

Mathematics Assessment Collaborative

MAC Report on the 2003 Tests

October 2, 2002



The Silicon Valley Mathematics Initiative

MAC & STAR technical report
September 24, 2003 by



Educational Data Systems
15850 CONCORD CIRCLE, SUITE A
MORGAN HILL, CA 95037
WWW.EDDATA.COM
(408) 776-7696

Acknowledgements

There are several contributors to the Mathematics Assessment Collaborative's Final Report for 2003. The Mathematics Assessment Collaborative is a program of the Silicon Valley Mathematics Initiative supported by the Noyce Foundation, the Santa Clara Valley Mathematics Project, and the twenty-seven member public school districts.

The Mathematics Assessment Collaborative contracts with CTB-McGraw Hill for rights to the Balanced Assessment in Mathematics Exams, designed by the Mathematics Assessment Resource Service from U.C. Berkeley, Michigan State and the Shell Centre at Nottingham, England. The senior designer of tasks and rubrics is Rita Crust.

The *Tools for Teachers* section of the report was researched, analyzed and written by Linda Fisher and the MAC Senior Trainers: Melissa Adams, Ellen Cafferata, Lucy de Anda, Sandy Devlin, Sally Keyes, Ford Long, Anne Paterson, Linda Shields, Jean Short, Margie Trainer, and Jeff Trubey.

Data analysis was conducted by the Educational Data Systems of Morgan Hill, California. The data analysis team included Sonyo Montelongo, Gabriela Guerrero, Al Shannon, Gurvinder Gill, Sharon White, Renuka Dhar, and Richard Smith.

Contributors to the Final Report

The Mathematics Assessment Collaborative – Linda Fisher, Director
The Noyce Foundation – Amy Gerstein, Ph.D., Executive Director
The Santa Clara Valley Mathematics Project – Joanne Rossi Becker, Ph.D., Director
The Silicon Valley Mathematics Initiative - David Foster, Program Director
The Mathematics Assessment Resource Service – Hugh Burkhart, Ph.D., Director
CTB-McGraw Hill – Alan Shienker, Program Manager
Educational Data Systems – Caroline Fahmy, President
Second Language Testing Incorporated - Charles W. Stansfield, Ph.D., President

Member Districts

Alum Rock School District	Los Gatos School District
Berryessa School District	Moreland School District
Brisbane School District	Morgan Hill School District
Burlingame School District	Mountain View-Whisman S.D.
Cambrian School District	New Haven School District
Campbell Union School District	Palo Alto School District
Cupertino School District	Portola Valley School District
East Side High School District	San Carlos School District
Evergreen School District	San Jose School District
Fremont Unified School District	Santa Clara School District
Gilroy Unified School District	San Mateo – Foster City S.D.
Jefferson Elementary School District	Saratoga School District
Laguna Salada School District	Union School District
	Woodside School District

Table of Contents	
Summary/Overview	pg. 4
MAC Examinations/Level Descriptions	pg. 6
Comparing Student Performance	pg. 9
NCTM Principles and Standards	pg. 17
Third Grade	
Blue Print	pg. 19
Overall Frequency Distribution	pg. 20
Demographic Data Analysis	pg. 21
Analysis of Data by Task	pg. 33
Fourth Grade	
Blue Print	pg. 44
Overall Frequency Distribution	pg. 45
Demographic Data Analysis	pg. 47
Analysis of Data by Task	pg. 57
Fifth Grade	
Blue Print	pg. 68
Overall Frequency Distribution	pg. 69
Demographic Data Analysis	pg. 70
Analysis of Data by Task	pg. 80
Sixth Grade	
Blue Print	pg. 91
Overall Frequency Distribution	pg. 92
Demographic Data Analysis	pg. 93
Analysis of Data by Task	pg. 103
Seventh Grade	
Blue Print	pg. 114
Overall Frequency Distribution	pg. 115
Demographic Data Analysis	pg. 117
Analysis of Data by Task	pg. 127
Eighth Grade	
Blue Print	pg. 138
Overall Frequency Distribution	pg. 139
Demographic Data Analysis	pg. 140
Analysis of Data by Task	pg. 151
Course One	
Blue Print	pg. 162
Overall Frequency Distribution	pg. 163
Demographic Data Analysis	pg. 164
Analysis of Data by Task	pg. 174
Course Two	
Blue Print	pg. 185
Overall Frequency Distribution	pg. 186
Demographic Data Analysis	pg. 187
Analysis of Data by Task	pg. 197
Appendix 1 – Star Data Comparisons	pg. 208
Appendix 2 – Scoring Reliability	pg. 234
Appendix 3 - Status of Data/Statistical Measures	pg. 243

Summary

The Silicon Valley Mathematics Assessment Collaborative uses performance assessment of the kind used to assess mathematics in most countries worldwide. The assessments and their information are then used for professional development and data for improving instruction.

