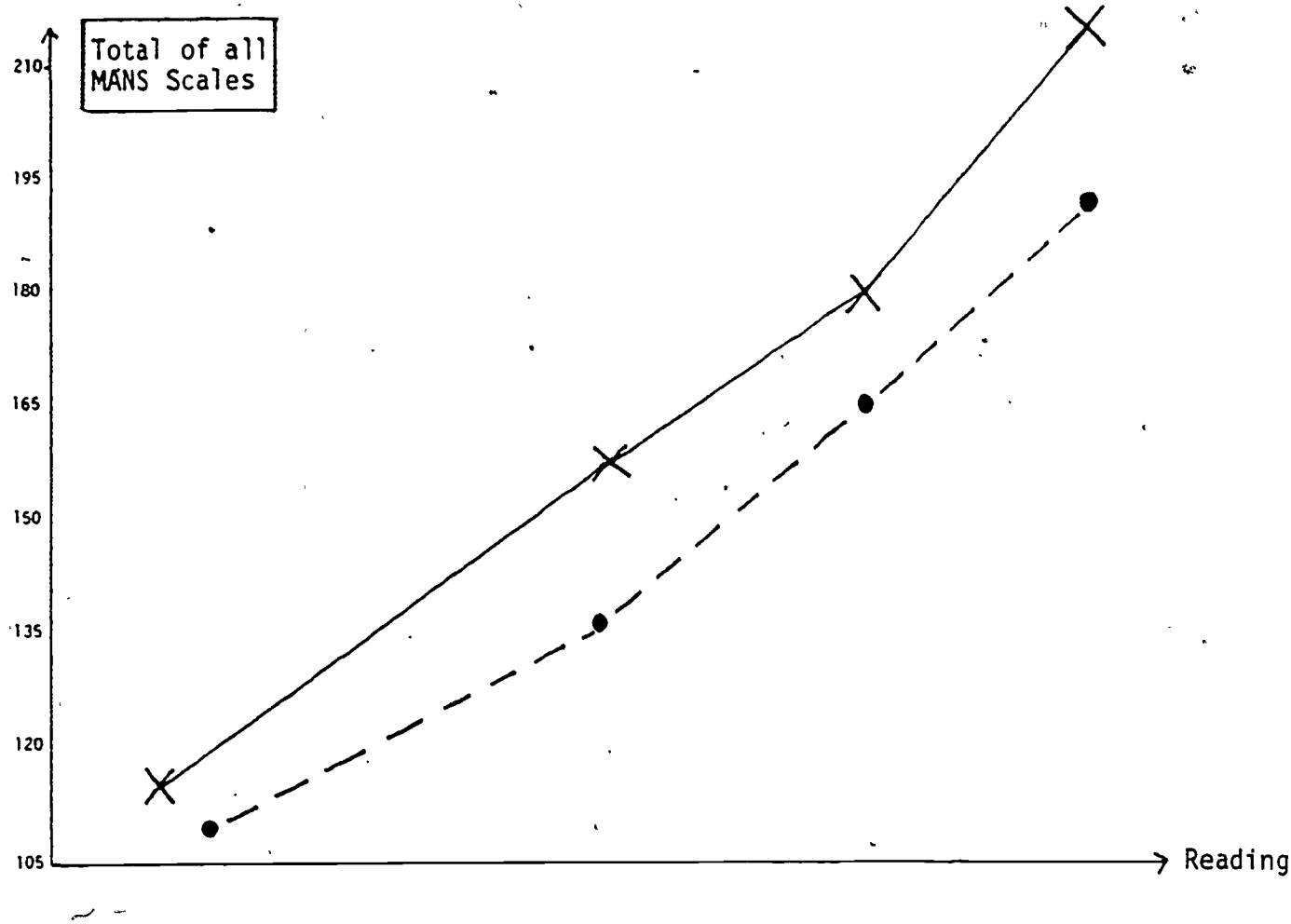

Total: Probability (P1, P3)
23.5
19.5
15.5
11.5
Reading


Total: Number Relations (R1, R2)
16.5
13.5
10.5
7.5
Reading


Total: Word Problems (W2, W3)
4.5
3.5
2.5
1.5
Reading


U1: Elucidation
22
18
14
10
Reading


O1: Weight Graph
8.0
7.0
6.0
5.0
Reading



From the graph above, for Total MANS, it can be seen that there was a smaller, though still clear, CSMP advantage at the lowest reading level than at other reading levels. This finding is not reflected on each of the other graphs, which seem to show one of two rather different things.

a) In most categories, the lowest level of readers in CSMP do quite as well, compared to their non-CSMP counterparts, as the other levels of readers.

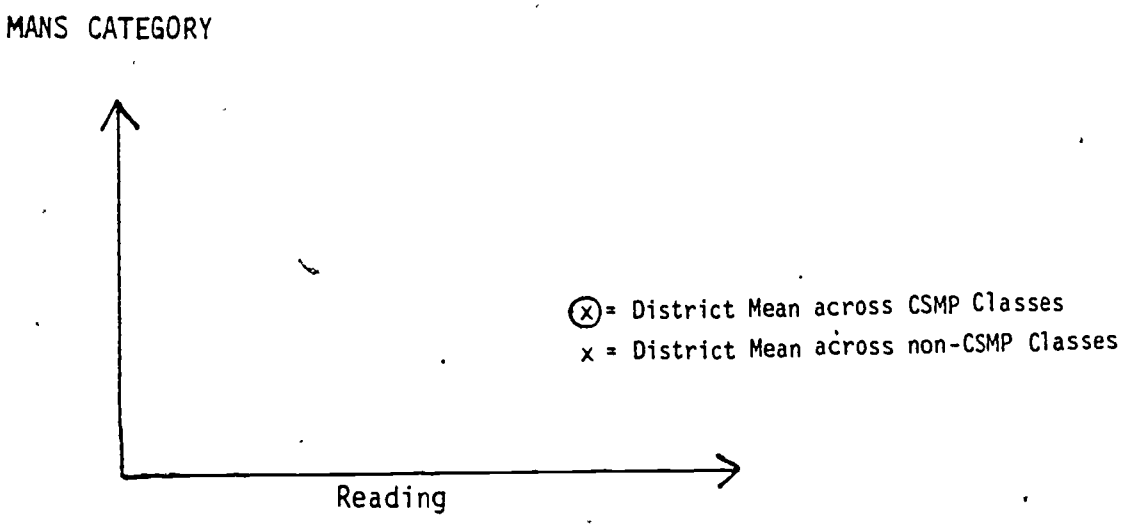
b) In four categories, the lowest level of readers in CSMP have scores virtually equal to their non-CSMP counterparts, while at other reading levels, CSMP students are doing better. This is true for Probability, Word Problems, Mental Arithmetic and Estimation. In the latter two categories however, the graphs indicate that this may result from outstanding performance by the low-reading non-CSMP students. (See the graphs of class means, pages A19, A25.)

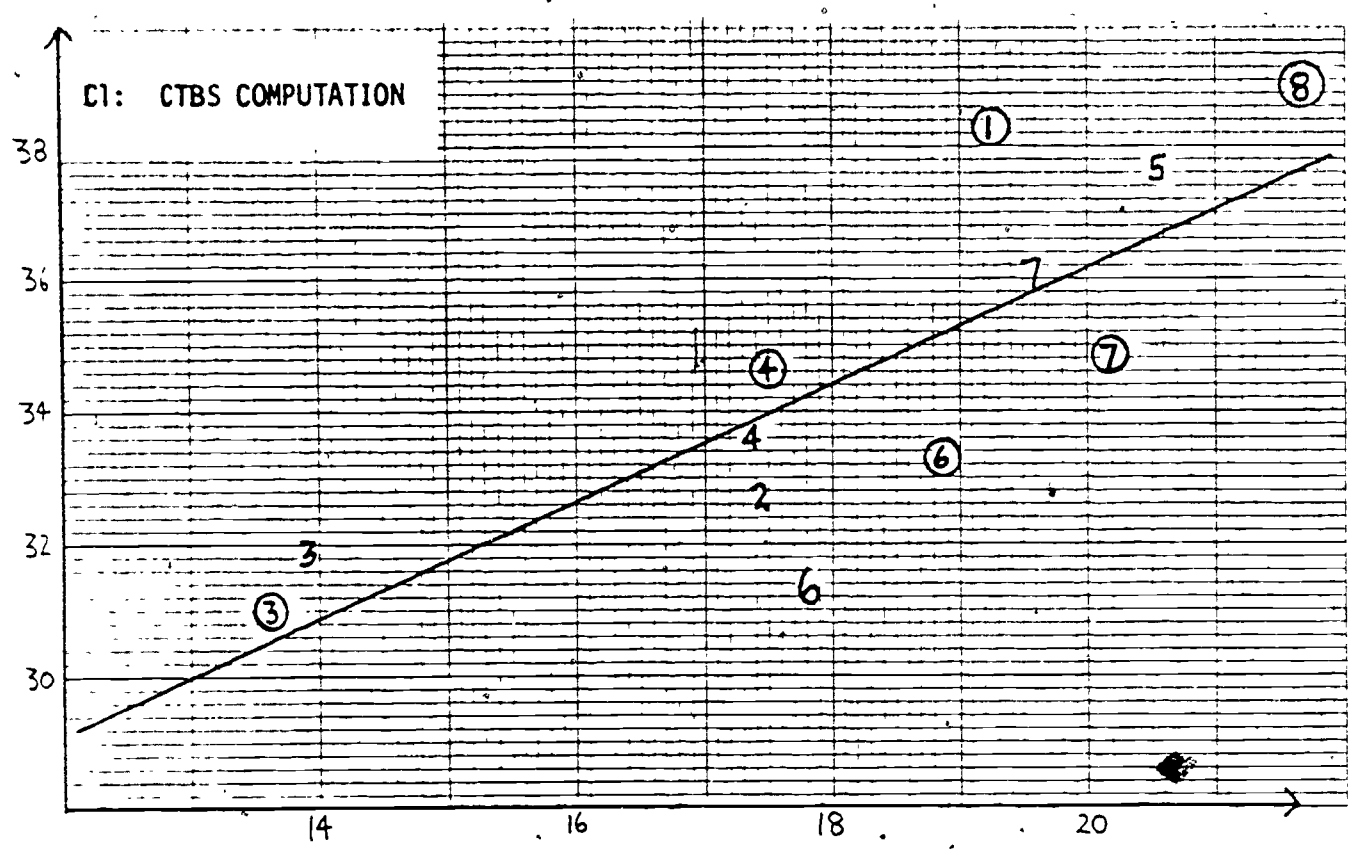
Graphs of District Means

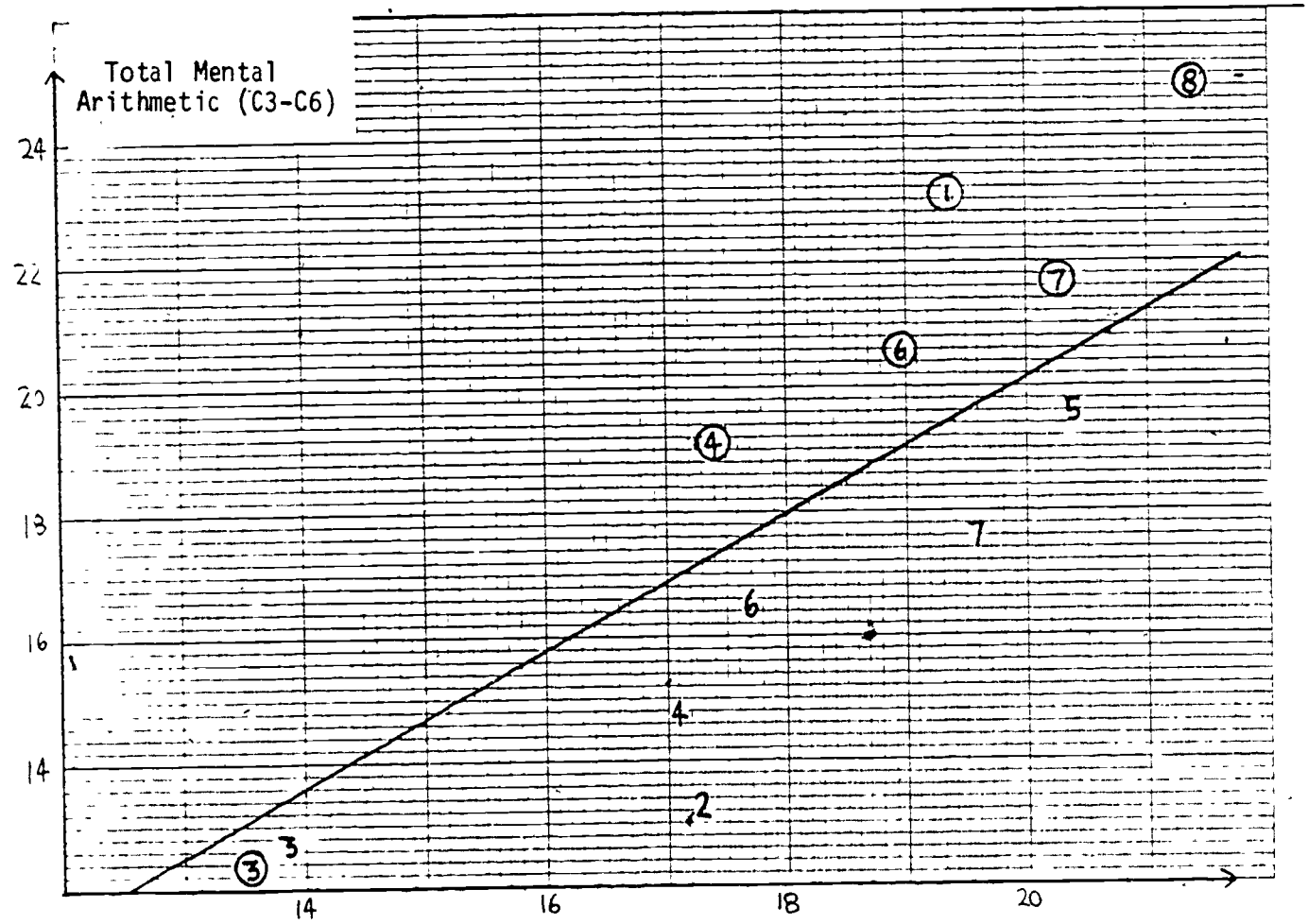
In the previous section, graphs were presented to show comparisons between CSMP and non-CSMP students according to the ability level of the students. In this section, a similar presentation is given, except the mean scores for the various districts are compared. It is the case that different methods of analysis (whether data is aggregated at the student, classroom, school, or district level - see Appendix B) produce very similar results. However the graphs presented in this section are probably much easier to interpret than the more detailed graphs of Appendix A and in fact they need little explanatory comment.

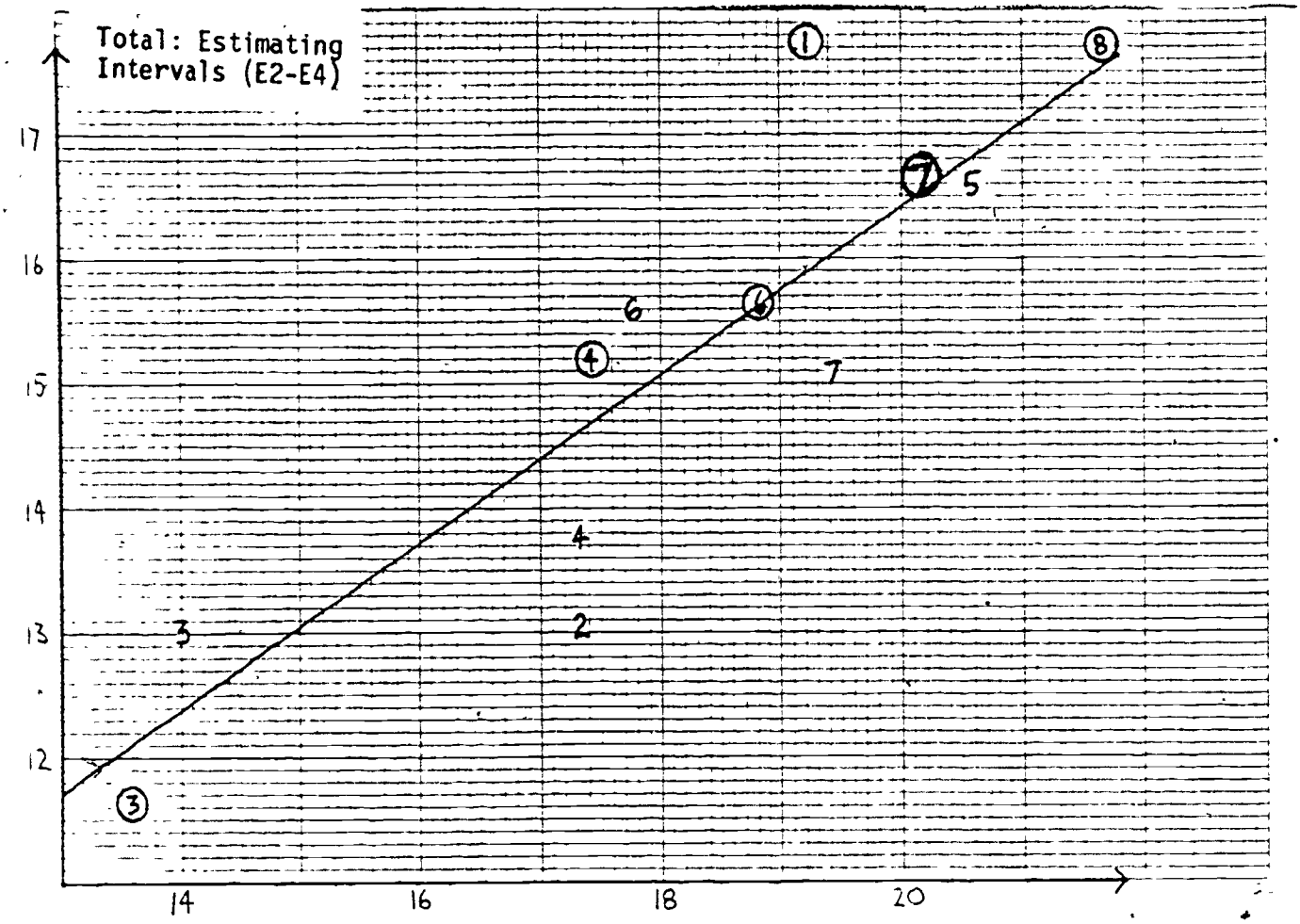
For each district, a CSMP and a non-CSMP mean were calculated for each category or grouping of scales. This was done by taking the means across classes. Then these sets of means were plotted against the corresponding mean reading score, and a regression line drawn for this set of district means.

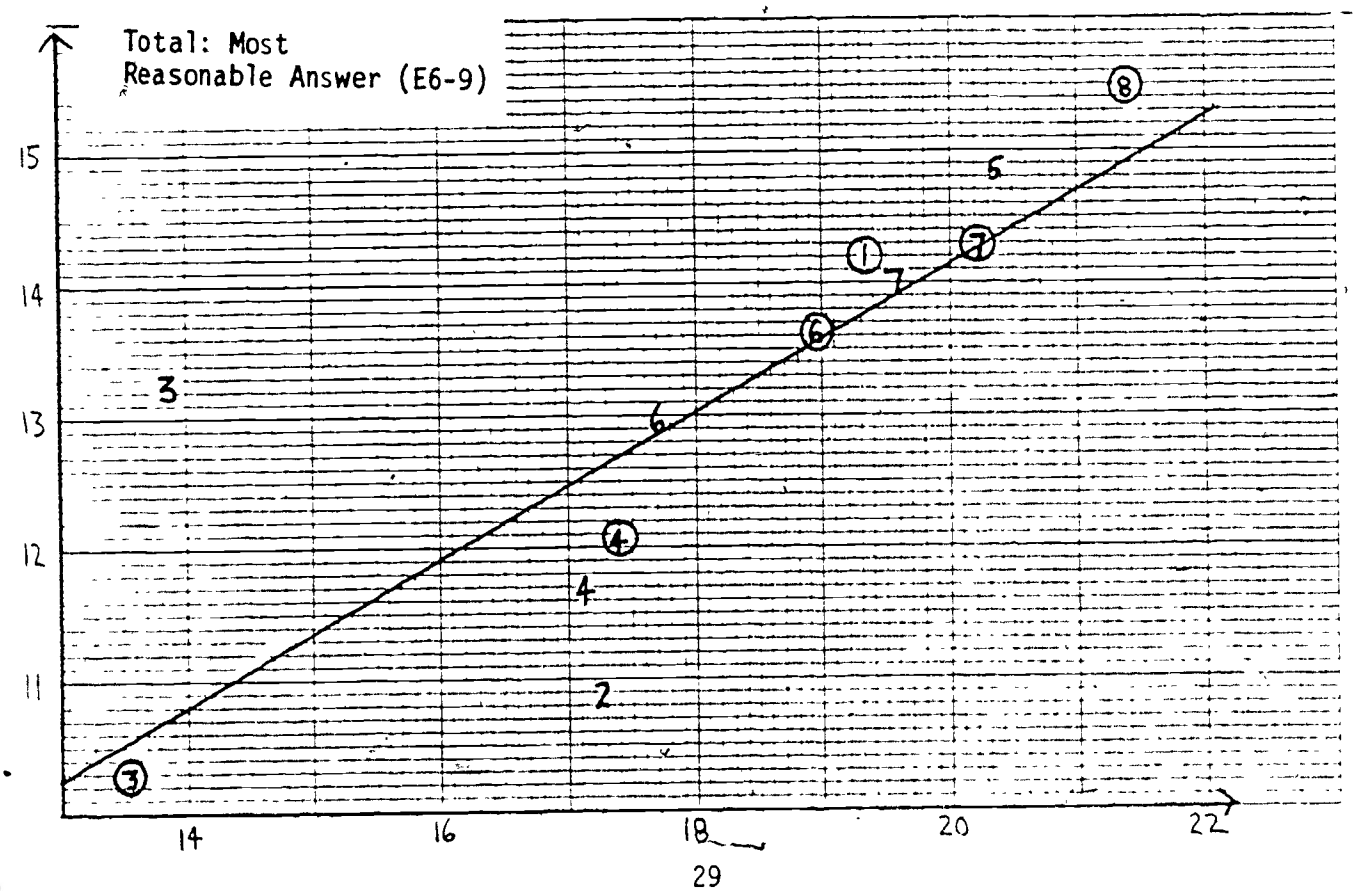
Each district mean is shown by a different numeral. Circled numerals stand for CSMP, uncircled numerals stand for non-CSMP. The graphs are in the form:

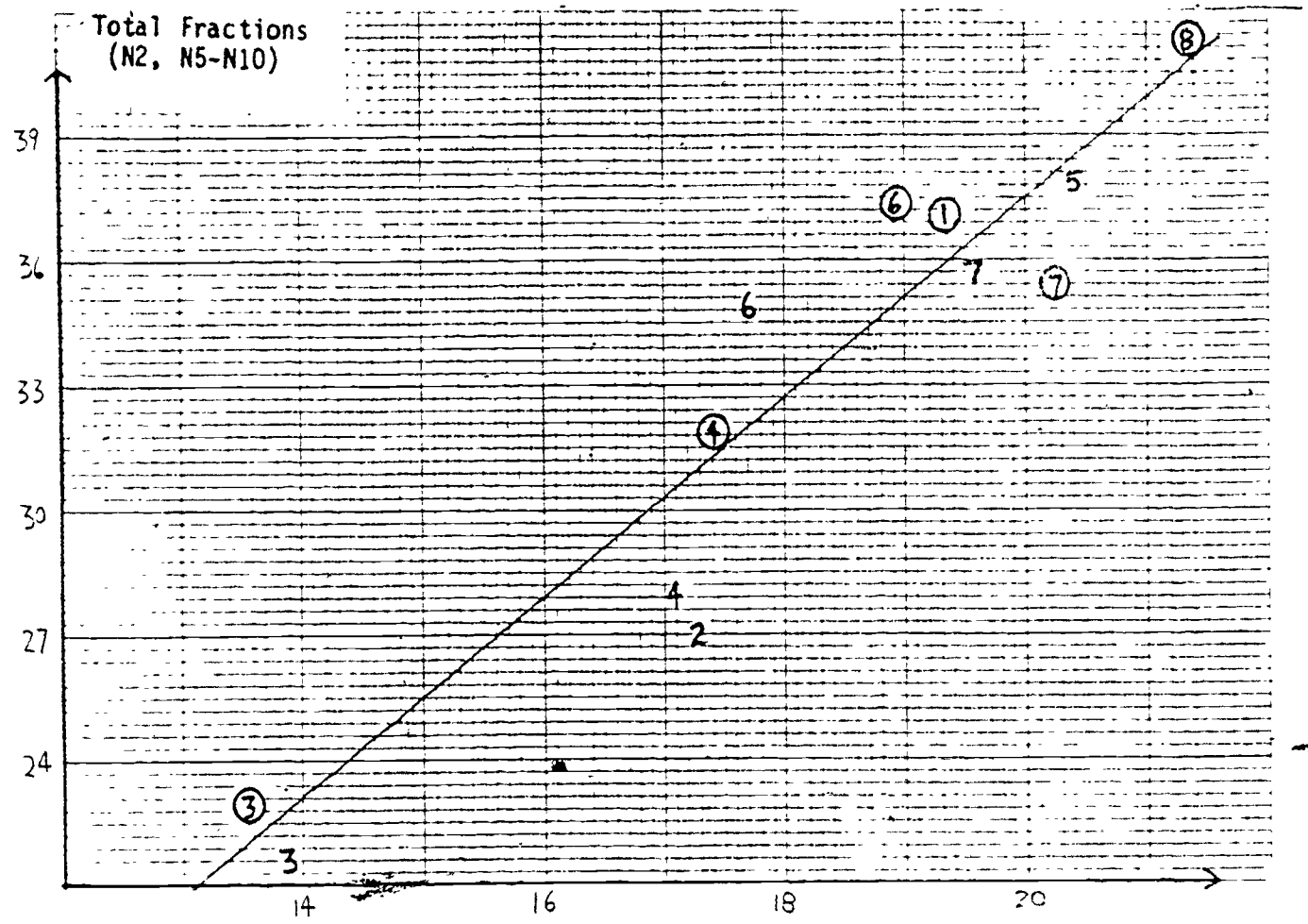


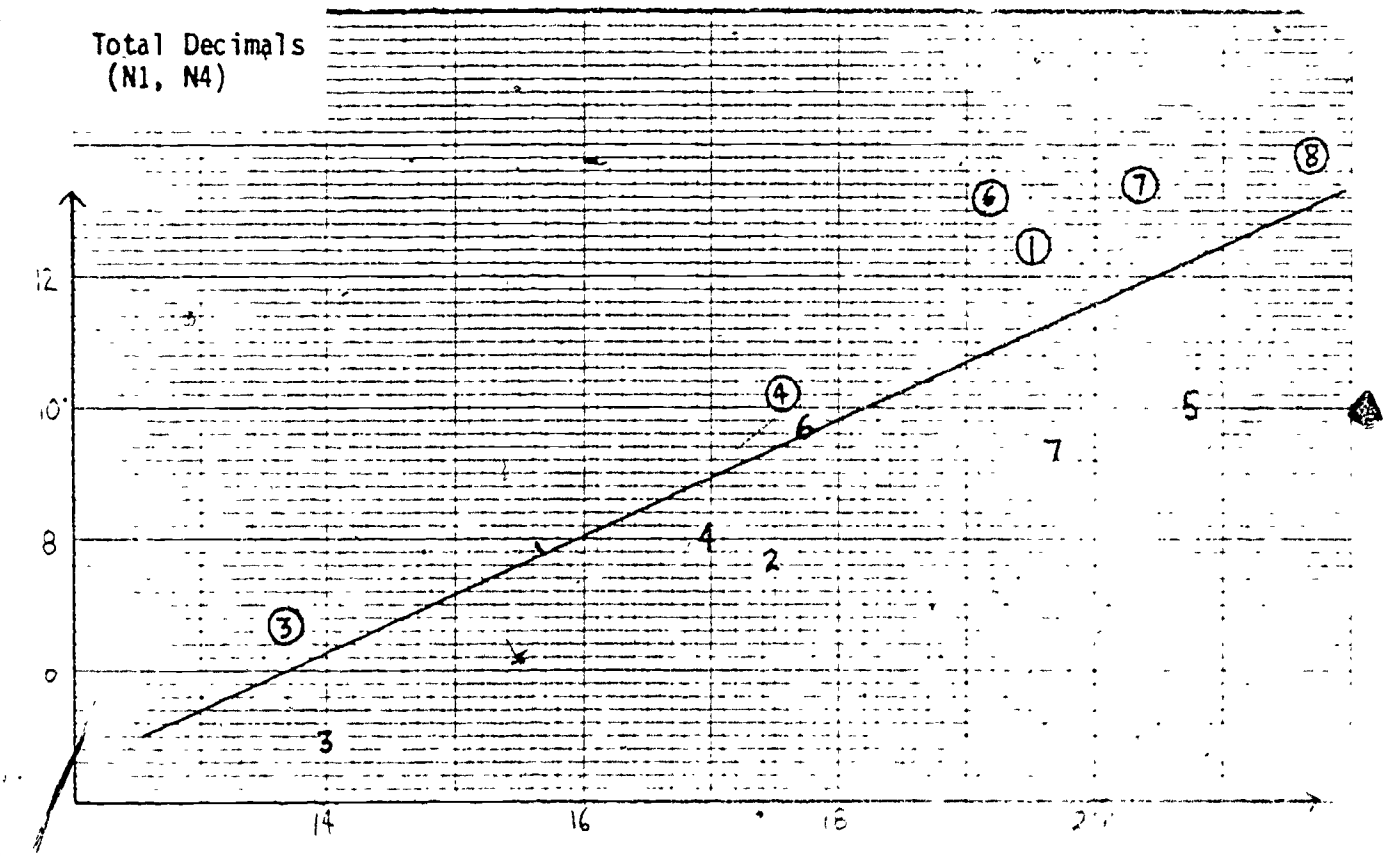

C1: CTBS COMPUTATION
38
36
34
32
30
14
16
18
20

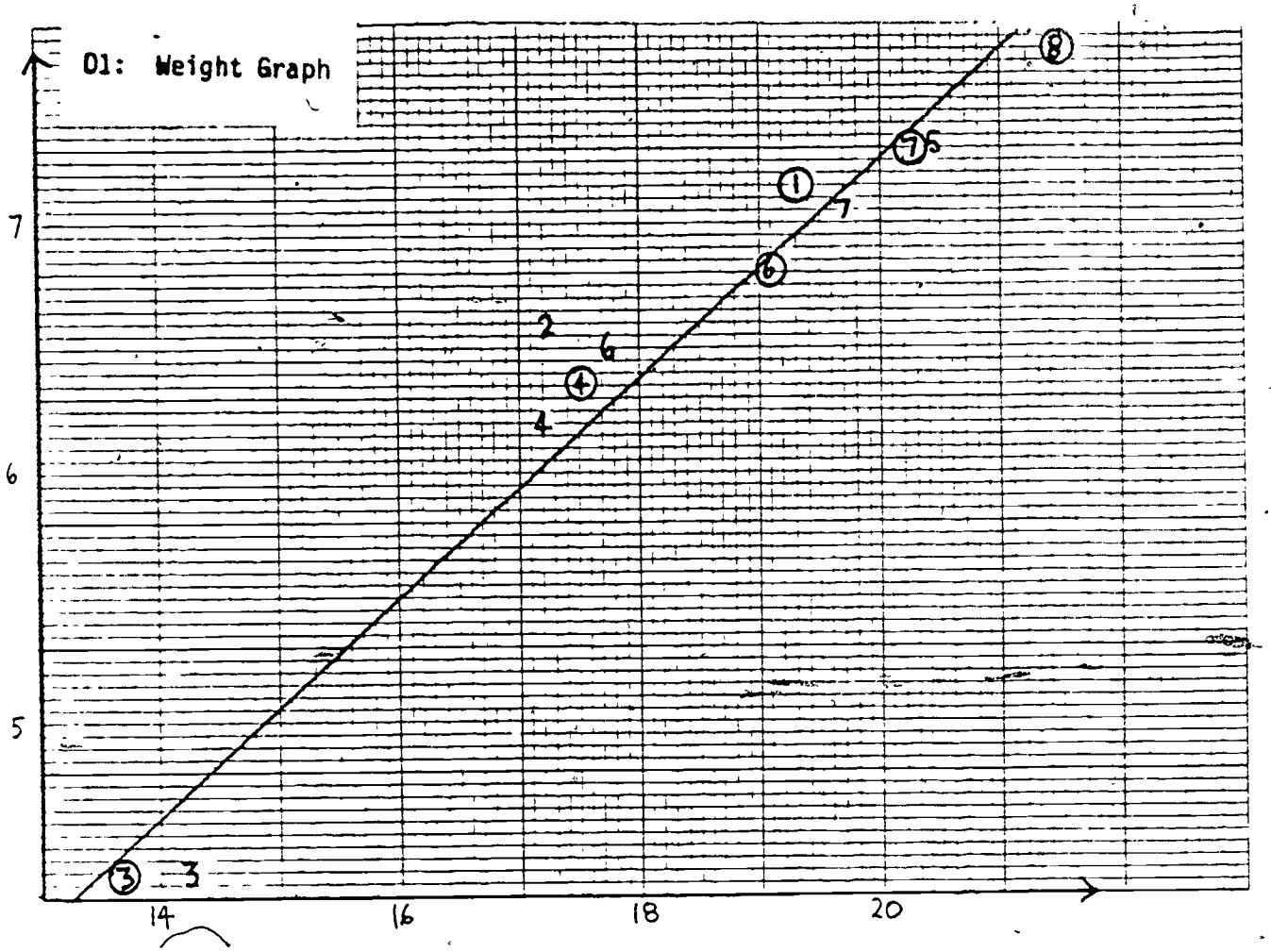

Total Mental
Arithmetic (C3-C6)
24
22
20
18
16
14
14
16
18
20

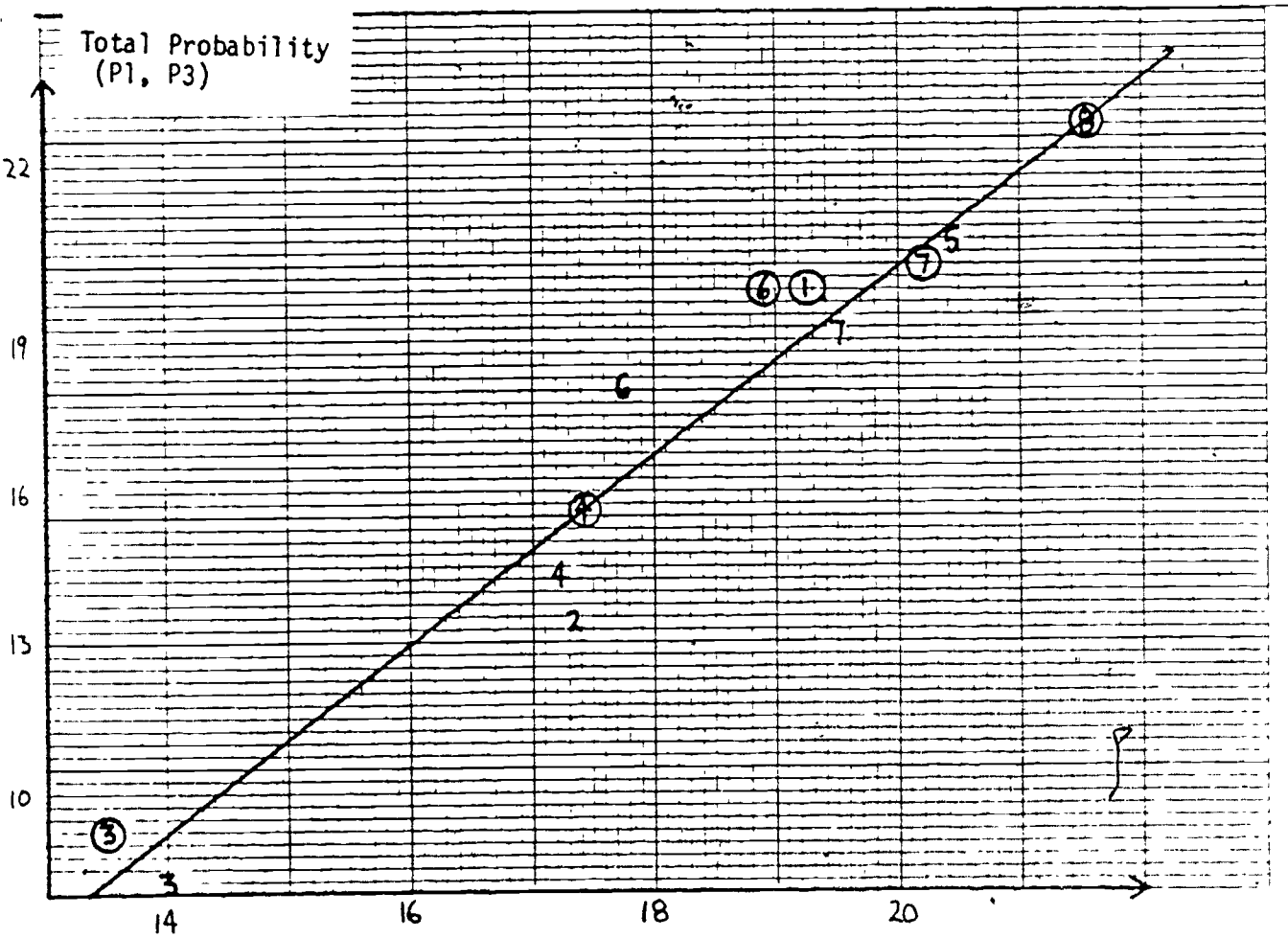

Total: Estimating Intervals (E2-E4)
17
16
15
14
13
12
14
16
18
20
1
8
7
5
6
4
7
4
2
3
3

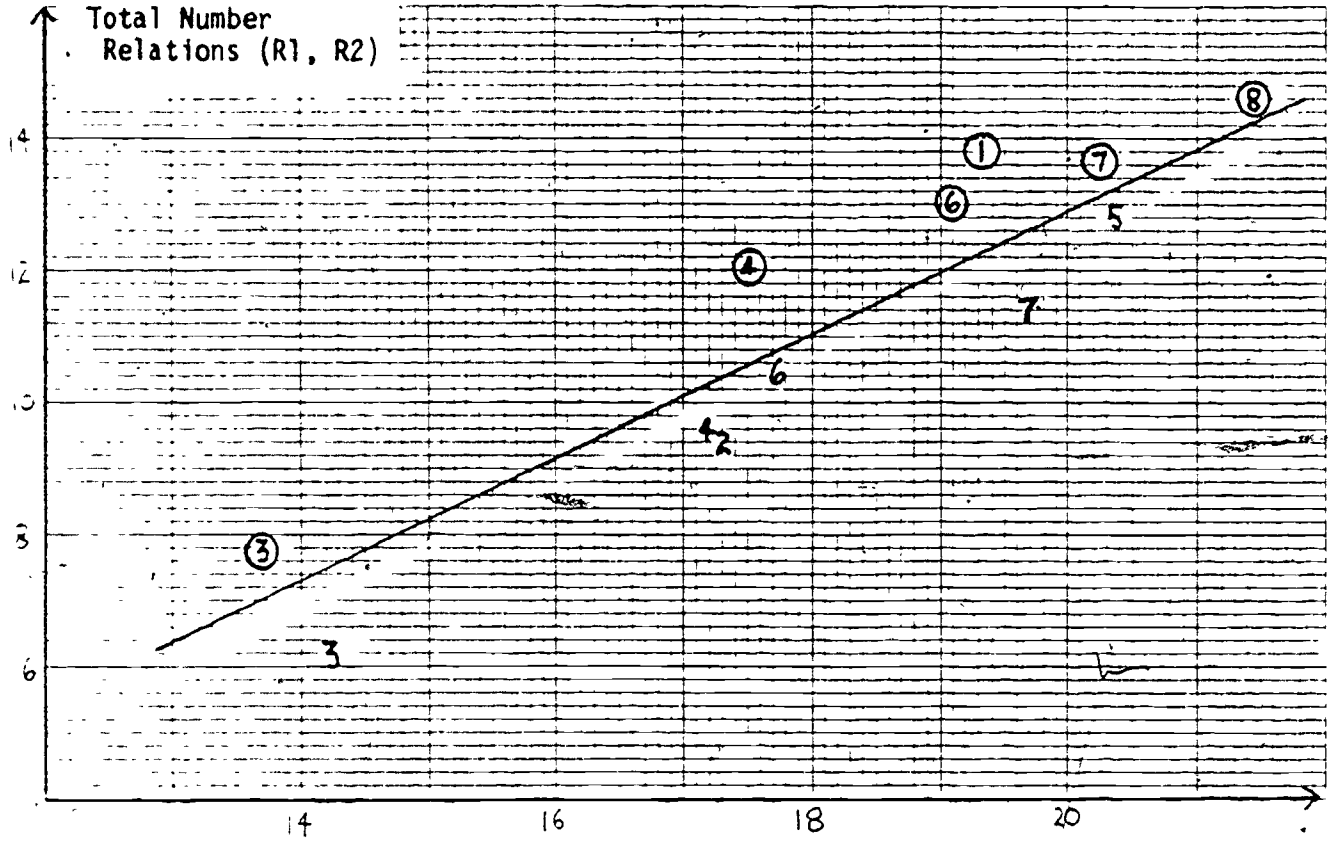

Total: Most Reasonable Answer (E6-9)
15
14
13
12
11
14
16
18
20
22
8
5
1
7
7
6
6
4
4
2
3