This year MAC has a membership base of 27 districts. Over 200 teachers were trained in March to lead district scoring- sessions. This year we tested over 70,000 students from grades 3 through Course Two. MAC also developed a Practice Test and Pilot Test for Second Grade Students. The results come from 26 of MAC's 27 districts. One district used information for professional development purposes only.

This report describes test results across the collaborative as a whole. The results are based on the 2003 tests, which are the fifth annual cycle of tests administered by the collaborative. For each grade level the report provides a statistical analysis of the test as a whole, and the results task-by-task. The report is supplemented by the Tool Kit for Teachers, which provides a more in-depth look at student work, student misconceptions, instructional implications, and task commentary.

Statistical Analyses show that scoring has been done to high standards, across MAC, at all grade levels. Local scoring is working well. Data analysis provides clear evidence that MARS tests produce information on student attainment in mathematics, which complements rather than duplicates evidence from other tests (e.g. CST, CAT-6, SAT-9).

Over the five years of testing, there have been year-on-year gains in student performance at grades 3, 4 and 5. Student performance continues to improve and excel in elementary grades, where most of our coaches focus their efforts. Data shows that students at middle grades are not making the same kind of progress.

Districts, with support from the Collaborative, are encouraged to provide continued professional development for teachers to help use this information and data to improve mathematics education. In an effort to make data more accessible to teachers, MAC will provide Toolkit Workshops in October of 2003 to help development teacher leaders to work at school sites with the Tools for Teachers. The SCVMP provided an August Coaching Institute to focus on key ideas. That work will be followed by a series of MAC Professional Development Sessions during the school year, focusing on student thinking and misconceptions and developing better instructional strategies. Dr. Becker of the SCVMP with support of a state Math Professional Development Institute grant facilitates funding for this professional development. The Robert Noyce Foundation also provides mathematics coaches for many school districts to help teachers in their classroom.

Overview

MAC, The Silicon Valley Mathematics Assessment Collaborative, is composed of twenty-seven public elementary and high school districts in San Mateo, Santa Clara, and Alameda Counties, in association with the Santa Clara Valley Math Project and the Robert Noyce Foundation. The purpose of the Collaborative is to produce, score and report on tests of student performance in mathematics. In addition, the

Collaborative provides professional development opportunities for teachers and teacher leaders, helping to build the capacity of all member districts.

The assessment of students' mathematics requires far more than the administration of psychometric tests. Assessment tasks exemplify the core mathematics to be attained, therefore examples are given within the report to illustrate the educational ambitions of MAC. Through the process of giving performance exams, the teaching community of the member districts is building a common vision of the mathematical standards for each grade, a clear set of expectations for good student work, and how mathematical standards and expectations change as students move through the grades. This shared vision promotes a truly standard-based mathematics curriculum.

For the year 2002-2003, MAC has collaborated with and become a part of a larger team. For the last three years MAC has worked with the MARS (Mathematics Assessment Resource Service) team to develop practice tests, performance tests, and professional development materials. MARS has also been a valuable resource for scoring training. MARS is now working with CTB McGraw Hill to make tests available nationally. MAC trainers hosted the meeting to work with a national team in finalizing the rubrics used for 2003.

Through grants written by Dr. Becker, professor at San Jose State University, many more teachers were able to benefit from the professional development sessions this past year. During the first year of the project, much of the staff development centered on the concepts of performance assessment and scoring protocols. This last year the focus of the staff development revolved around looking at MAC results, students misconceptions, and learning instructional strategies for overcoming those misconceptions. The depth of the content is having a more direct impact on classroom instruction. As a professional community, MAC districts are becoming more sophisticated in looking at and learning from student work and developing better tools for helping teachers improve instruction.