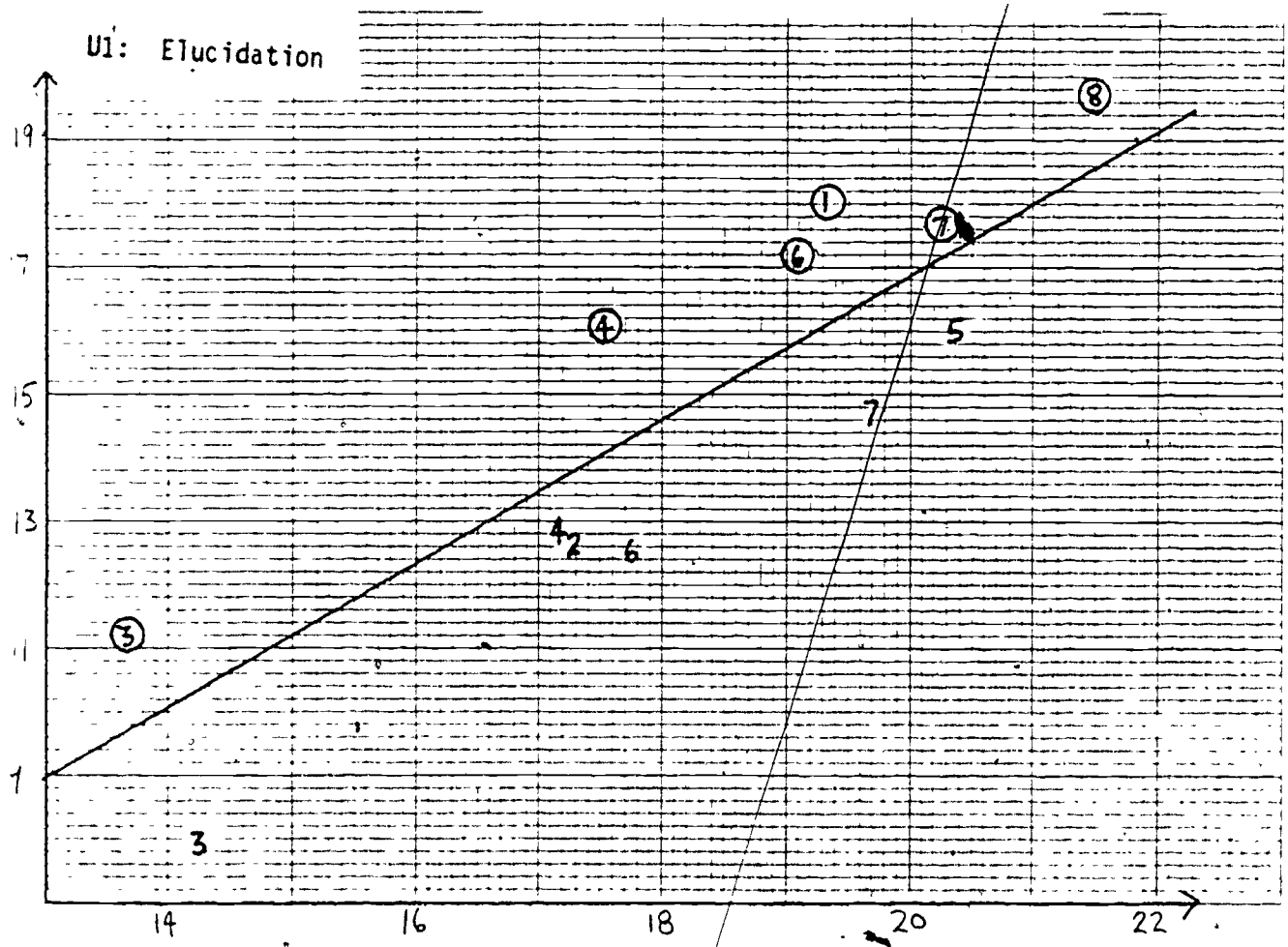

Total Fractions
(N2, N5-N10)
39
36
33
30
27
24
14
16
18
20


Total Decimals
(N1, N4)
12
10
8
0
14
16

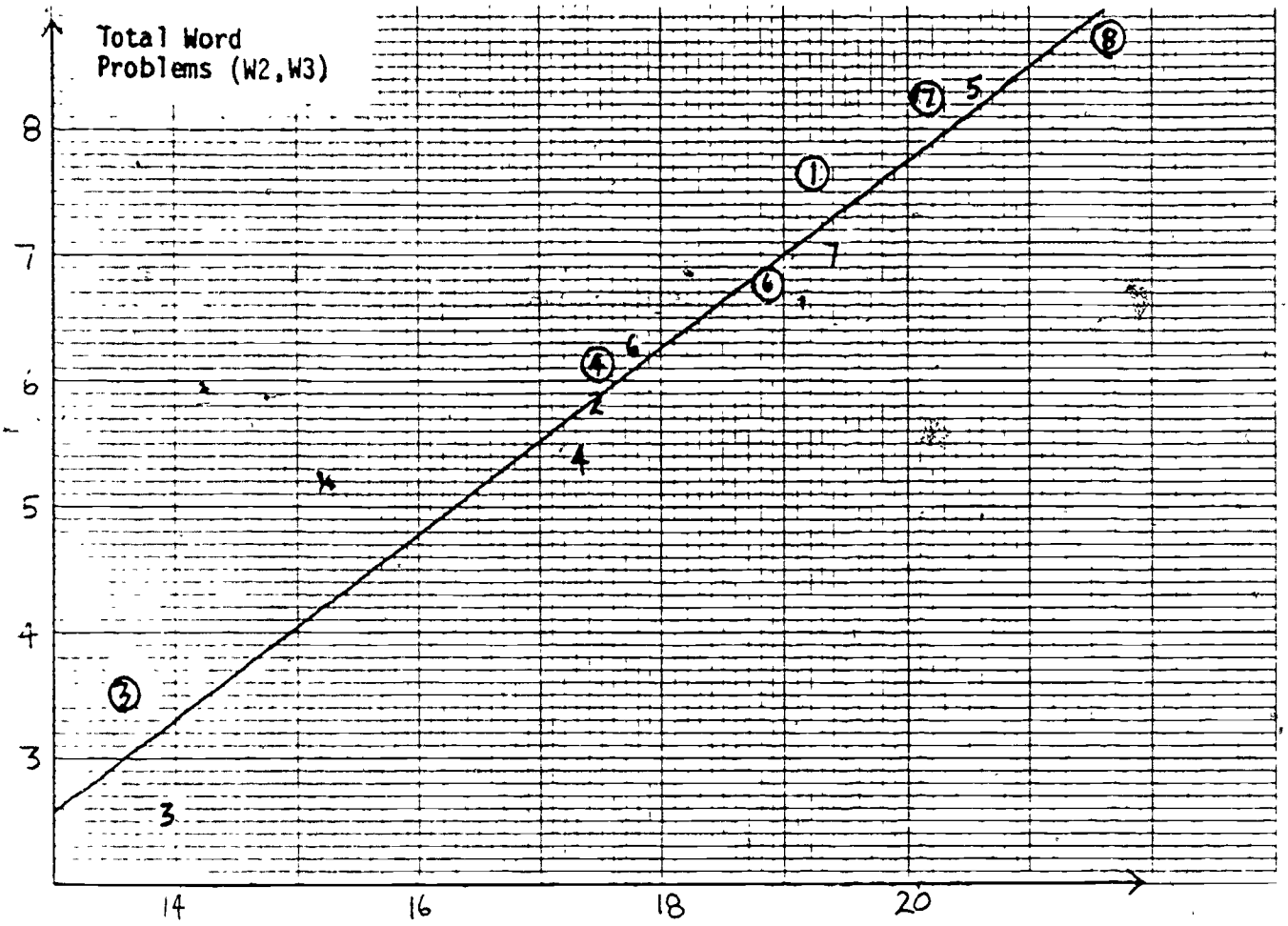

01: Weight Graph
7
6
5
14
16
18
20

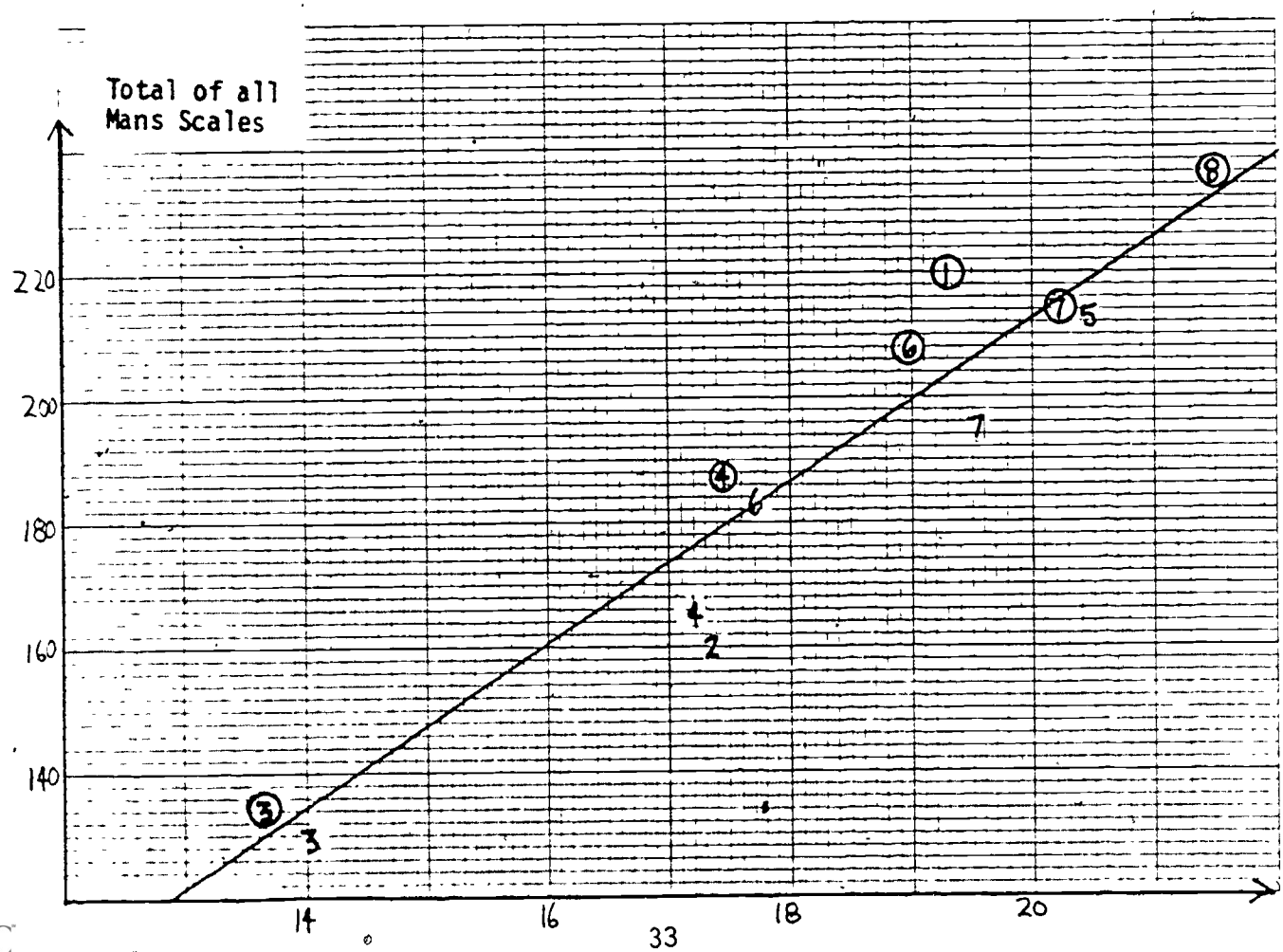

Total Probability
(P1, P3)
22
19
16
13
10
14
16
18
20


Total Number
Relations (R1, R2)
14
12
10
8
6
14
16
18
20
8
1
7
6
5
4
7
6
4 2
3
3


U1: Elucidation
19
17
15
13
11
9
14
16
18
20
22
8
1
7
6
4
5
7
4 2
6
3
3

35


Total Word
Problems (W2,W3)
8
7
6
5
4
3
14
16
18
20


Total of all
Mans Scales
220
200
180
160
140
14
16
18
20

The graphs reflect the numerical data from the analysis of class means; in particular they show dramatically how great the CSMP advantage is in the areas of Mental Arithmetic, Decimals, Number Relations and Elucidation.

New Students

Separate mean scores were calculated for two special groups of students:

New students, who moved or were transferred during the previous summer (these students were included in the various data in this report)

Late students, who moved to their new school after September 30 (these students were not included in the various data).

On the average, there were 1 or 2 new students per class and 1 late student. However, the distribution across classes was very uneven; for example, many classes had no new students, while others had 5 or more.

Table 3, below gives the mean scores for each of these groups for CSMP and for non-CSMP students.

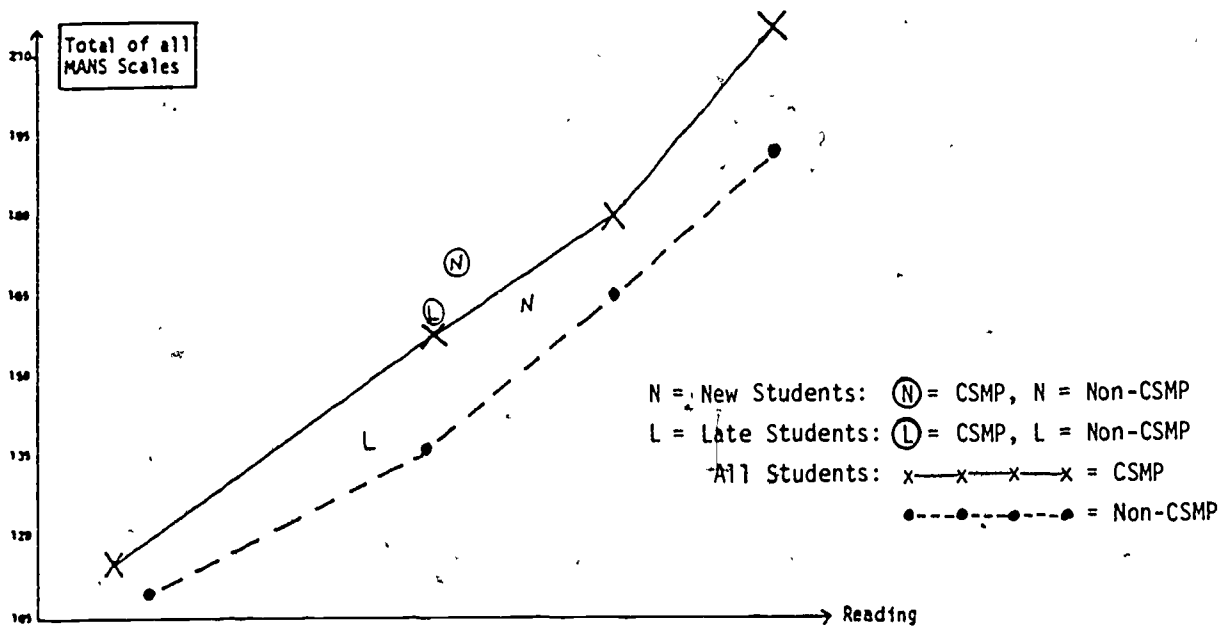
Table 3

Mean Scores, New and Late Students

Category	New Students		Late Students	
	CSMP	Non-CSMP	CSMP	Non-CSMP
CTBS Reading	16.7	18.1	16.8	15.2
Computation	34.8	33.6	31.1	31.9
Mental Arithmeiic	17.0	14.5	14.5	11.2
Estimating Intervals	14.8	13.7	12.2	11.5
Most Reasonable Answer	12.4	11.8	11.1	10.4
Measurement Estimation	1.6	1.9	1.6	1.8
Negative Numbers	5.6	5.5	4.8	4.1
Decimals	9.6	8.1	9.4	5.8
Fractions	28.0	28.2	27.6	23.9
Organizing Data	6.0	6.2	5.3	5.6
Probability	14.0	14.9	14.4	11.8
Number Relations	10.6	10.0	10.0	6.2
Eludication	14.4	12.7	11.8	8.6
Word Problems	5.8	5.9	5.3	5.3
Total	174.6	167.0	158.9	137.9
Number of Students	55	31	24	25

It is somewhat difficult to interpret this data. Clearly, CSMP students did better then their non-CSMP counterparts, and their advantage was greatest in Mental Arithmetic, Decimals and Elucidation, i.e. scales which also produced large differences in the original analyses.

Furthermore, if one plots the total MANS score against reading and superimposes this graph onto the graph on page 26 (the graph of mean scores by reading level for all students) one gets the graph pictured below.



Surprisingly, each of the groups had higher scores than would have been expected from the graphs of all students, though the CSMP advantage remained. Whether or not the same results would have occurred at various ability levels as well was not investigated because the numbers of students was too small to be subdivided in this way. Nevertheless, this data, tentative as it is, does not support the view that students transferring to a new school at the beginning of school, or even later in the year, suffer in their performance; this finding is true for both CSMP and non-CSMP students.

Appendix A

Scale-by-Scale Statistical Analysis

This Appendix contains information on each MANS scale. There are two kinds of information given. First, the actual test items and various item statistics are given for each scale (see page A2). Then, for each category or grouping of similar scales, a graph of class means is given for reading versus the total score on that category (see page A3). The scales appear in alphabetical order by scale category as shown below with their page number in this Appendix.

Item Statistics

Test Items	Percent Correct	
	CSMP	Non-CSMP
(1)	(2)	
Number of Students	(3)	
KR 20 Reliability	(4)	
Reading Score	(5)	
Correlation: Reading Versus Scale	(6)	

(7) Means by Ability Level:

	1	2	3	4	All
CSMP					
Non-CSMP					
t-test					

1 The test items listed here are shown in a greatly compressed and sometimes altered form so that what required a full page on the student copy of the test can be sqeezed into this space.

2 These are the percentages of students getting the item correct.

3 The number of CSMP students will usually be about 630 except when the scale was taken by a random half of each class. For non-CSMP, it is usually about 500.

4 The KR 20 reliability coefficient is a measure of the degree to which the items in a scale are testing a single underlying ability. A large KR 20 (above say, .8) means a high correlation among the items; a low KR 20 (below say .5) means a low correlation among the items and not a single underlying ability.

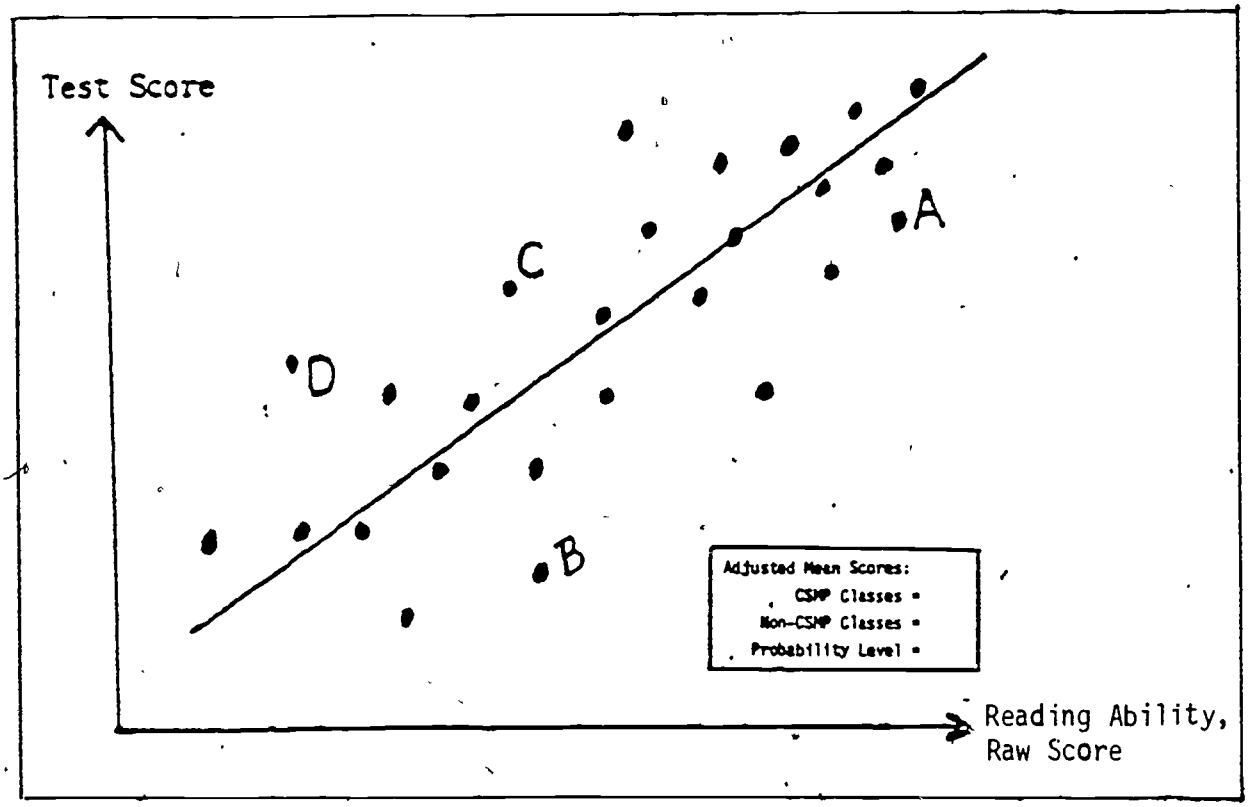
5 The mean raw score on the sampled items of the CTBS Reading Comprehension Test for those students who took this particular scale.

6 The correlation between scores on the reading test and scores on this particular MANS test.

7 Students were divided into four quartiles according to their reading score, with roughly equal numbers of students in each quartile. In each quartile an Analysis of Covariance across students was employed. In the table, for each of CSMP and non-CSMP the adjusted means are given (taking into account differences in reading scores, which adjustments were always small because of the restricted range of the reading scores). Also given below this pair of mean scores, in the third row of the table, is the resulting t-statistic with degrees of freedom in the hundreds. A rough rule of thumb would be to consider t-values above 2 to be significant. This is aggregation by students rather than classes, so this is a much more liberal test of differences. Positive t-values indicate a CSMP advantage, negative t-values a non-CSMP advantage.

Graphs of Class Means

The right hand pages are for the graphs of class means. In the hypothetical graph below, each class is represented by a dot whose location is determined by the average reading scores (horizontal axis) and MANS Test score (vertical axis) for the students in that class.



Based on this set of class means thus graphed, the regression line has been drawn. This line is the best linear prediction of mean class test score that can be made from knowing the ability level of the class. Note in the example that classes A and B fall well below the regression line, or are scoring well below what would be predicted for them knowing the ability level of the class, while classes C and D fall well above the regression line. Note also that, although class A had a slightly higher mean score on the test than did class C, class C did much better given relative ability scores of the two classes. When the class means generally fall close to the regression line, test scores are well predicted by the covariate; when they are more dispersed from this line, the covariate is a less effective predictor.

In the box in the lower right hand corner, the mean scores across CSMP and across non-CSMP classes, adjusted for reading ability, are also given, together with the p-value obtained from the F-test. (The p-value is the probability that a difference in mean scores that large could have occurred by chance alone. Hence, the smaller the p-value (especially below, say, .05), the more likely it is that there are "real" differences between CSMP and non-CSMP classes).

In the actual graphs, CSMP classes are represented by "solid" symbols, non-CSMP by corresponding "empty" symbols. (See page for key.)

C1a) Standardized Computation-Addition (2 forms)

Form 1 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP	Biserial	Form 2 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
36 + 29	96	96	50,04	46 + 5	96	96
179 + 430 + 245	80	71	29,42	$13 + 2\frac{3}{4}$	90	89
36.418 4.893 25.153 + 8.030	76	67	50,33	2.713 49 3.574 + 2.020	77	80
$\frac{1}{2} + \frac{1}{2}$	71	61	34,24	346 159 + 350	84	79
0 64 44 56 5 89 + 6.25	75	66	48,53	$12\frac{1}{3}$ + $4\frac{1}{4}$	37	47
$20 00 0 75 4 00 + 1 25	92	84	52,56	30 4 + 6 5	89	88
Number of Students	315	263		Number of Students	315	256
KR20 Reliability	.52	.51		KR20 Reliability	.33	.29
Mean Reading Score	18.1	18.2		Mean Reading Score	18.0	17.8
Correlation, Reading versus Scale	.37	.38		Correlation, Reading versus Scale	.26	.38

Across students, regardless of form:

Means By Ability Level

	1	2	3	4	All
CSMP	4.3	4.6	5.0	5.2	4.8
non CSMP	4.1	4.4	4.8	5.1	4.6
t-test	2.0	1.3	2.1	1.2	2.4

1. Together with the other C1 scales (b, c, and d), each of Forms 1 and 2 constituted a set of 24 items (half the items of the CTBS, Level 2, Form S, Computation Test). This set of items was preceded by a sample item, as per the CTBS directions, and with a time limit of 20 minutes. All items were multiple choice (not shown).

C1b) Standardized Computation - Subtraction (2 forms)

Form 1 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP	Biserial	Form 2 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
490 - 130	91	90	32,64	647 - 159	77	76
5,681 - 796	83	78	56,48	648 - 105	89	91
6,341 - 457	76	73	66,52	7,605 - 4,327	70	70
$\frac{1}{3} - \frac{1}{3}$	93	71	33,25	$\frac{3}{7} - \frac{1}{7}$	70	77
36 2 - 3 6	69	59	39,44	$35\frac{2}{3}$ - 12	88	88
$25 00 - 1 75	64	66	51,72	48 3 - 4 8	71	68
Number of Students	315	256		Number of Students	315	256
KR20 Reliability	.59	.63		KR20 Reliability	.57	.61
Mean Reading Score	18.1	18.2		Mean Reading Score	18.0	17.8
Correlation, Reading versus Scale	.44	.43		Correlation, Reading versus Scale	.35	.36

Means By Ability Level

	1	2	3	4	All
CSMP	3.8	4.5	4.9	5.2	4.6
non CSMP	3.7	4.3	4.8	5.2	4.5
t-test	0.8	1.0	1.0	-0.1	1.3

1. See Note 1. for C1a).

C1c) Standardized Computation - Multiplication (2 forms)

Form 1 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP	Biserial	Form 2 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
$\begin{array}{r} 300 \\ \times\ 3 \end{array}$	92	90	54,48	$\begin{array}{r} 25 \\ \times\ 4 \end{array}$	96	91
$\begin{array}{r} 706 \\ \times\ 8 \end{array}$	83	85	52,65	$\begin{array}{r} 3,057 \\ \times\ 6 \end{array}$	75	79
$\begin{array}{r} 237 \\ \times\ 506 \end{array}$	59	63	51,40	$\begin{array}{r} 33 \\ \times\ 24 \end{array}$	77	78
$\frac{3}{5} \times \frac{7}{8}$	71	77	44,44	$\frac{1}{4} \times \frac{1}{4}$	54	45
$8 \times \frac{1}{2}$	64	43	47,24	6.68×9	61	63
$\begin{array}{r} 7.45 \\ \times\ 9 \end{array}$	69	67	62,56	$\begin{array}{r} \$13.30 \\ \times\ 12 \end{array}$	62	66
Number of Students	315	263		Number of Students	315	256
KR20 Reliability	.64	.58		KR20 Reliability	.58	.48
Mean Reading Score	18.1	18.2		Mean Reading Score	18.0	17.8
Correlation, Reading versus Scale	.41	.46		Correlation, Reading versus Scale	.32	.26

Means By Ability Level

	1	2	3	4	All
CSMP	3.6	4.0	4.5	4.9	4.3
non-CSMP	3.5	4.0	4.5	4.9	4.3
t-test	0.4	0.3	0.2	0.1	0.3

1. See Note 1. for C1a).

C1d) Standardized Computation - Division (2 forms)

Form 1 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP	Biserial
$28 \div 7$	84	84	60,46
$5\overline{)550}$	78	81	70,54
$6\overline{)12{,}000}$	73	79	56,48
$330 - 5$	77	78	68.65
$4 \div \frac{1}{2}$	17	17	17,10
$21\overline{)1008}$	46	39	33,30
Number of Students	315	263	
KR20 Reliability	.62	.55	
Mean Reading Score	18.1	18.2	
Correlation, Reading versus Scale	.40	.45	

Form 2 Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
$7\overline{)427}$	78	83
$9\overline{)183}$	73	76
$39\overline{)3{,}370}$	39	39
$\frac{1}{5} \div \frac{1}{5}$	52	56
$\$3.00\overline{)\$24.00}$	71	59
$\$12.00 \div 4$	80	77
Number of Students	315	256
KR20 Reliability	.53	.57
Mean Reading Score	18.0	17.8
Correlation, Reading versus Scale	.46	.43

Means By Ability Level

	1	2	3	4	All
CSMP	3.0	3.6	4.0	4.6	3.8
non-CSMP	2.9	3.4	4.2	4.5	3.8
t-test	0.4	1.0	-1.1	0.5	0.2

1. See Note 1. for C1a).

Scale C1

Notes:

1. In each class, half the students took one set of 24 items, the other half took the other set of 24 items. To calculate the class means on the total of 48 items of this test (the CTBS Level 2 Form S Computation Test), the means across the two 24 items sets were added together. The graph on the facing page shows the distribution of class means; also given are the adjusted means across CSMP and non-CSMP classes, and the p-value of the ANCOVA F-test.

2. A similar procedure was used for the 12-item subsets (half of each class taking a 6-item set) for addition, subtraction, multiplication and division. The adjusted means and p-value for these subsets were as follows:

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
Addition	9.6	9.3	.07
Subtraction	9.2	9.0	.53
Multiplication	8.5	8.4	.68
Division	7.6	7.6	.95

3. The scatter plot on the facing page clearly indicates that, not only is there very little difference between CSMP classes (solid figures) and non-CSMP classes (empty figures) but there is also not a particularly strong relationship between reading and computation scores; the classes are widely dispersed from the regression line. Across students the correlation between reading score and the various 6-item sets of items were only in the .3's and .4's' the correlations between the reading and 24-item sets were about .5.

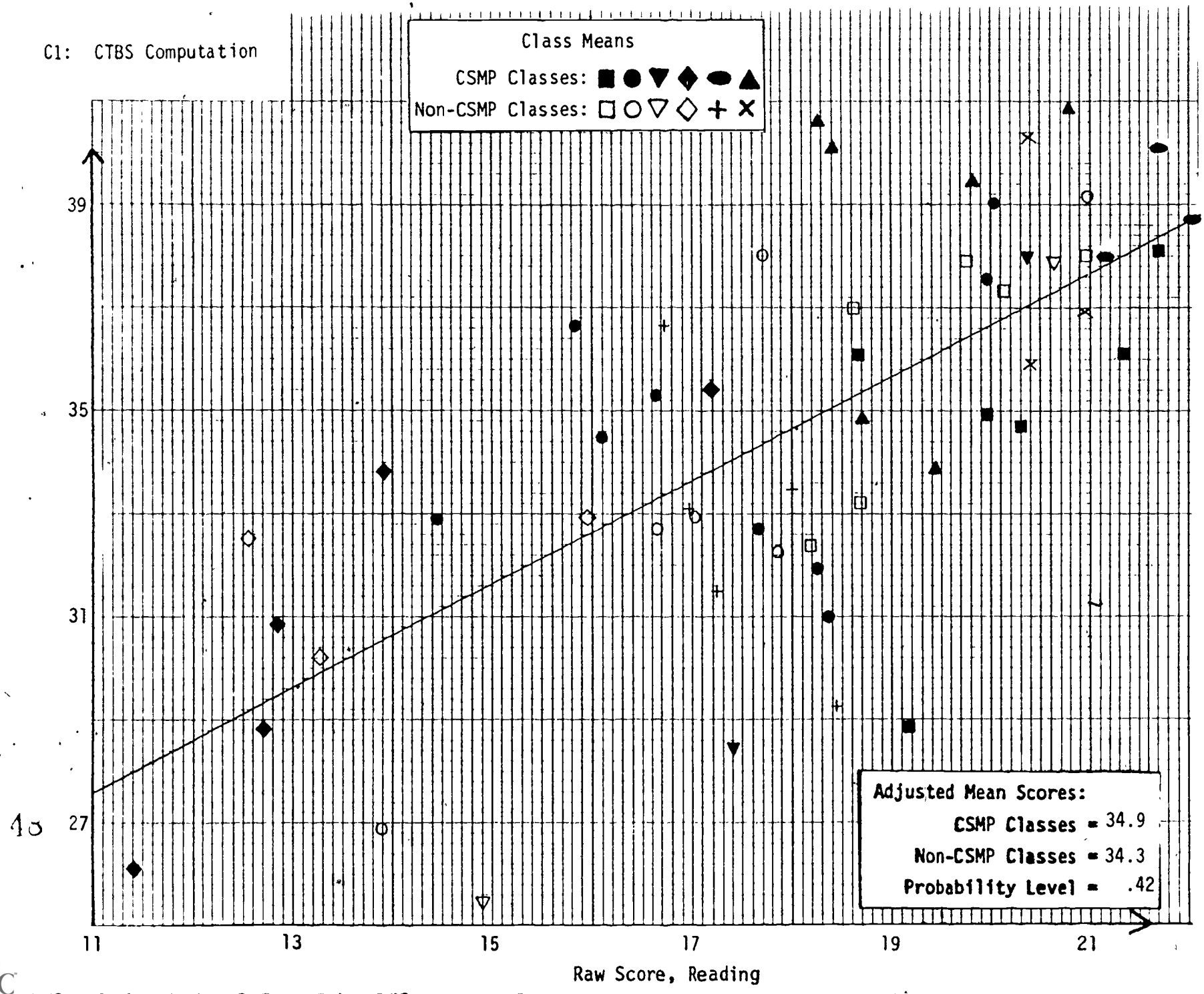
4. On most indivdual items, there was very little difference in percent correct by CSMP students versus non-CSMP students. CSMP students had a slight advantage on the 12 items involving decimals (a mean score of 8.5 versus 8.0) and on some of the column addition items. On items involving fractions, CSMP students did a little better on one type:

$$(8 \times \tfrac{1}{2} \qquad \tfrac{1}{4} \times \tfrac{1}{4} \qquad \tfrac{1}{3} - \tfrac{1}{3} \qquad \tfrac{1}{2} + \tfrac{1}{2},$$

non-CSMP on another:

$$(\tfrac{3}{5} \times \tfrac{7}{8} \qquad \tfrac{3}{7} - \tfrac{1}{7} \qquad 12\tfrac{1}{3} + 4\tfrac{1}{4}).$$

C1: CTBS Computation



C3 Mental Arithmetic - Addition

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
501 + 501 + 501 + 501 = ☐	78	73
☐ + 125 = 250	83	70
9,001 + ☐ = 9,100	71	58
125 + 125 + 225 + 225 = ☐	63	58
4,999 + ☐ = 10,000	47	30

Means By Ability Level

	1	2	3	4	All
CSMP	2.6	3.3	3.5	4.1	3.4
non-CSMP	2.3	2.7	2.9	3.6	2.9
t-test	1.4	2.5	2.5	3.2	4.7

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.55	.59
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.42	.34

1. Students were not allowed to do paper and pencil calculations; but had to do the calculations in their head and write down only the final answer.

C4 Mental Arithmetic - Subtraction

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
459 — 359 = ☐	85	85
7,001 — 6,999 = ☐	50	40
1,000 — 5 = ☐	78	61
700 — 401 = ☐	56	44
☐ — 250 = 150	39	26

Means By Ability Level

	1	2	3	4	All
CSMP	1.8	2.9	3.3	3.9	3.1
non-CSMP	1.7	2.2	2.9	3.2	2.6
t-test	0.3	3.1	2.1	3.9	4.7

	CSMP	Non-CSMP
Number of Students	337	276
KR20 Reliability	.66	.64
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.54	.42

1. See Note 1. for Scale C3

C5 Mental Arithmetic - Multiplication

Test Items	Percent Correct CSMP	Non-CSMP
$7 \times 30 = \square$	83	83
$3 \times 125 = \square$	77	70
$\square \times 30 = 900$	68	51
$7 \times \square = 280$	65	53
$\square \times 250 = 500$	69	48
$12 \times 500 = \square$	40	32
$30 \times 20 \times 5 = \square$	40	27
$11 \times 273 = 3{,}003$ $22 \times 273 = \square$	53	33
$25 \times 32 = 900$ $26 \times 32 = \square$	36	20
$\square \times 585 = 0$	87	85
$(8 \times 29) + (2 \times 29) = \square$	27	12

Means By Ability Level

	1	2	3	4	All
CSMP	3.4	5.9	6.8	8.6	6.4
non-CSMP	3.1	4.6	5.8	6.5	5.1
t-test	1.0	3.2	3.0	5.9	6.9

	CSMP	Non-CSMP
Number of Students	337	276
KR20 Reliability	.51	.74
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.64	.49

1. See Note 1. for Scale C3

C6 Mental Arithmetic - Division

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
210 Divided by 3 = ☐	66	58
500 Divided by 2 = ☐	75	50
700 Divided by 10 = ☐	84	71
800 Divided by ☐ = 200	60	45
360 Divided by 90 = ☐	46	37
☐ Divided by 3 = 30	63	38
1,200 Divided by ☐ = 4	55	45
3,600 Divided by 15 = 240 3,615 Divided by 15 = ☐	38	31
1,200 Divided by 30 = 40 1,200 Divided by 15 = ☐	38	24
524 Divided by 524 = ☐	76	70
498 Divided by ☐ = 498	78	68

Means By Ability Level

	1	2	3	4	All
CSMP	3.9	6.3	7.4	9.1	6.7
non-CSMP	3.2	4.3	5.7	7.7	5.3
t-test	1.6	3.7	3.8	4.1	6.5

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.85	.84
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.62	.53

1. See Note 1. for Scale C3

Scales C3-C6

Notes:

1. The graphs and covariate statistics for class means on the total of the four mental arithmetic scales are given on the facing page. The adjusted means across classes and the p-value for the individual scales are given below.

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
C3 Addition	3.4	2.9	.01
C4 Subtraction	3.1	2.5	.01
C5 Multiplication	6.4	5.1	.01
C6 Division	6.7	5.4	.01

2. CSMP students did relatively best on items requiring some strategy:

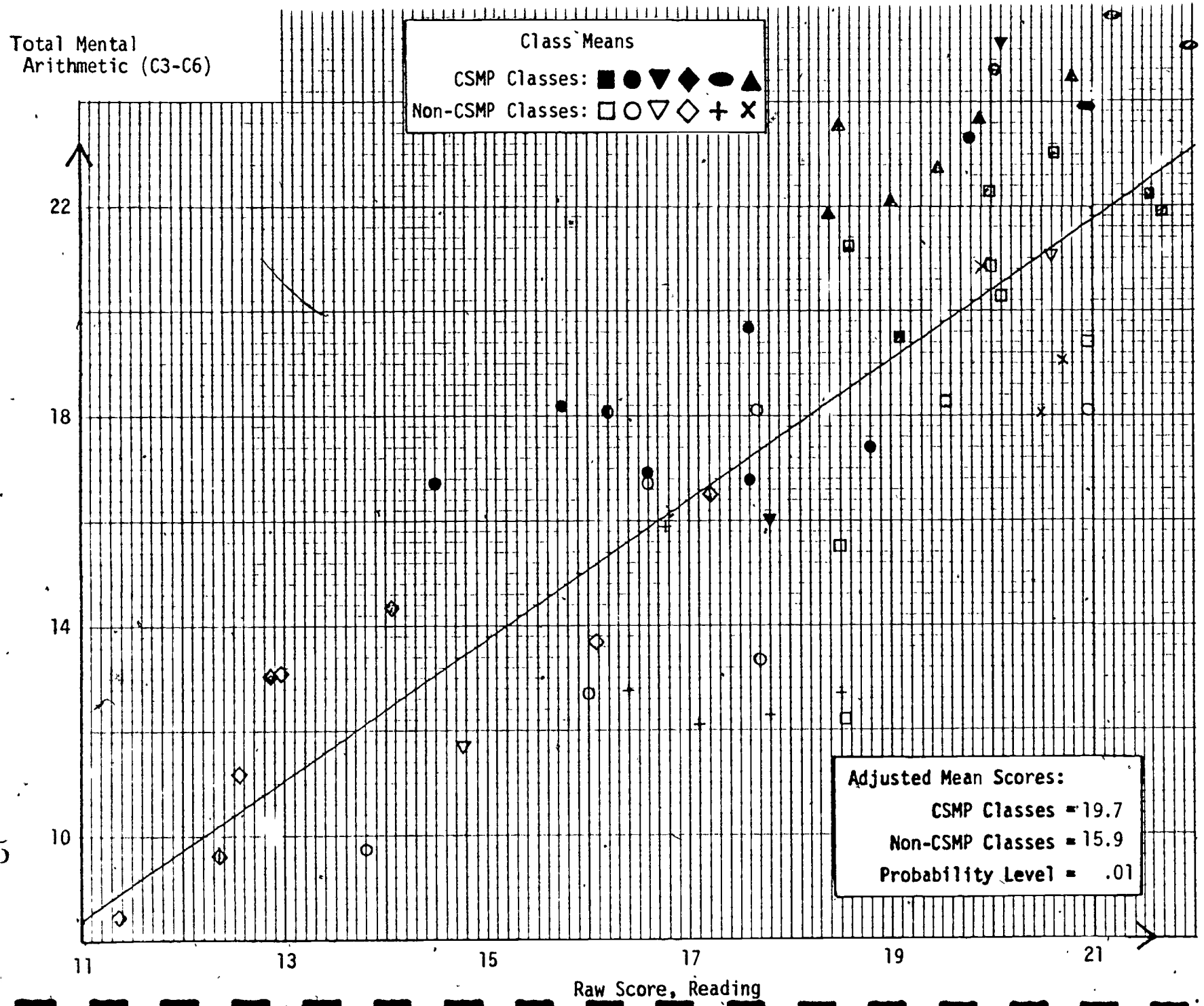
   (8 x 29) + (2 x 29); 30 x 20 x 5; hints such as 11 x 273 = 3,003 22 x 273 = ?

   and 3,600 ÷ 15 = 240 3615 ÷ 15 = ?

   On the six items of these types, the mean percent correct was 39 for CSMP students versus 21 for non-CSMP students.

   Of the remaining items, 13 had the answer box on the left of the equal sign. On these items CSMP students averaged 65% correct versus 52% for non-CSMP.

   The remaining 13 items all had the answer box on the right of the equals sign; the mean percent correct for these items was 69 for CSMP and 60 for non-CSMP.


Total Mental
Arithmetic (C3-C6)
Class Means
CSMP Classes:
Non-CSMP Classes:
22
18
14
10
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 19.7
Non-CSMP Classes = 15.9
Probability Level = .01

E2 Estimating Intervals - Addition

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
ADDITION 185 + 97 0 10 50 100 500 1000	90	87
(Other items used this format.)		
24 + 24	90	85
59 + 39	91	81
479 + 86	82	82
279 + 165	86	84
19 + 29	84	80
257 + 294	70	69
19 + 19 + 19	60	60

Means By Ability Level

	1	2	3	4	All
CSMP	5.5	6.4	6.8	7.2	6.5
non-CSMP	5.6	6.0	6.6	6.9	6.3
t-test	-0.6	1.6	1.4	2.0	1.8

	CSMP	Non-CSMP
Number of Students	630	519
KR20 Reliability	.69	.72
Mean Reading Score	18.0	18.0
Correlation, Reading versus Scale	.47	.34

1. Three sample items were done to illustrate that an x was to be placed between the two numbers (e.g. 50 and 100) which bounded the answer. Working quickly, not figuring out the exact answer, and not getting stuck on one item were stressed. There was a time limit of $1\frac{1}{2}$ minutes.

2. This scale was intended to be, and was, quite easy (mean percent correct = 82), to prepare students for the format and short time limit (but more difficult items) of the next two scales.

57

E3 Estimating Intervals - Multiplication

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
MULTIPLICATION		
5 x 109 0 10 50 100 500 1000	80	72
(Other items used this format.)		
2 x 19	83	80
40 x 10	81	66
4 x 23	80	80
11 x 50	59	47
2 x 49	78	78
4 x 29	53	50

Means By Ability Level

	1	2	3	4	All
CSMP	3.8	5.1	5.3	6.0	5.1
non CSMP	3.9	4.2	5.1	5.5	4.7
t-test	-0.2	4.4	1.3	3.5	4.2

	CSMP	Non-CSMP
Number of Students	630	519
KR20 Reliability	.69	.64
Mean Reading Score	18.0	18.0
Correlation, Reading versus Scale	.46	.36

1. See Note 1. for Scale E2.

E4 Estimating Intervals - Division

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP	Biserial
300 DIVIDED BY 4 0 1 10 20 100	78	73	
(Other items used this format.)			
190 DIVIDED BY 10	63	50	
1 DIVIDED BY 2	55	43	
101 DIVIDED BY 9	58	53	
133 DIVIDED BY 50	40	36	
18.230 DIVIDED BY 1,000	33	34	
850 DIVIDED BY 101	33	32	
180 DIVIDED BY 21	30	21	

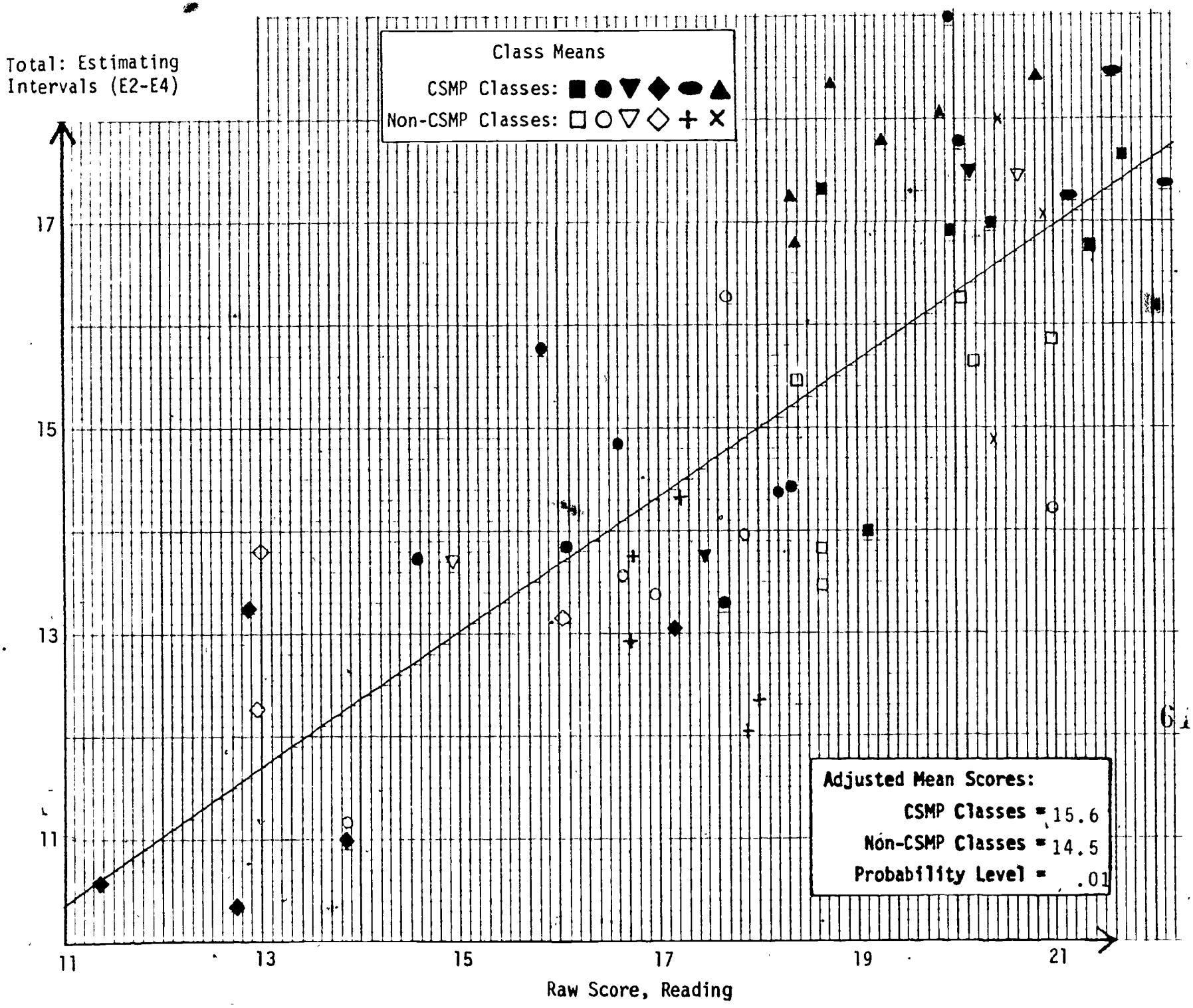
Means By Ability Level

	1	2	3	4	All
CSMP	2.5	3.4	4.2	5.2	3.9
non-CSMP	2.5	2.8	3.6	4.5	3.4
t-test	-0.2	2.5	2.9	2.8	4.3

	CSMP	non-CSMP
Number of Students	630	519
KR20 Reliability	.69	.67
Mean Reading Score	18.0	18.0
Correlation, Reading versus Scale	.47	.34

1. See Note 1. for Scale E2.
2. The facing page shows class means for the total of scales E2-E4. The class mean statistics for the individual scales were as follows:

	Adjusted Class Means CSMP	Adjusted Class Means non-CSMP	p-value
E2, Addition	6.6	6.3	.14
E3, Multiplication	5.2	4.7	.01
E4, Division	3.9	3.4	.01


Total: Estimating Intervals (E2-E4)
Class Means
CSMP Classes:
Non-CSMP Classes:
17
15
13
11
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 15.6
Non-CSMP Classes = 14.5
Probability Level = .01

E6 Most Reasonable Answer - Addition

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
ADD 47,377 26,896 + 31,456 = 51,377 58,377	88	81
931 836 + 31 + 26 + 19 + 27 = 1,131 1,331	53	46
3,740 1,022 + 1,713 + 1,991 = 4,740 11,740	73	64
10,604 10,278 + 558 + 4 = 15,604 19,604	50	50
310 105 + 97 + 106 + 98 + 104 = 410 510	60	58
15,030 5,079 + 5,076 + 5,075 = 15,230 17,230	69	71

Means By Ability Level

	1	2	3	4	All
CSMP	3.1	3.7	4.1	4.5	3.9
non-CSMP	3.4	3.3	3.9	4.2	3.7
t-test	-1.5	2.0	1.0	1.9	1.7

	CSMP	Non-CSMP
Number of Students	337	266
KR20 Reliability	.33	.34
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.41	.24

1. A sample item was done with emphasis on not taking the time to figure out the exact answer. All three alternatives were wrong, but one of them was a lot better than the other two. Scale E6 and E9 (done by half the students) had a combined time limit of $3\frac{1}{2}$ minutes; similarly for Scales E7 and E8.