New This Year

Under the leadership of Melissa Adams and David Foster, a committee of MAC teachers developed a set of MAC Core Ideas for Kindergarten through Second Grade. The committee developed a 5 task Practice Performance Assessment and a 5 task Assessment for Second Grade students. In May of 2003, the Mathematics Assessment Collaborative (MAC) piloted a Second Grade Performance Assessment. The assessments were administered over two consecutive days in early May. Eleven districts participated in the piloting of the assessment. In all, 1800 students were tested. The data from these assessment tasks has been published in a Second Grade Toolkit. The toolkit has been distributed to all districts in the Collaborative to help direct their professional development.

Funding from the Noyce Foundation covered costs for two additional leaders to assist in the staff development workshops, which allowed these sessions to have grade-level meetings as well as look at general issues. Melissa Adams assisted with the primary grades. Sally Keyes worked with three-five teachers. In addition the MAC senior trainers volunteered to help with the Professional Development Sessions.

David Foster and Linda Fisher were asked to speak at the National Academy of Sciences in Washington, D.C. They were seeking experts from the international community, educational research community, and state testing groups to develop a position paper on best practices in assessment. The draft version of the paper, *Assessment in Support of Instruction and Learning: Bridging the Gap between Large Scale and Classroom Assessments* will be published in November 2003.

This year the Collaborative also produced a prototype for individual student reports. These reports were field-tested in some districts. More development needs to go into the initial data entry to make the reporting process timelier.

The MAC Examinations

The purpose of the MAC project is to develop a high stakes test that districts can use for multiple measures on student performance in mathematics. While many tests are currently available to measure how students can compute under time pressure, we need a test that measures students' ability to apply mathematics in problem situations and explain their thinking. These tests are designed to measure the student performance in using the key curriculum content of their grade level in tackling worthwhile tasks.

The concept is for the testing process to be transparent to all the stakeholders. Teachers and students know in advance what will be covered. Likewise they are familiar with the type of rubric to be used and they have the opportunity to have a practice test before the exam, so that the format will not be new and confusing. This helps to demystify the testing process, so that it can come closer to measuring what students know and are able to do.

Research has also shown that a well-designed test can help drive quality instruction. This test is designed to show how well the student understands the mathematical tools learned and how well they can be applied in problem solving situations, nonroutine tasks. An unexpected, but significant benefit is the opportunity to increase content knowledge of teachers and develop common, consistent expectations of good student performance. By having clearly defined Core Ideas for each grade level, tests matched to those Core Ideas, and a consistent criteria for measuring student performance, MAC is truly building meaningful, working mathematical standards.

Designing the tests: The range, variety and balance of the tasks are critical to the success or otherwise of this endeavor. The task sets used here included the kinds of tasks that are found on the National Assessment of Education (NAEP) and on the Third International Mathematics and Science Study (TIMSS); because these assessments are constrained in a number of practical ways, MAC was able to achieve better balance.

The assessment covers a familiar range of mathematical topics: Number Operations and Properties, Geometry, Measurement, Patterns and Functions, Data Analysis, Probability, Algebraic Symbols and Properties as well as Mathematical Representation, Communication, Connections, Reasoning and Proof, and Problem Solving. For MAC, teachers met to determine which core ideas comprised the mathematical foci specific to each grade level. An item related to each core idea was selected to sample the overall range of student thinking. In addition the content dimension, the tasks were chosen to achieve reasonable balance in other important dimensions of performance – types of task, and of practical context, problem solving process, etc.

Scoring student work: Examining student responses to such tasks provides teachers with the opportunity to look at student thinking, see common misunderstandings, and make adjustments in curriculum and instruction. In the scoring and in-service process, teachers can start to see the road map in mathematical instruction and how expectations and sophistication in responses is stepped up as students move up in grades.

At each scoring session, teachers and students were first trained on the point scoring process and then on a rubric specific to the problem they would be scoring.

After discussion of the question and its rubric, scorers were given 5 sample student responses to score. Their scores were then compared to the score determined by the “senior assessor” – a MARS consultant or a MAC scoring facilitator. The differences between how they scored compared to the established score were then discussed, to clear up their thinking and help everyone develop the same internal standards and understandings. After this process was completed, another set of 10 Standardizing Papers were given to them for practice. If these papers were correctly scored against the established score, scorers were then allowed to start correcting student work.