E7 Most Reasonable Answer - Subtraction

Test Items		Percent Correct CSMP	Percent Correct Non-CSMP
	4,744		
SUBTRACT	7,907 - 2,249 = 5,644	68	75
	7,744		
	7,250		
	78,412 - 5,879 = 23,650	72	74
	72,550		
	940		
	10,153 - 719 = 9,340	67	63
	10,040		
	332		
	1,213 - 888 = 842	31	37
	1,322		
	9,780		
	101,787 - 1,989 = 19,780	70	67
	99,780		
	1,162		
	3,105 - 1,986 = 2,162	23	24
	2,862		

Means By Ability Level

	1	2	3	4	All
CSMP	2.7	3.0	3.4	4.0	3.3
non-CSMP	2.8	3.2	3.5	3.9	3.4
t-test	-0.5	-1.1	-0.1	0.2	-0.7

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.42	.40
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.39	.28

1. See Note 1. for Scale E6.

E8 Most Reasonable Answer - Multiplication

Test Items		Percent Correct CSMP	Percent Correct Non-CSMP
MULTIPLY	9 x 1,120 = 980 / 1,980 / 10,080	83	78
	21 x 123 = 257 / 2,557 / 25,557	57	55
	8 x 123,456 = 1,000,100 / 10,000,100 / 100,000,100	22	29
	15 x 2,111 = 3,173 / 20,173 / 31,173	59	59
	52 x 99 = 1,483 / 5,183 / 9,883	45	45
	11 x 989 = 1,900 / 10,900 / 19,900	44	39

Means By Ability Level

	1	2	3	4	All
CSMP	2.5	2.8	3.0	3.8	3.1
non-CSMP	2.5	3.0	3.1	3.5	3.0
t-test	0.4	-0.5	-0.2	1.2	0.5

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.50	.45
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.29	.28

1. See Note 1. for Scale E6.

E9 Most Reasonable Answer - Division

Test Items		Percent Correct CSMP	Percent Correct Non-CSMP
DIVIDE	3		
1,513 ÷ 498 =	30	32	34
	300		
	2,000		
181,832 ÷ 9 =	20,000	55	55
	200,000		
	15		
980 ÷ 11 =	40	37	40
	100		
	5		
3,641 ÷ 69 =	50	58	58
	500		
	10		
13,980 ÷ 1,402 =	50	45	41
	100		
	10		
2,082 ÷ 39 =	50	43	46
	100		

Means By Ability Level

	1	2	3	4	All
CSMP	1.9	2.4	2.7	3.5	2.7
non-CSMP	2.1	2.1	2.9	3.6	2.7
t-test	-0.8	1.3	-0.8	-0.5	-0.4

	CSMP	Non-CSMP
Number of Students	337	266
KR20 Reliability	.52	.51
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.35	.33

1. See Note 1. for Scale E6.

Scales E6-E9

Notes:

1. These scales had low reliabilities and low correlations with reading scores. This may have been in part due to guessing; the average score was about 12.9 out of 24 but random guessing alone would have produced on expected score of 8.

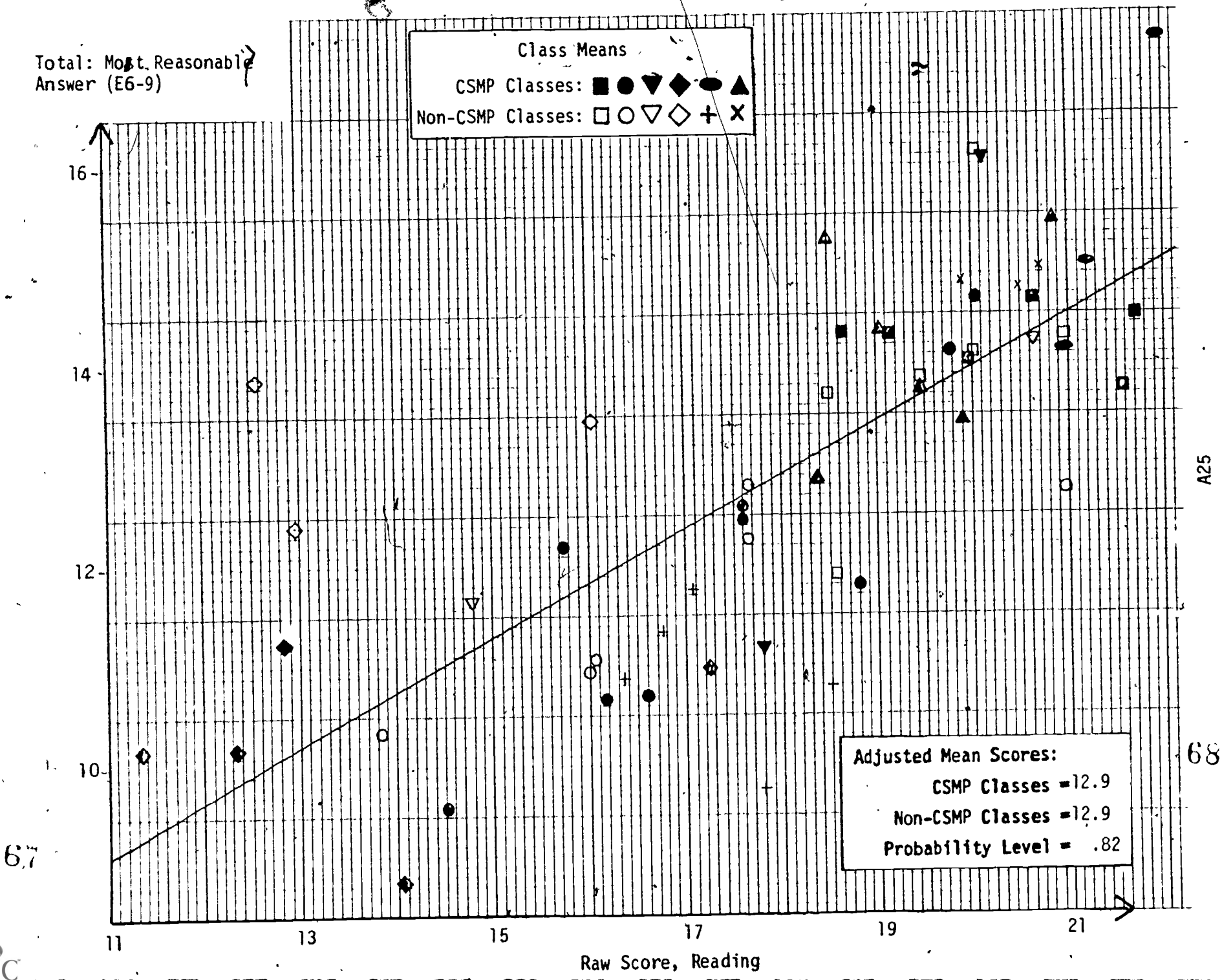
2. CSMP students in the lowest quartile (lowest reading scores) did relatively poorly compared to their non-CSMP couterparts. It can be seen from the graph page 25, that this was mainly due to the non-CSMP students (and classes) at the lowest reading level who did nearly as well as students at the next highest reading level.

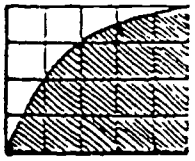
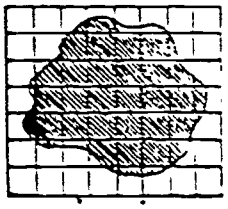
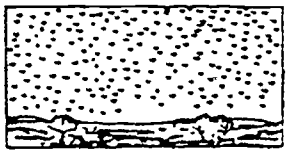
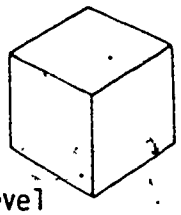
3. The class mean statistics for the individual scales were as follows:

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
E6, Addition	3.9	3.7	.09
E7, Subtraction	3.3	3.4	.38
E8, Multiplication	3.1	3.0	.86
E9, Division	2.7	2.7	.69

Except for addition, the scores were virtually identical.

4. For several items, surprisingly few students got the correct answer (for example, the last item on E7, the last two on E8, the first and third items on E9).


Total: Most Reasonable Answer (E6-9)
Class Means
CSMP Classes:
Non-CSMP Classes:
16
14
12
10
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes =12.9
Non-CSMP Classes =12.9
Probability Level = .82

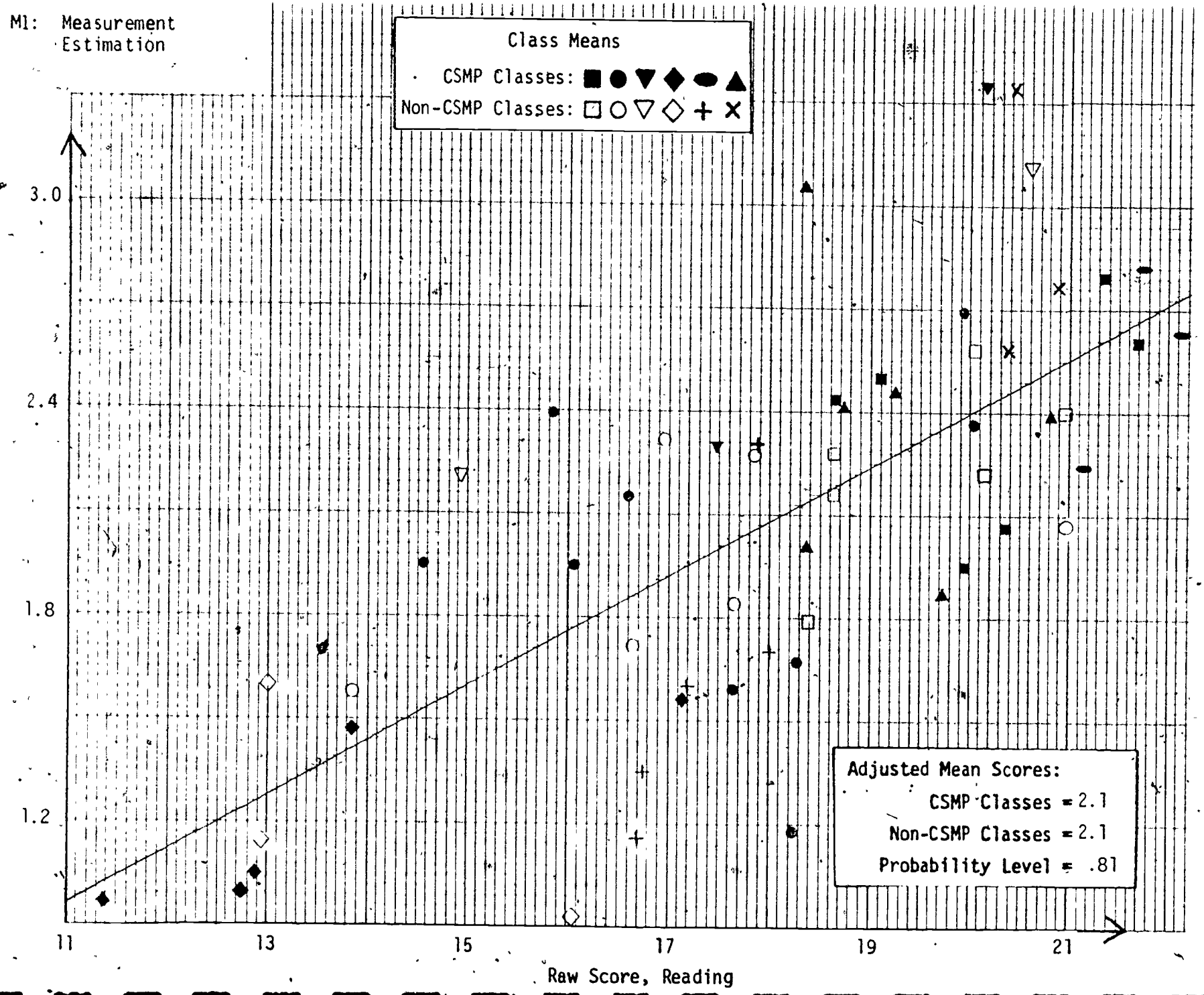
Test Items: Allowable range of answers given on answer blanks.	Percent Correct CSMP	Non-CSMP
 This playground is divided into 20 sections. It takes one gallon of paint to cover one section. About how many gallons of paint would it take to cover the shaded part of the playground? 13-16	56	47
 About how many gallons of paint would it take to cover the shaded part of this playground? 27-33	31	29
 This is a picture of birds flying south. You should not count them all. But about how many birds are in the picture? 61-299	42	44
Your school desk is about 70 centimeters high. About how many centimeters high is the average doorway? 175-245	38	41
If it takes a gallon of paint to cover this, → □ About how many gallons would it take to cover this? 6-8 → 	18	21
About how many blocks like this ⬡ would fit into the box below? 50-220 	24	25

Means By Ability Level

	1	2	3	4	All
CSMP	1.5	1.9	2.1	2.6	2.1
non-CSMP	1.3	1.9	2.2	2.7	2.1
t-test	1.0	0.1	-0.2	-0.4	0.3

	CSMP	Non-CSMP
Number of Students	630	519
KR20 Reliability	.35	.39
Mean Reading Score	18.0	18.0
Correlation, Reading versus Scale	.33	.36

1. Students were told they would not be able to figure exact answers, but to make their best estimate.
2. The percentages correct were greatly affected by the allowable range decided upon. This was intentionally fairly narrow so that either fairly good intuitive estimation or some strategy was needed to get the items correct.
3. As one might expect for this scale, the correlation with reading was fairly low. However the reliability was also low and this together with the rather low scores in general on this scale indicate that much "uneducated" guessing took place and that this kind of task is not one with which students get much practice.


M1: Measurement Estimation
Class Means
CSMP Classes:
Non-CSMP Classes:
3.0
2.4
1.8
1.2
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 2.1
Non-CSMP Classes = 2.1
Probability Level = .81

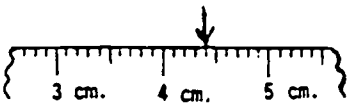
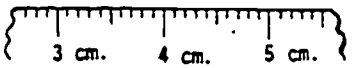
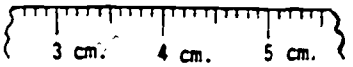
N1. Decimal Gas

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
1. Peter has 6.5 gallons. Then he spills 1.2 gallons. How much gas will he have left? _____	91	89
2. Tom has 6.5 gallons. He buys 3.5 more gallons. How much gas will he have then? _____	78	68
3. John has 6.5 gallons. He uses up four gallons. How much gas will he have left? _____	73	61
4. Bill has 6.5 gallons. He buys another half gallon. How much gas will he have then? _____	61	40
5. Ron has 6.5 gallons. Next week he will use ten times this much. How much gas will he use next week? _____	61	43
6. Joe has 6.5 gallons. He sells each gallon for $2. How much money will he get altogether? _____	41	26
7. Ken has 6.5 gallons of gas. He gives away half of it. How much gas will he have left? _____	51	24

Means By Ability Level

	1	2	3	4	All
CSMP	2.9	4.1	4.8	5.8	4.5
non-CSMP	2.3	3.1	3.7	4.6	3.5
t-test	3.0	4.6	5.7	8.1	9.2

	CSMP	Non-CSMP
Number of Students	632	511
KR20 Reliability	.73	.64
Mean Reading Score	18.2	18.0
Correlation, Reading versus Scale	.59	.46

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
The arrow is pointing at _____ cm. 	71	47
Put an arrow at 3.4 cm. 	88	76
Put an arrow at 4.25 cm. 	37	17
Which is larger?		
6.1 or 6.01	76	50
1.5 or 0.58	81	71
4.077 or 4.155	83	82
4.999 or 5.1	69	51
0.9 or 0.111	52	32
Each bucket holds 1 gallon. How many gallons are shown? Circle the best answer  3.0 3.3 3.5 3.8 4.0	57	42
How many gallons are shown? Circle the best answer.  1.0 1.2 1.5 1.8 2.0	50	30

Means By Ability Level

	1	2	3	4	All
CSMP	4.1	5.7	7.5	8.3	6.6
non-CSMP	3.3	4.6	5.3	6.2	4.9
t-test	2.5	3.0	7.0	7.0	9.9

	CSMP	Non-CSMP
Number of Students	337	226
KR20 Reliability	.78	.67
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.62	.44

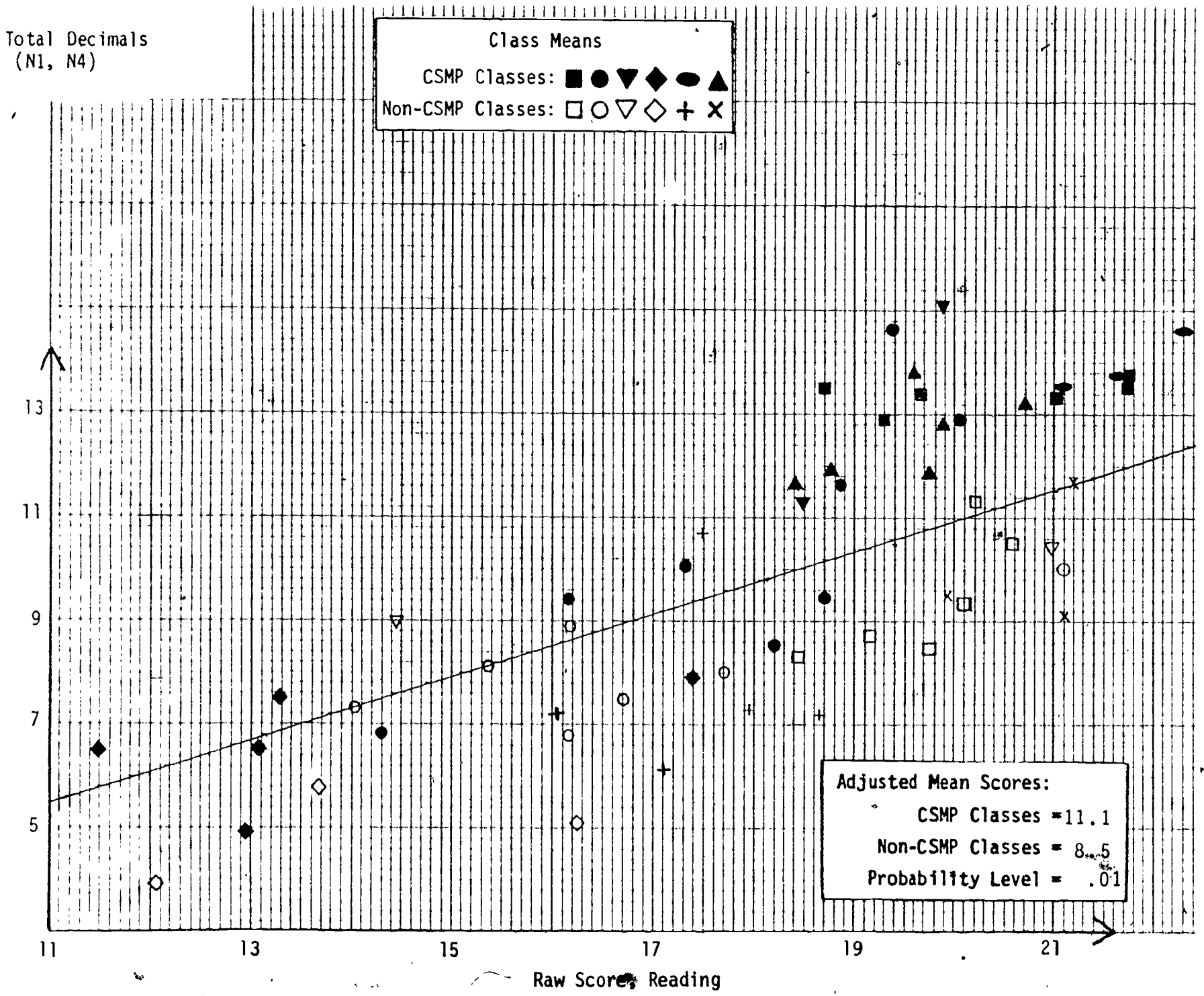
Scales N1,N4

Notes:

1. The class mean statistics for these two scales were as follows:

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
N1, Decimal Gas	4.5	3.5	.01
N4, Decimal Magnitudes	6.6	5.0	.01

2. There was little variation in the various items regarding differences in percent correct between CSMP and non-CSMP students; the CSMP advantage was quite consistent for all types of items.


Total Decimals
(N1, N4)
Class Means
CSMP Classes:
Non-CSMP Classes:
13
11
9
7
5
11
13
15
17
19
21
Raw Score: Reading
Adjusted Mean Scores:
CSMP Classes = 11.1
Non-CSMP Classes = 8.5
Probability Level = .01

N2 Negative Hits and Misses (Form 1)

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
ABOVE ZERO - BELOW ZERO Each Hit: Gain 5 points Each Miss: Lose 1 point		
Jim — Started with a score of: 0; Number of Hits: 1; Number of Misses: 1; Ended with a score of:	78	76
Sue — Started with a score of: 8 below zero; Number of Hits: 1; Number of Misses: 0; Ended with a score of:	63	56
Rick — Started with a score of: 3 below zero; Number of Hits: 1; Number of Misses: ; Ended with a score of: 0	68	62
Pam — Started with a score of: 4; Number of Hits: ; Number of Misses: 6; Ended with a score of: 3	57	56
Joel — Started with a score of: ; Number of Hits: 0; Number of Misses: 2; Ended with a score of: 7 below zero	60	51
Number of Students	315	263
KR20 Reliability	.83	.81
Mean Reading Score	18.1	18.2
Correlation, Reading versus Scale	.53	.56

1. The rules for gaining and losing points in the game were explained and two examples (one of which used "below zero") were done.

N2 Negative Hits and Misses (Form 2)

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
ABOVE ZERO - BELOW ZERO Each Hit: Gain 5 points Each Miss: Lose 1 point		
Bill — Started with a score of: 0 — Number of Hits: 1 — Number of Misses: 1 — Ended with a score of:	80	73
Jane — Started with a score of: 3 — Number of Hits: 0 — Number of Misses: 7 — Ended with a score of:	68	68
Peter — Started with a score of: 10 below zero — Number of Hits: 1 — Number of Misses: — Ended with a score of: 12 below zero	63	54
Beth — Started with a score of: 3 below zero — Number of Hits: — Number of Misses: 2 — Ended with a score of: 5	54	41
John — Started with a score of: — Number of Hits: 2 — Number of Misses: 0 — Ended with a score of: 15 below zero	47	37
Number of Students	315	256
KR20 Reliability	.81	.78
Mean Reading Score	18.0	17.8
Correlation, Reading versus Scale	.57	.53

1. See Note 1. for N2, Form 1.

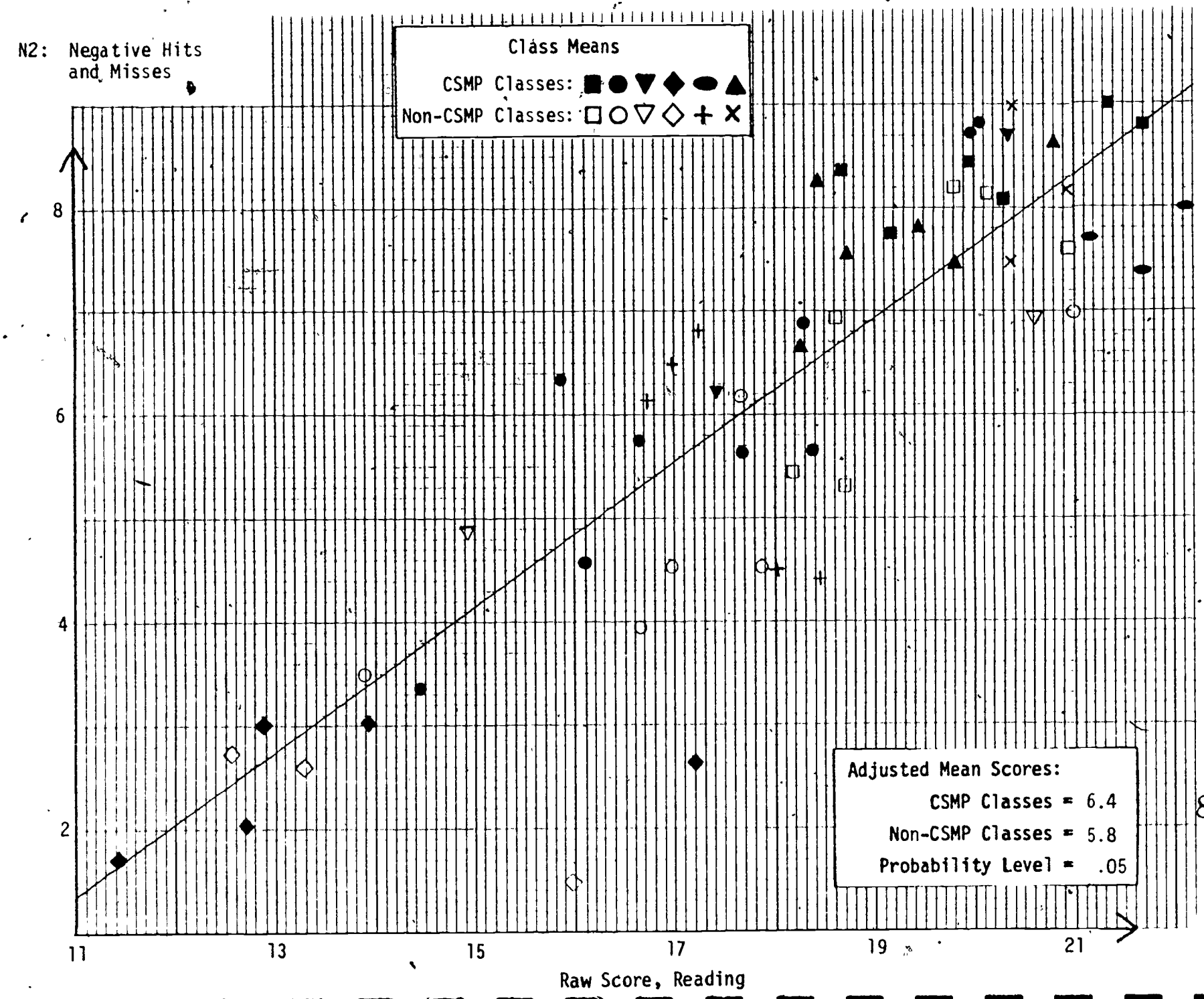
Scale N2

Notes:

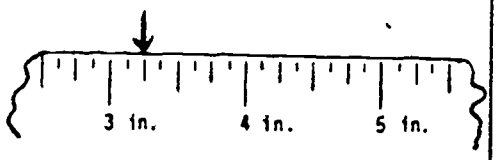
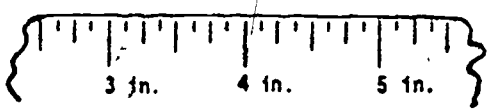
1. The table below shows means by ability level regardless whether the student took Form 1 or Form 2.

Means By Ability Level

	1	2	3	4	All
CSMP	1.6	2.9	3.6	4.3	3.2
non-CSMP	1.4	2.3	3.3	4.1	2.8
t-test	0.9	2.9	1.6	1.9	3.0


N2: Negative Hits and Misses
Class Means
CSMP Classes:
Non-CSMP Classes:
8
6
4
2
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 6.4
Non-CSMP Classes = 5.8
Probability Level = .05

N3 Measuring Fractional Inches

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
The arrow is pointing at _____ in. 	31	41
Put an arrow at $4\frac{1}{2}$ in. 	77	71
Put an arrow at $3\frac{3}{4}$ in. 	36	46
Number of Students	337	266
KR20 Reliability	.63	.67
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.52	.49

Notes:

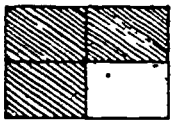
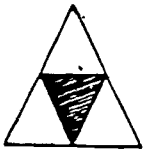
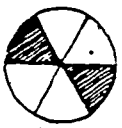
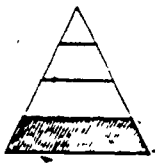
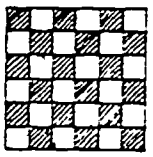
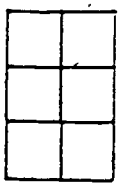
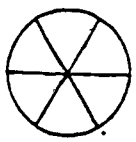
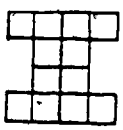
The class mean statistics for the individual fraction scales (N3 above, and N5-10 on the following pages) are given below; the graph and statistics for the total of these 7 scales appears on page A43.

	Adjusted Class Means CSMP	Adjusted Class Means non-CSMP	p-value
N3, Measuring Fractional Inches	1.4	1.6	.17
N4, Fractional Areas	4.1	3.9	.25
N6, Equivalent Fractions	14.0	13.4	.19
N7, Fractional Open Sentences	3.2	2.6	.01
N8, Which Fraction's Larger	3.4	3.1	.13
N9, Fractional Word Problems	3.0	2.6	.01
N10, Other Representations	3.9	3.9	.85

It can be seen that CSMP classes did significantly better on two of these individual scales.

82

N5 Fractional Areas

Test Items					Percent Correct CSMP	Percent Correct Non-CSMP
	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{3}{4}$	none of these	77	80
	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	none of these	82	83
	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{4}$	none of these	18	14
	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	none of these	20	15
	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$	none of these	40	30
Shade $\frac{1}{2}$ of the figure. 					89	83
Shade $\frac{2}{3}$ of the figure. 					45	47
Shade $\frac{1}{4}$ of the figure. 					37	36

Means By Ability Level

	1	2	3	4	All
CSMP	2.6	3.8	4.1	5.5	4.1
non-CSMP	2.9	3.5	4.2	4.9	3.9
t-test	-1.1	1.2	-0.3	2.8	1.5

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.66	.63
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.58	.44

Note:

1. Both groups of students did surprising poorly on the 3rd and 4th items.

Test Items	Percent Correct CSMP	Non-CSMP
General Format		
Circle the fractions that are equal to the one in the box.		
$\boxed{\frac{1}{3}}$ $\frac{2}{6}$ $\frac{50}{150}$ $\frac{11}{31}$ $\frac{3}{15}$ $\frac{4}{12}$		
$\frac{1}{3}$ → $\frac{2}{6}$	89	83
→ $\frac{50}{150}$	29	27
→ $\frac{11}{31}$	88	87
→ $\frac{3}{15}$	70	64
→ $\frac{4}{12}$	70	68
$\frac{3}{4}$ → $\frac{9}{12}$	65	64
→ $\frac{31}{41}$	86	82
→ $\frac{6}{8}$	74	74
→ $\frac{13}{14}$	82	81
→ $\frac{300}{400}$	45	42
$\frac{1}{5}$ → $\frac{20}{100}$	48	45
→ $\frac{5}{1}$	80	74
→ $\frac{11}{15}$	85	80
→ $\frac{5}{25}$	63	68
→ $\frac{2}{10}$	79	70
$\frac{2}{3}$ → $\frac{25}{45}$	88	89
→ $\frac{4}{6}$	86	80
→ $\frac{30}{45}$	32	28
→ $\frac{3}{2}$	79	74
→ $\frac{5}{15}$	80	76

Means By Ability Level

	1	2	3	4	All
CSMP	12.0	13.3	14.2	16.2	14.1
non-CSMP	9.9	12.9	14.2	16.3	13.6
t-test	3.4	0.6	0.0	-0.2	1.8

	CSMP	Non-CSMP
Number of Students	337	266
KR20 Reliability	.83	.84
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.40	.51

1. A completed example was provided.

2. In each of the four groups of items one item was much harder than the others, namely that equivalent fraction with large numerators and denominators (e.g. 50/150, 300/400, etc.). It was also true that students did slightly better on fractions which were not equivalent (i.e. which should not have been circled).

N7 Fractional Open Sentences

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
Complete the sentences:		
$\frac{2}{3} \times 1 = \square$	70	56
$\frac{1}{2} + \square = 1$	71	67
$\frac{1}{2} \times \square = 10$	38	13
$\frac{1}{2} \div 2 = \square$	29	14
$1 - \frac{3}{4} = \square$	33	39
$\frac{1}{4} + \square = \frac{1}{4}$	76	72

Means By Ability Level

	1	2	3	4	All
CSMP	1.8	2.8	3.2	4.5	3.1
non-CSMP	1.5	2.3	2.9	3.7	2.6
t-test	1.4	2.2	1.0	3.7	4.3

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.76	.69
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.55	.53

N8 Which Fraction is Larger?

Test Items	Percent Correct CSMP	Non-CSMP	
Which is larger?			
$\frac{1}{2}$ or $\frac{1}{3}$	69	63	
$\frac{3}{4}$ or $1\frac{1}{4}$	79	86	
$\frac{5}{2}$ or $\frac{5}{4}$	62	48	
$\frac{3}{4}$ or $\frac{5}{10}$	61	54	
$\frac{1}{100}$ or $\frac{1}{2}$	72	64	95,88

Means By Ability Level

	1	2	3	4	All
non-CSMP	2.0	3.4	3.7	4.4	3.4
CSMP	2.1	2.8	3.3	4.2	3.1
t-test	-0.1	2.5	1.4	1.4	2.3

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.79	.77
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.52	.43

N9 Fractional Word Problems

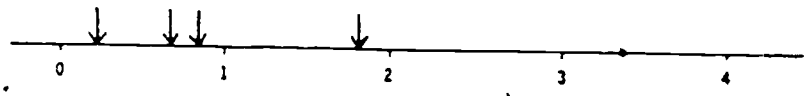
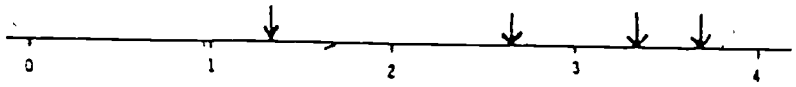
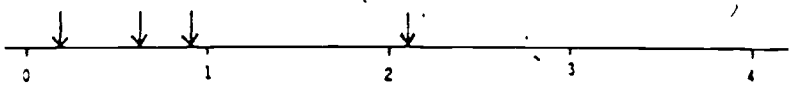
Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
Complete the sentences:		
$\frac{1}{2}$ of a 40-pound bag of dog food is _____ pounds.	84	73
$\frac{1}{4}$ of a 200-page book is _____ pages.	69	49
$\frac{1}{3}$ of a dozen eggs is _____ eggs.	72	49
$\frac{3}{4}$ of a dollar is _____ cents.	41	52
$\frac{2}{3}$ of a 30-ounce bottle is _____ ounces.	38	33

Means By Ability Level

	1	2	3	4	All
CSMP	1.3	2.6	3.2	4.3	3.0
non-CSMP	0.7	1.8	3.1	4.1	2.6
t-test	2.5	3.2	0.7	1.7	4.0

	CSMP	Non-CSMP
Number of Students	337	266
KR20 Reliability	.80	.84
Mean Reading Score	18.3	18.1
Correlation, Reading versus Scale	.62	.64

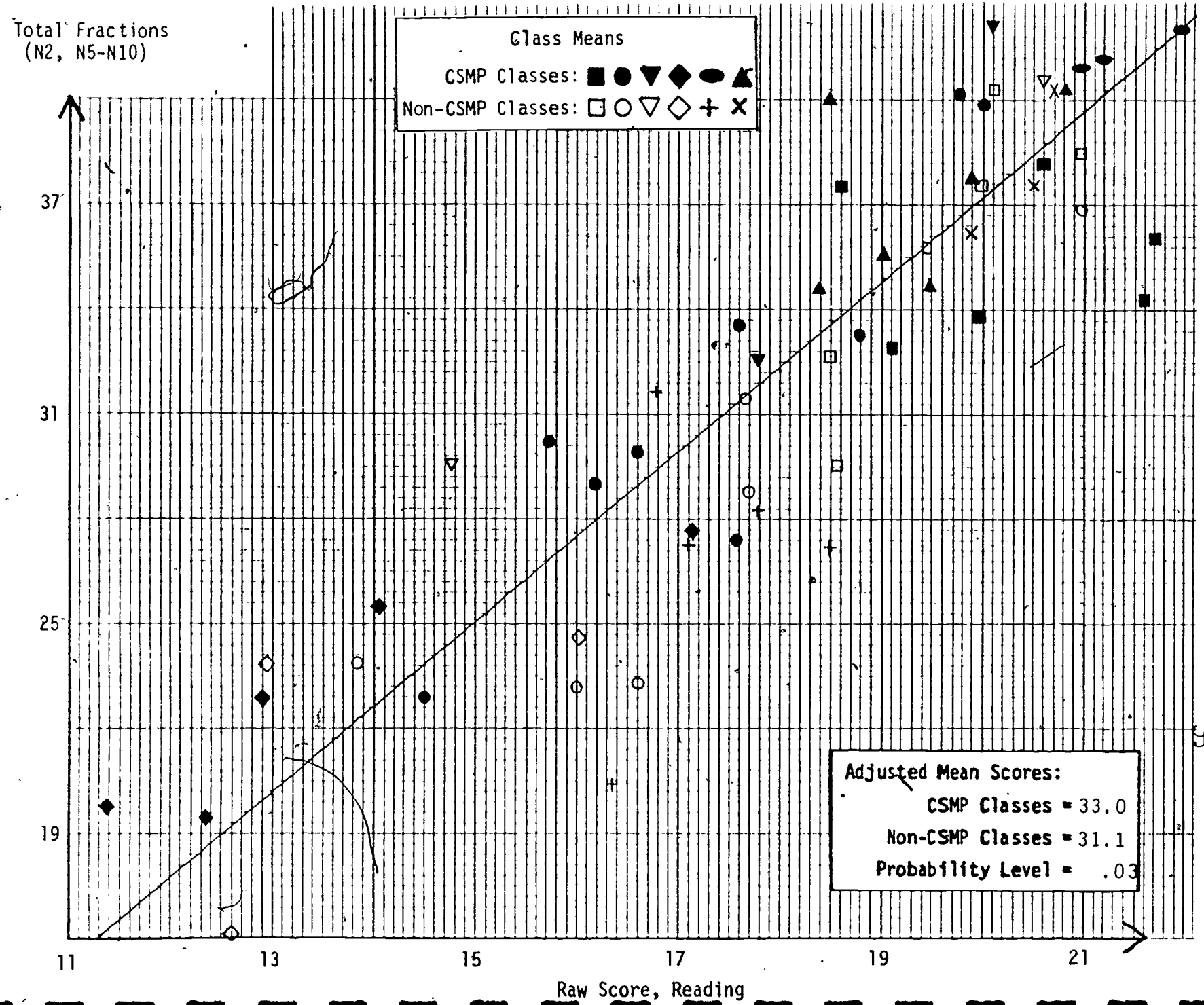
N10 Other Representations of Fractions

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
Circle the arrow that points to $\frac{1}{4}$ on the number line. 	47	41
Circle the arrow that points to $3\frac{1}{3}$ on the number line. 	79	74
Circle the arrow that points to $\frac{9}{10}$ on the number line. 	60	55
Circle the arrow that points to $2\frac{3}{4}$ on the number line. 	57	66
Each bucket holds 1 gallon. How many gallons are shown? Circle the best answer.  $1\frac{1}{4}$ $1\frac{1}{2}$ $1\frac{3}{4}$ 2 $2\frac{1}{2}$	83	81
How many gallons are shown? Circle the best answer.  $2\frac{1}{10}$ $2\frac{1}{4}$ $2\frac{1}{2}$ $2\frac{3}{4}$ 3	65	66

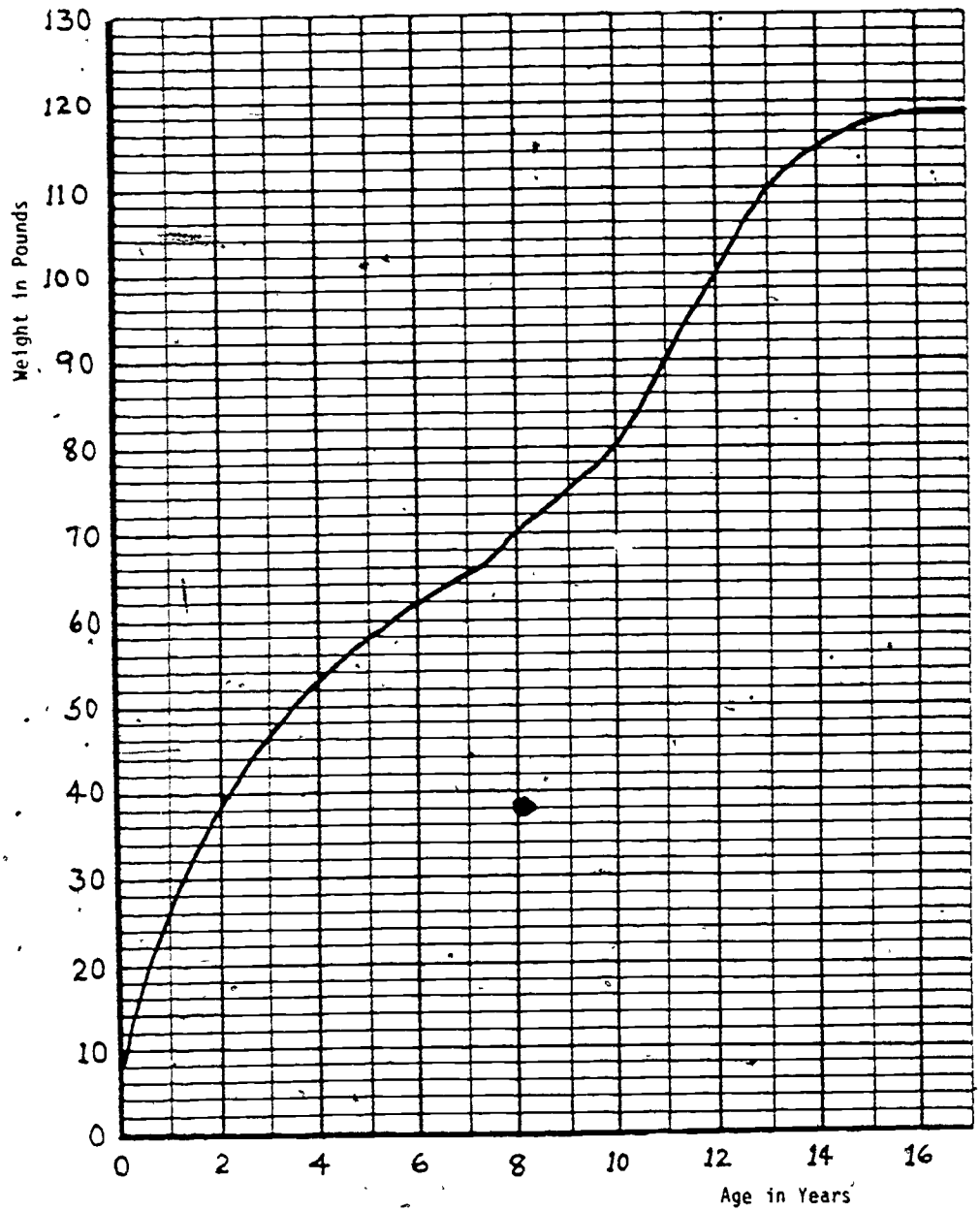
Means By Ability Level

	1	2	3	4	All
CSMP	2.7	3.8	3.9	5.0	3.9
non-CSMP	2.5	3.5	4.3	4.8	3.8
t-test	0.8	0.8	-1.5	1.0	0.5

	CSMP	Non-CSMP
Number of Students	327	276
KR20 Reliability	.56	.70
Mean Reading Score	17.9	17.8
Correlation, Reading versus Scale	.56	.47


Total Fractions
(N2, N5-N10)
Class Means
CSMP Classes:
Non-CSMP Classes:
37
31
25
19
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 33.0
Non-CSMP Classes = 31.1
Probability Level = .03

Weight Chart for Bill from Birth to Age Seventeen



01 Weight Graph

Test Items	Percent Correct CSMP	Non-CSMP
1. How much did Bill weigh at 8 years of age? _____	94	95
2. How old was Bill when he reached 80 pounds? _____	94	95
3. How much did Bill weigh at 13 years of age? _____	89	90
4. How much did Bill weigh at 2 years of age? _____	56	56
5. How much did Bill weigh at 7 years of age? _____	48	47
6. How much did Bill weigh at $5\frac{1}{2}$ years of age? _____	68	70
7. How old was Bill when he reached 90 pounds? _____	85	86
8. How old was Bill when he reached 50 pounds? _____	66	67
9. How much do you think Bill will weigh when he gets to be 18? _____	23	21
10. For how many years was Bill between 50 and 70 pounds? (Circle one) $3\frac{1}{2}$ 4 years $4\frac{1}{2}$ years 5 years $5\frac{1}{2}$ years	27	21

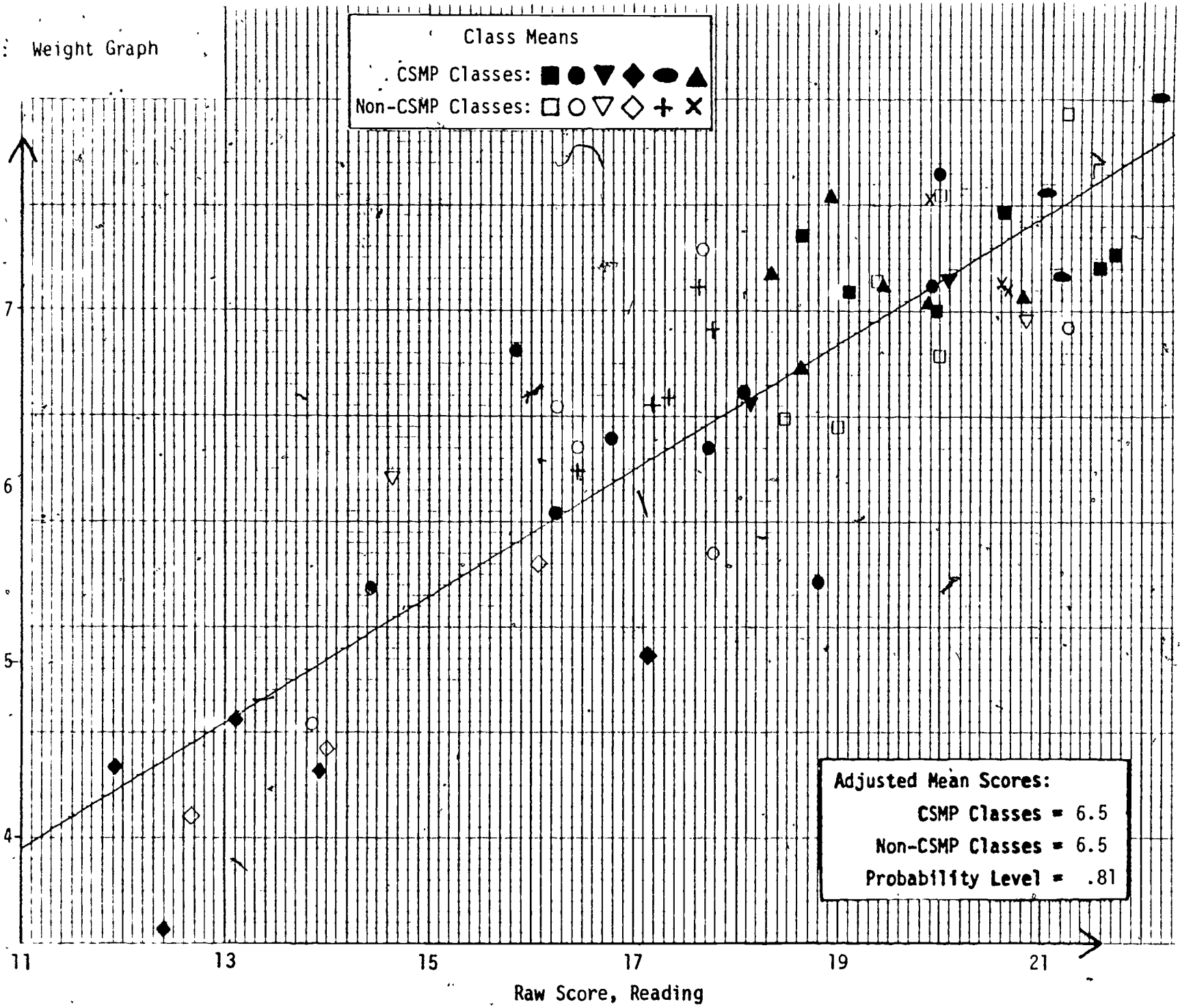
Means By Ability Level

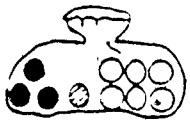
	1	2	3	4	All
CSMP	4.7	6.1	6.7	7.8	6.4
non-CSMP	5.0	6.2	6.9	7.5	6.5
t-test	-1.0	-0.5	-1.1	1.9	0.4

	CSMP	Non-CSMP
Number of Students	626	511
KR20 Reliability	.70	.66
Mean Reading Score	18.2	18.0
Correlation, Reading versus Scale	.53	.49

1. The meaning of the axes, and the several points on the graph were explained. A sample item was done.

2. Except for the last item, the percent correct for CSMP and non-CSMP students never varied more than 2 percentage points. The graph of class means on the next page indicates how similar the two groups of classes were; it also shows that most mean scores fall fairly close to the regression line (i.e. class score on this test is fairly well predicted by reading score).