After they completed their first folder of student papers, they traded with another scorer for checking. Differences were discussed and settled with the help of scoring facilitators. Periodically throughout the day random papers were double scored to check for accuracy and consistency.

Looking at the reliability analysis in the appendix, it appears that this training process provides us with relatively accurate and consistent scores across many different local scoring sessions.

Setting the Level boundaries: In international assessment, the setting of standards through decisions on level boundaries (“cut scores”) is through informed professional judgments on the basis of three kinds of input. The following procedure, based on international practices outlined above, was used in the MAC tests:

- Analysis of the tasks and rubrics in relation to the Standards. For each task, an analysis of the elements of performance and the rubric suggests a typical score that represents the core mathematics of the task. This points to provisional level boundaries.
- Each member of the Awarding Committee (A Board of top international Educators and Educational Leaders brought together by CTB McGraw Hill) makes holistic judgments of a sample of student papers around provisional level boundaries. The question for each student paper is “Does this student performance, viewed as a whole, meet the standard?”
- Statistical information on the score distributions for each task, and for the test, informs these professional judgments. It includes cross-comparison data between different tasks, which links the demand levels of parallel tests, from one year to another, for example.

Performance Level Descriptions

4 – Achieves Standards at a High Level

The student performed at a high level on the tasks and consistently achieved the standards.

3 – Performance at Standards

The student performed at a level that met the standards.

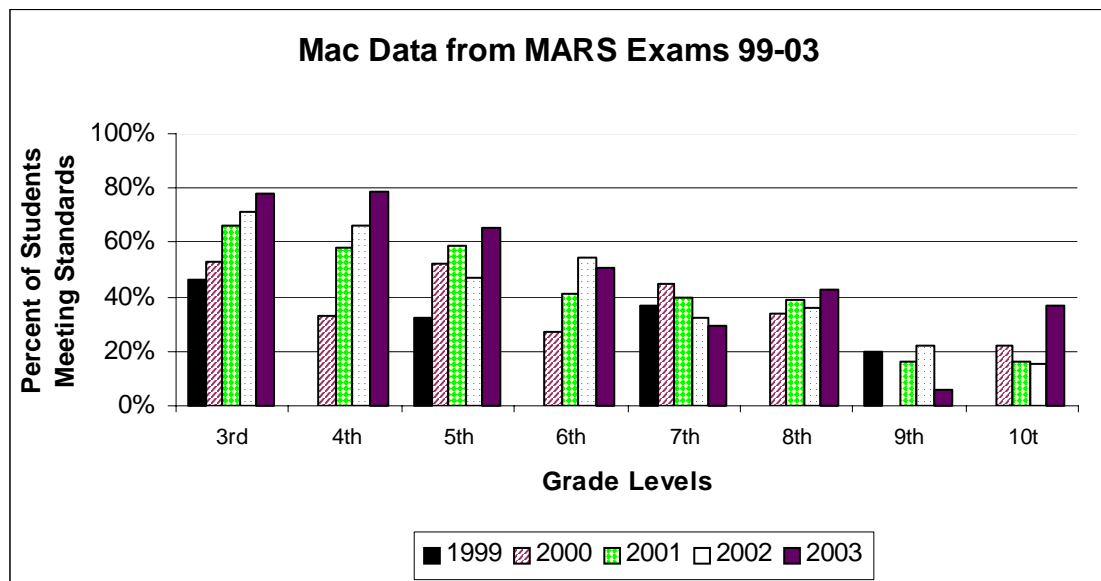
2 – Performance Below Standards

The student showed some evidence of performing at the level of the standards, but overall the performance did not meet the standards.

1 – Demonstrates Minimal Success

The student showed minimal evidence of performing at the level of the standards. Much of the response was either fragmented or non-existent.

MAC Data from MARS Exams 99-03



Comparing Student Performance

Data are available from 1999-2002 at grades 3, 5, and 7. Data are available from 2000-2002 at grades 4, 6, 8, and 10. Data for grade 9 are available for 1999, 2001, and 2002. These data are analyzed here.

Different tests are given each year in MAC, and while standards remain the same each year, no attempt is made to create parallel test forms. It is impossible, therefore, to compare performance in adjacent years by examining raw scores.

Test performance is described each year in terms of levels of performance. A careful process is used to allocate points on tests to levels, and so changes from year to year can be made in terms of the numbers of students achieving different performance levels.