01: Weight Graph
Class Means
CSMP Classes:
Non-CSMP Classes:
7
6
5
4
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 6.5
Non-CSMP Classes = 6.5
Probability Level = .81

Test Items	Percent Correct CSMP	Non-CSMP
In 100 trials, how often would each of the following occur?		
1. 		
a. Black marble?____ } Counted correct if within 5 best answers.	44	40
b. White marble?____ }	51	44
c. Shaded?____ }	44	36
d. Not white?____ }	39	44
Correct order: b > a and b > c	69	61
a = c	64	54
Correct relative size of answers: a = 2b or c = 2b	47	38
d = 100 -b or d = a+c	50	40
2. 		
a. Black marble?____	60	53
b. White marble?____	47	42
c. Shaded marble?____	62	56
d. White or shaded?____	42	35
Correct order: c > b	76	76
C > a	78	74
Correct relative size: a = 3c } Counted as only 2 responses because of linear dependence	46	39
b = bc }	44	39
b = 2a }	49	43
d = 100 -a or d = b+c	50	40
3. 		
a. Black part?____	59	56
b. White part?____	63	57
c. Shaded part?____	51	43
Correct order: a < c and b < c	76	70
a = b	69	63
Correct size: c = 2b or c = 2a	52	41
4. If you wanted black to win, which game should you play?	52	45

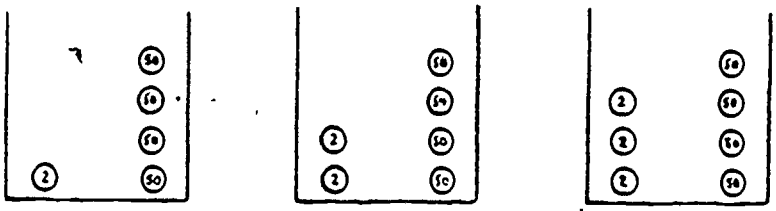
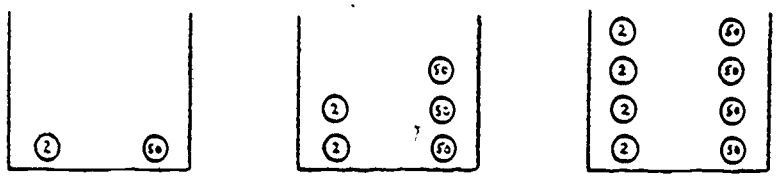
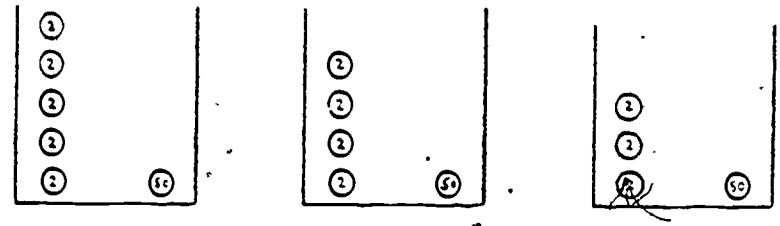
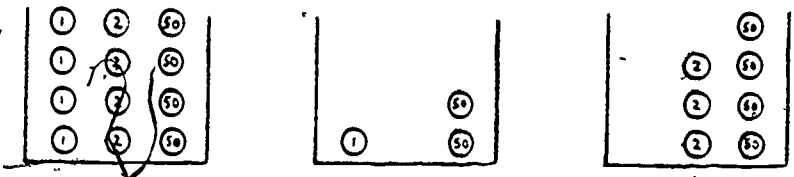
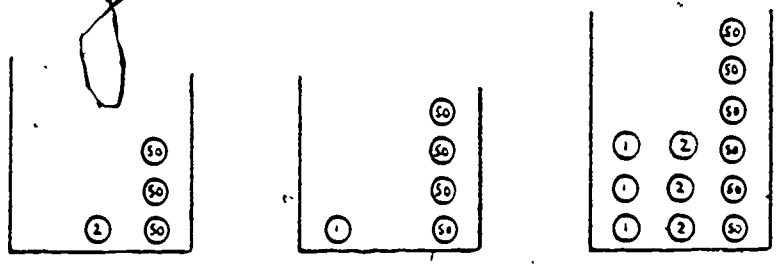
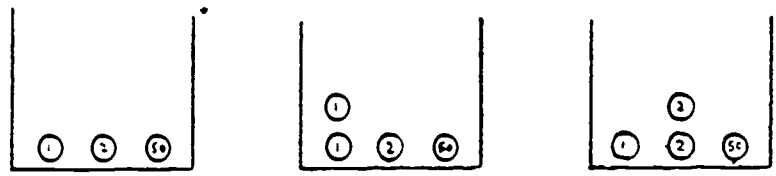
Means By Ability Level

	1	2	3	4	All
CSMP	7.6	11.8	14.4	18.9	13.5
non-CSMP	7.2	10.6	13.4	16.8	12.3
t-test	0.5	1.6	1.4	3.4	3.6

	CSMP	Non-CSMP
Number of Students	664	544
Mean Reading Score	18.1	18.0
Correlation, Reading versus Scale		

1. The directions for Game 1 were reviewed, with an emphasis on not being able to tell ahead of time what would happen, but to make the best guess.
2. For items 1-3, students only answer a., b., c. and d. Based on these responses scoring was done for "correct order" (large answers given for more likely outcomes), "correct relative size" (for example, an event twice as likely got a response twice as big), and whether the response for d was consistent with the responses for a-c.

P3 Which Box?

Test Items	Percent Correct	
	CSMP	Non-CSMP
WHICH BOX WOULD YOU CHOOSE? MONDAY 	71	70
TUESDAY 	65	54
WEDNESDAY 	66	67
THURSDAY 	48	47
FRIDAY 	57	53
SATURDAY 	60	59

Means By Ability Level

	1	2	3	4	All
CSMP	2.3	3.5	3.9	4.7	3.7
non-CSMP	2.5	3.2	3.7	4.4	3.5
t-test	-0.6	1.3	0.8	1.7	1.6

	CSMP	Non-CSMP
Number of Students	630	579
KR20 Reliability	.84	.84
Mean Reading Score	18.0	18.0
Correlation, Reading versus Scale	.40	.31

1. A sample box was discussed to illustrate how a blind draw of one ball would be made. Then the students had to decide which of three given boxes they would like to use for their hypothetical draw.

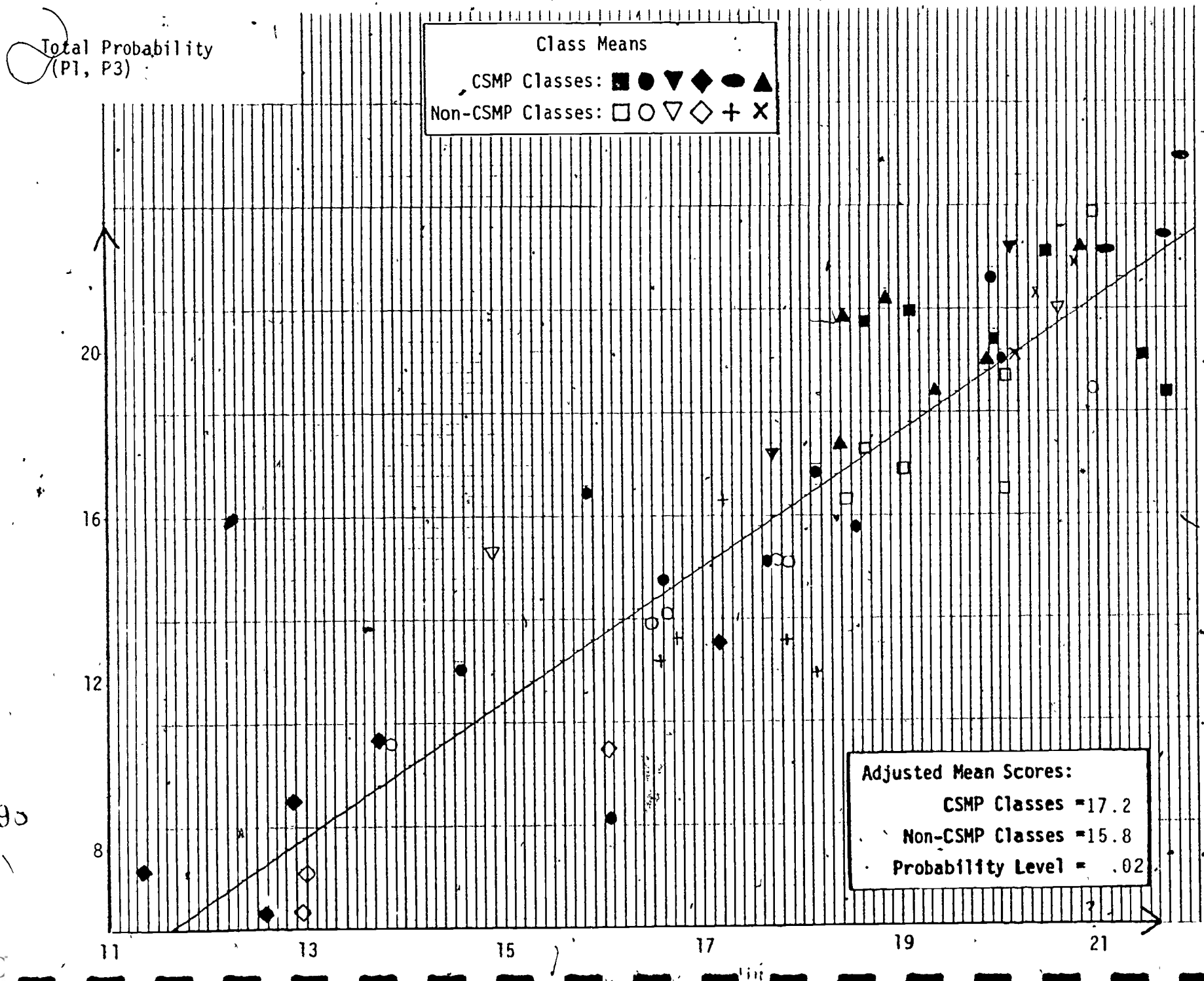
Scales P1,P3

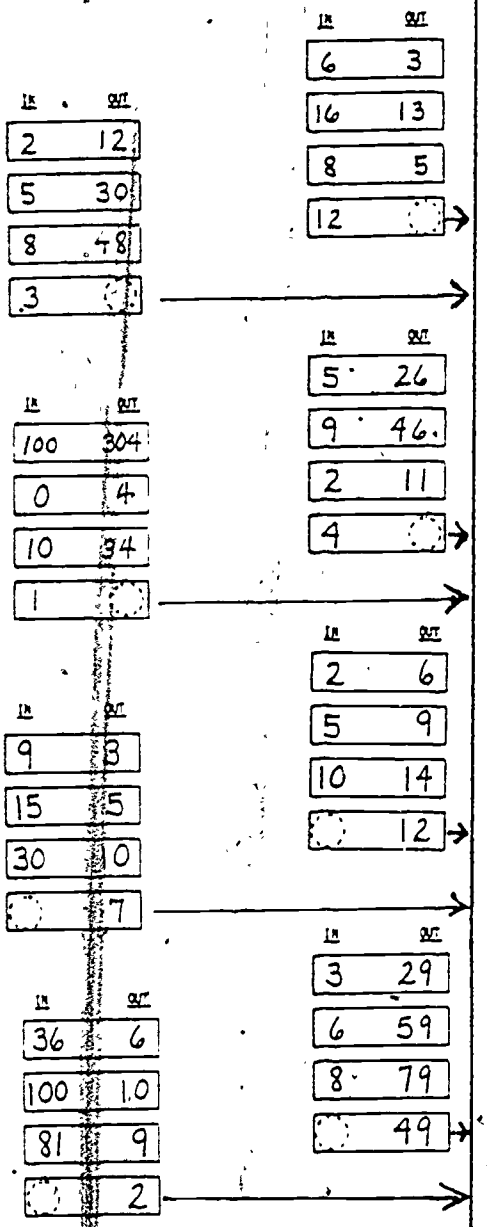
Notes:

1. The class mean statistics for the individual scales were as follows:

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
P1, 100 Outcomes	13.5	12.3	.02
P3, Which Box?	3.7	3.5	.40

2. P3 had very high reliability for such a short scale (.84) and low correlation with reading (.3 and .4). In other words, the items were homogeneous, measuring a single task, but that task was not particularly highly related to reading ability.


Total Probability
(P1, P3)
Class Means
CSMP Classes:
Non-CSMP Classes:
20
16
12
8
11
13
15
17
19
21
Adjusted Mean Scores:
CSMP Classes =17.2
Non-CSMP Classes =15.8
Probability Level = .02

Test Items	Percent Correct CSMP	Non-CSMP
		
1. 6 → 3; 16 → 13; 8 → 5; 12 → ☐	83	75
2. 2 → 12; 5 → 30; 8 → 48; 3 → ☐	79	74
3. 5 → 26; 9 → 46; 2 → 11; 4 → ☐	71	66
4. 100 → 304; 0 → 4; 10 → 34; 1 → ☐	54	40
5. 2 → 6; 5 → 9; 10 → 14; ☐ → 12	82	72
6. 9 → 3; 15 → 5; 30 → 10; ☐ → 7	72	59
7. 3 → 29; 6 → 59; 8 → 79; ☐ → 49	51	43
8. 36 → 6; 100 → 10; 81 → 9; ☐ → 2	59	52

Means By Ability Level

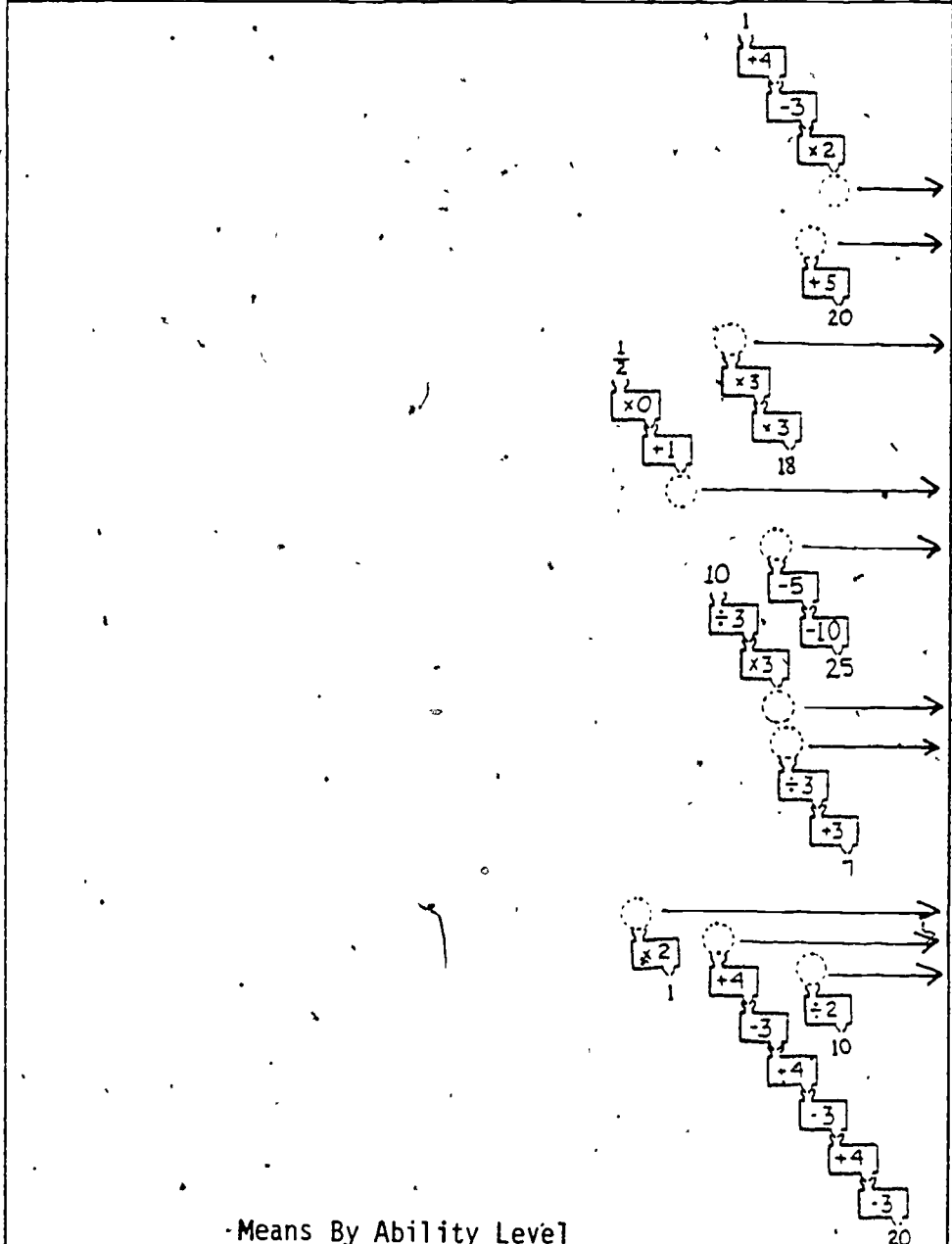
	1	2	3	4	All
CSMP	3.2	5.1	6.0	7.0	5.5
non-CSMP	2.6	4.2	5.4	6.4	4.8
t-test	2.5	3.2	2.6	4.1	6.0

	CSMP	Non-CSMP
Number of Students	632	511
KR20 Reliability	.81	.81
Mean Reading Score	18.2	18.0
Correlation, Reading versus Scale	.62	.58

1. Sample items were done in which it was stressed that students had to figure out from the first three numbers going in and out, what the machine was doing, and then complete the fourth line.

2. The first, second, fifth, and sixth items required what might be called a one-step operation (e.g. -3, x6, +4 ÷ 3) while the others were two-step (e.g. (x5 + 1), x3 = 4), (x10 - 1), square root). Naturally items of the first type were always easier than items of the second type, though the CSMP advantage was about the same on each.

R2 Using Functions

Test Items	Percent Correct CSMP	Non-CSMP
		
→	93	88
→	93	89
→	79	66
→	53	52
→	78	68
→	32	18
→	59	43
→	53	30
→	58	38
→	75	54

Means By Ability Level

	1	2	3	4	All
CSMP	4.5	6.2	7.3	8.3	6.7
non-CSMP	3.6	4.8	5.8	7.1	5.4
t-test	3.5	4.8	6.7	6.9	8.9

	CSMP	Non-CSMP
Number of Students	632	511
KR20 Reliability	.75	.74
Mean Reading Score	18.2	18.0
Correlation, Reading versus Scale	.61	.54

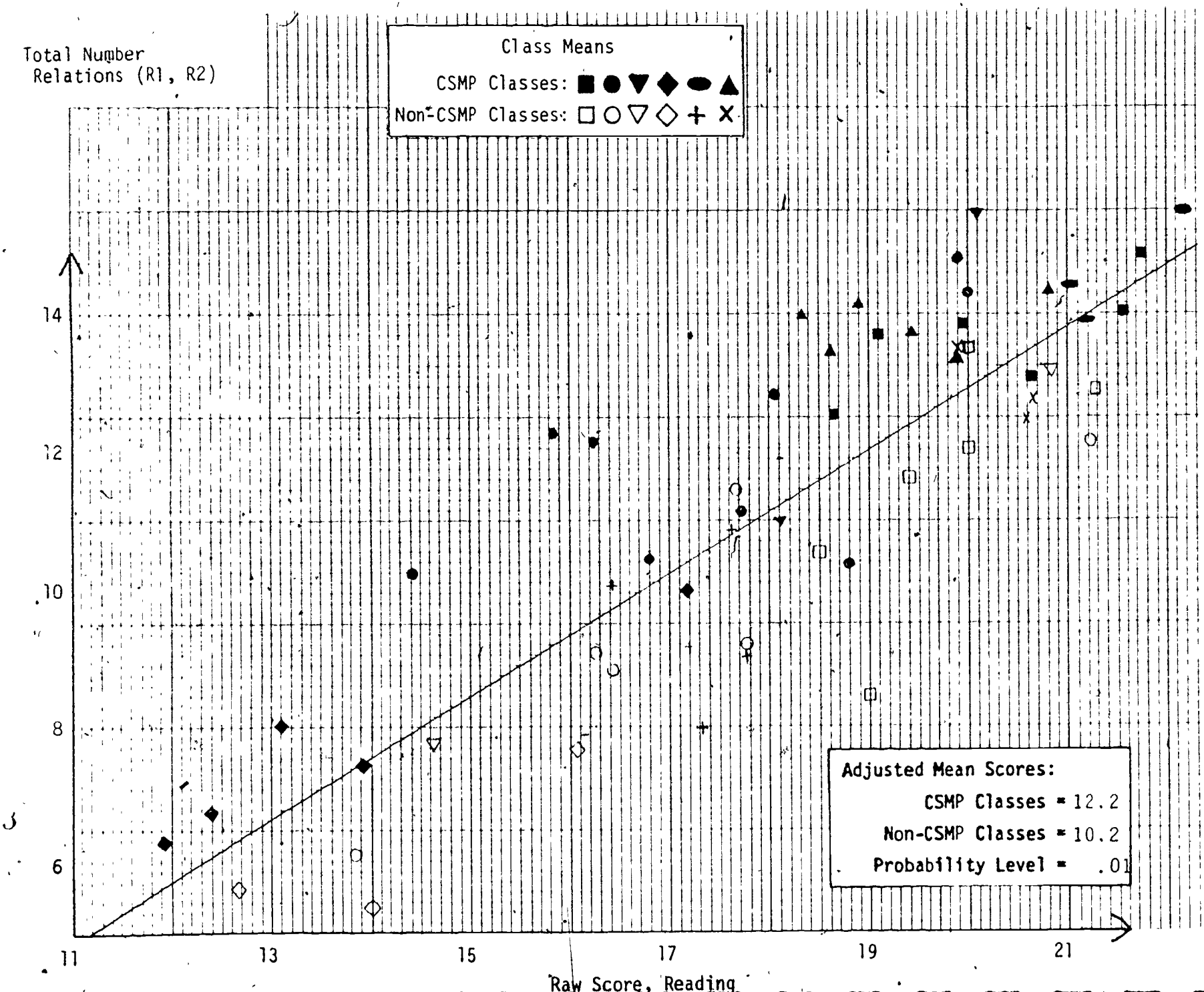
1. This scale was done after Scale R1 so that students were familiar with these "machine" formats. Three examples were done to illustrate how machines could be combined.

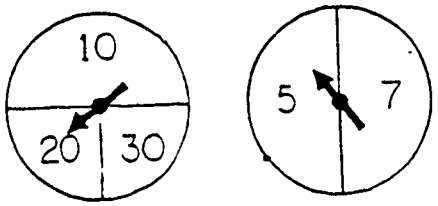
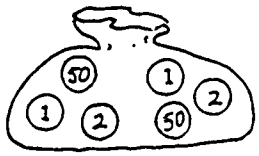
Scales R1, R2

Note:

1. The class mean statistics for the individual scales were as follows:

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
R1, Solving Functions	5.5	4.8	.01
R2, Using Functions	6.7	5.5	.01


Total Number
Relations (R1, R2)
Class Means
CSMP Classes:
Non-CSMP Classes:
14
12
10
8
6
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 12.2
Non-CSMP Classes = 10.2
Probability Level = .01

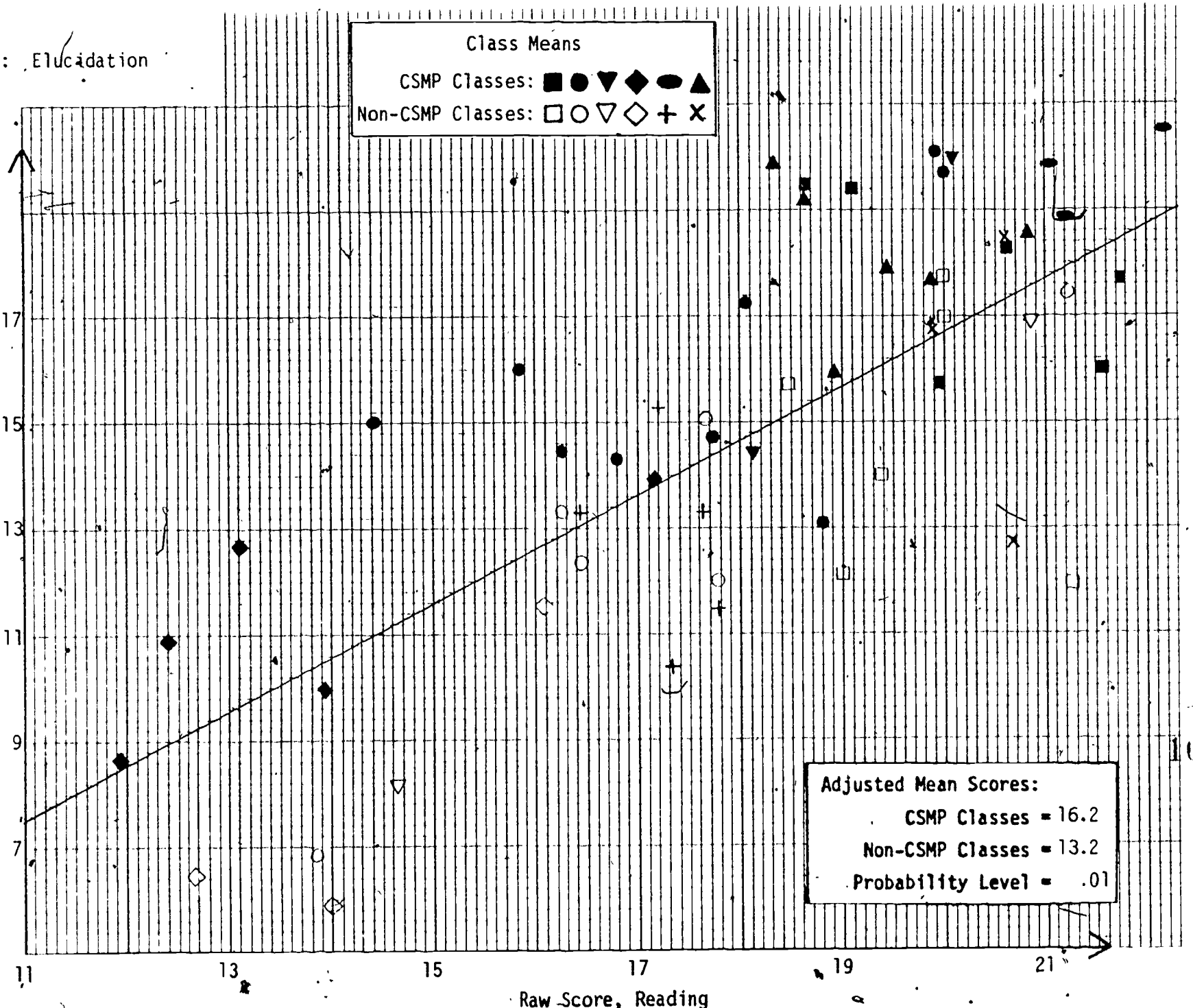
Test Items	Responses:	Percent Correct CSMP	Percent Correct Non-CSMP
① Spin both spinners at the same time. Your score is the total from the two spinners.  What are the possible total scores? 25,	15	91	87
	17	92	86
	27	90	85
	35	90	85
	37	88	83
	Sub Total	4.51	4.26
② Start at zero. Counting by? End up at 24. What could you be counting by? 1,	2	90	87
	3	84	78
	4	89	85
	6	88	83
	8	83	76
	12	71	57
	24	47	29
	Sub Total	5.52	4.95
③ Close your eyes. Pick out three balls. Add to get a total score.  What are the possible total scores? 52,	4	70	69
	5	67	65
	53	83	76
	54	75	66
	101	71	59
	102	71	56
	Sub Total	4.37	3.91
④ Multiple of 2 Multiple of 3 Smaller than 50 For what numbers are all three statements true? 24,	6	85	75
	12	84	73
	18	71	57
	30	55	42
	36	54	32
	42	42	21
	48	43	25
	Sub Total	4.34	3.25

Means By Ability Level

	1	2	3	4	All
CSMP	11.0	15.0	17.4	20.1	16.2
non-CSMP	8.5	11.2	14.6	17.3	13.2
t-test	3.6	5.2	4.5	5.5	8.1

	CSMP	Non-CSMP
Number of Students	626	511
KR20 Reliability	.87	.87
Mean Reading Score	18.2	18.0
Correlation, Reading versus Scale	.57	.53

1. The problems were reviewed, one at a time, with an explanation of the one given correct answer. Some time was allowed after each problem (with additional time allowed as needed after the last problem) for students to give as many correct answers as possible.


U1: Elucidation
Class Means
CSMP Classes:
Non-CSMP Classes:
17
15
13
11
9
7
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 16.2
Non-CSMP Classes = 13.2
Probability Level = .01

W2 Two Stage Word Problems

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
1. Shirts cost $10 each and ties cost $5 each. Peter bought 2 shirts and 3 ties. What was his total cost? _____	82	79
2. Joan starts with $40. Each week she spends $2. How much will she have left after 5 weeks? _____	61	55
3. The cost of gum is 3 pieces for 10¢. How many pieces can we buy for 40¢? _____	68	68
4. Pam gets 50¢ each week. She always spends 30¢ and saves the rest. How much will she save in 4 weeks? _____	69	61
5. On Saturday Amy and Susan made $13 selling lemonade. On Sunday they made $5. They put their money together and divided it evenly. How much did each girl get? _____	75	67
6. Jim has $10 in his bank now. Each week he will add $5 to his bank. In how many weeks will he have $30 in his bank? _____	57	50
7. John has 5¢ more than Tom. Ann has 3¢ less than Tom. If John has 20¢, how much does Ann have? _____	41	34

Means By Ability Level

	1	2	3	4	All
CSMP	1.3	2.2	2.5	2.9	2.3
non-CSMP	1.3	2.0	2.3	2.5	2.1
t-test	-0.1	0.5	0.6	1.4	1.4

	CSMP	Non-CSMP
Number of Students	315	256
KR20 Reliability	.78	.77
Mean Reading Score	18.0	17.8
Correlation, Reading versus Scale	.63	.56

W3 Three Stage Word Problems

Test Items	Percent Correct CSMP	Non-CSMP
1. Shirts cost \$10 each and ties cost \$5 each. Altogether Joe spent \$35 for shirts and ties. He bought 2 shirts. How many ties did he buy?_____	72	63
2. Joe puts boxes into piles. Each box is $\frac{1}{2}$ foot high. Each pile is 5 feet high. How many boxes does he need to make 3 piles? _____	39	35
3. Bill loads 6 boxes in 2 hours. Jonn loads 4 boxes in 2 hours. Together, how many boxes do they load in 6 hours? _____	29	31
4. Mary has 4 more marbles than Pete. Pete has 2 more marbles than Lisa. Lisa has 3 more marbles than Ed. If Mary has 20 marbles, how many does Ed have? _____	39	32
5. Monday, Tom ran 13 miles. Tuesday, he ran 8 miles. Wednesday, he ran some more. His average for the three days was 10 miles. How many miles did he run on Wednesday? _____	34	24

Means By Ability Level

	1	2	3	4	All
CSMP	0.4	0.8	1.1	1.8	1.1
non-CSMP	0.3	0.5	1.1	1.7	·0.9
t-test	0.8	2.1	0.0	0.5	1.3

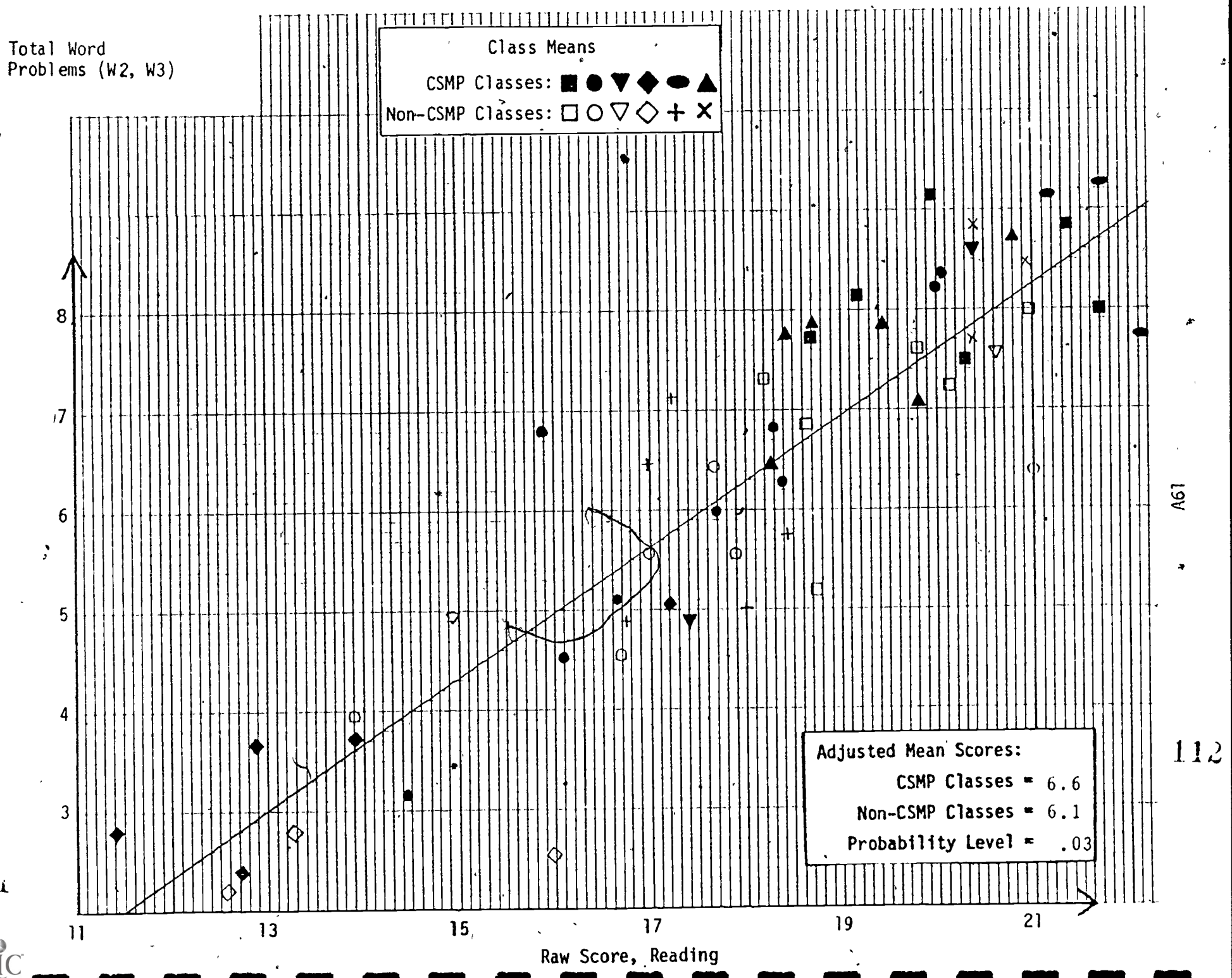
	CSMP	Non-CSMP
Number of Students	315	263
KR20 Reliability	.72	.69
Mean Reading Score	18.1	18.2
Correlation, Reading versus Scale	.60	.56

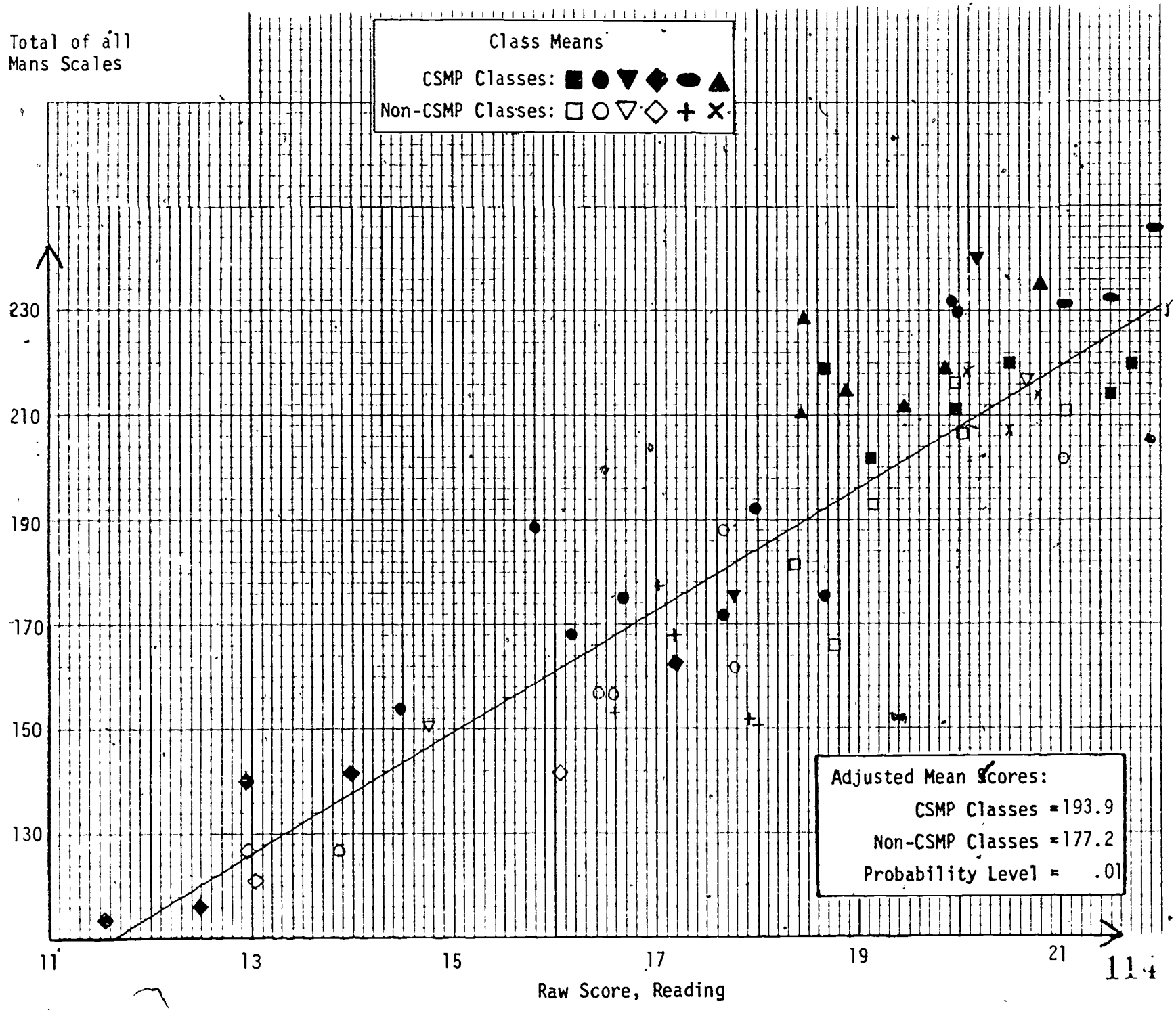
Scales W2,W3

1. The class mean statistics for the individual scales were as follows:

	Adjusted Class Means		
	CSMP	non-CSMP	p-value
W2, Two-Stage Word Problems	4.5	4.2	.11
W3, Three-Stage Word Problems	2.2	1.9	.12

It can be seen that neither of the scales, by itself, produced a significant difference. The total score on the sum of the two was, however, significantly in favor of CSMP as can be seen on the facing page.


Total Word
Problems (W2, W3)
Class Means
CSMP Classes:
Non-CSMP Classes:
8
7
6
5
4
3
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 6.6
Non-CSMP Classes = 6.1
Probability Level = .03


Total of all
Mans Scales
Class Means
CSMP Classes:
Non-CSMP Classes:
230
210
190
170
150
130
11
13
15
17
19
21
Raw Score, Reading
Adjusted Mean Scores:
CSMP Classes = 193.9
Non-CSMP Classes = 177.2
Probability Level = .01

A62

11.3

11.4

Appendix B

Comparison of Results Using Different Units of Analysis

An argument can be made for using different units of analyses. The question turns on whether one views normal variation in the treatment (i.e. curriculum) to be taking place at the student, class, school or district level. This report has used class as the unit of analysis, though data are presented in Appendix A for student-level analysis, which is a liberal interpretation. (With large numbers of students, relatively small differences produce significance; this is what occurred in this study as well.)

In order to determine whether different patterns of results - significant differences - would occur using the larger units of analysis, separate analyses of covariances were calculated for school and district as units of analysis.

These results are compared in the table, next page, with those obtained from the class-level analysis. The t-statistic is used in each case; this is appropriate because the t-statistic in these one-way analyses of covariances is simply the ratio of the CSMP-non-CSMP differences in adjusted means to the standard deviation (again adjusted for covariate) of the means.

Comparison of Results When Aggregation is at Class, School and District Level

Scale Category	Adjusted Class Means CSMP (n=31)	non-CSMP (n=25)	R*	Adjusted School Means CSMP (n=12)	non-CSMP (n=12)	R*	Adjusted District Means CSMP (n=6)	non-CSMP (n=6)	R*
Computation C1: CTBS Computation	34.9	34.3	0.8	35.4	34.8	0.7	34.8	34.1	0.7
Mental Arithmetic C3-C6	19.7	15.9	6.9	20.4	16.5	6.2	19.6	16.2	4.5
Estimation E2-E4,E6-E9	28.7	27.1	1.8	29.2	27.8	1.9	28.6	27.9	0.6
Fractions N3,N5-N10	33.0	31.1	2.3	34.0	32.1	2.0	33.3	31.8	1.1
Decimals N1,N4	11.1	8.5	7.1	11.5	8.8	6.8	11.2	8.7	5.6
Probability P1,P3	17.2	15.8	2.5	17.8	16.5	2.3	17.4	16.3	1.6
Number Relations R1,R2	12.2	10.2	6.7	12.5	10.5	5.9	12.0	10.3	6.2
Word Problems W1,W3	6.6	6.1	2.3	6.8	6.3	2.2	6.5	6.2	1.8
Weight Chart O1	6.5	6.5	0.9	6.6	6.6	0.2	6.4	6.4	0.5
Elucidation U1	16.2	13.2	5.5	16.7	13.6	5.4	16.2	13.2	8.7
Grand Total	193.9	177.2	4.5	200.3	181.9	4.8	194.3	179.3	4.1

*R = t-statistic for differences in means. If underlined, they are significant at .05, but different sizes of t are needed to reach significance because of different number of cases (56 classes, 24 schools, 12 districts).

It can be seen there are few differences obtained between the results from analysis at the class level versus school level. The means are higher at the school level (indicating perhaps a slight tendency for lower scoring classes to be more "densely" concentrated in schools). But the ratios, R (of difference in adjusted means to adjusted error) are very similar. The one case, Fractions, where they yield a different decision re significance is simply because of smaller n's in the t-table look-up. Thus, in effect, schools behave the way classes do; one interpretation is that, not unexpectedly, teachers within the same school implement the program in similar ways.

But analysis at the district level shows 3 categories to have a smaller ratio (Word Problems, Fractions and Probability). The differences are no longer significant and this finding has nothing to do with smaller n's. There is also a drop in R for Estimation. The corresponding interpretation is that within districts, there is considerable variation in the way individual schools implement the program.