A year-on-year comparison is not as straightforward as it might first appear. For example, if the demographic composition of samples changed dramatically between 2000 and 2001, then observed differences in student performance may be due to this factor, and not to factors such as changes in teaching approaches, increased familiarity with previously unfamiliar task types, and so on.

Over the four years of testing, there have been considerable changes in the sample of students tested. In particular, large numbers of students have been tested in later years. The analyses, which follow, should, therefore, be treated with caution. A better understanding of changes in student attainment will be gained by detailed analysis of District Reports, where the nature of any demographic changes can be assessed with some accuracy.

Comparing Student Performance

Grade 3

	1999		2000		2001	
	Count	Percentage	Count	Percentage	Count	Percentage
Level 1	2136	24.8%	2081	18.9%	1422	14.3%
Level 2	2478	28.8%	3169	28.7%	2010	20.2%
Level 3	2941	34.2%	3800	34.5%	4047	40.7%
Level 4	1054	12.2%	1975	17.9%	2458	24.7%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	1174	8.2%	1256	8.6%
Level 2	2878	20.2%	2117	14.5%
Level 3	4411	30.9%	4815	33.0%
Level 4	5804	40.7%	6421	44.0%

	N	Mean	Std. Deviation
1999	8609	2.34	0.98
2000	11025	2.51	0.99
2001	9937	2.76	0.98
2002	14267	3.04	0.97
2003	14609	3.12	0.96

Again, there has been a steady increase in the percentage of students reaching Level 4, although the increase from 2002 to 2003 is not as large as the previous year's increase. The percentage of students reaching level 3 also increased slightly between 2002 (30.9%) and 2003 (33.0%).

The increase in mean performance between 2002 and 2003 is statistically significant and corresponds to an effect size of .08 standard deviation units.

Grade 4

	2000		2001	
	Count	Percentage	Count	Percentage
Level 1	235	17.3%	1185	18.0%
Level 2	446	32.8%	1519	23.1%
Level 3	469	34.5%	2590	39.4%
Level 4	209	15.4%	1275	19.4%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	504	7.6%	717	7.6%
Level 2	1772	26.6%	1304	13.9%
Level 3	2492	37.4%	3417	36.4%
Level 4	1889	28.4%	3961	42.1%

	N	Mean	Std. Deviation
2000	1359	2.48	0.95
2001	6569	2.6	0.99
2002	6657	2.87	0.91
2003	9399	3.13	0.92

Until 2003, there had been a steady increase in the percentage of students reaching Level 4, while the percentage of students reaching Level 3 remained relatively constant. Between 2002 and 2003, the percentage increase of students reaching Level 4 was just over 48% (from 28.4% to 42.1%), and the percentage of students reaching Level 3 decreased by 1%.

The increase in mean performance between 2002 and 2003 is statistically significant and corresponds to an effect size of .28 standard deviation units.

Grade 5

	1999		2000		2001	
	Count	Percentage	Count	Percentage	Count	Percentage
Level 1	2141	24.9%	1165	11.8%	1234	11.3%
Level 2	3678	42.8%	3798	38.3%	3259	29.8%
Level 3	1901	22.1%	3111	31.4%	3984	36.4%
Level 4	870	10.1%	1836	18.5%	2459	22.5%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	2601	21.9%	1479	10.3%
Level 2	3697	31.1%	3562	24.7%
Level 3	2959	24.9%	4715	32.7%
Level 4	2615	22.0%	4645	32.3%

	N	Mean	Std. Deviation
1999	8590	2.17	0.92
2000	9910	2.57	0.92
2001	10936	2.7	0.94
2002	11872	2.47	1.06
2003	14401	2.87	0.98

The percentage of students reaching Level 4 increased fairly steadily from 1999 until 2001 then remained about the same in 2002. The percentage of students reaching Level 4 in 2003 (32.3%) increased by almost 47% from 2002 (22.0%). The percentage of students reaching Level 3 also increased between 2002 (24.9%) and 2003 (32.7%).

The increase in mean performance between 2002 and 2003 is statistically significant and corresponds to an effect size of .41 standard deviation units.

Grade 6

	2000		2001	
	Count	Percentage	Count	Percentage
Level 1	166	11.1%	983	15.7%
Level 2	456	30.6%	2682	42.7%
Level 3	572	38.4%	1773	28.2%
Level 4	296	19.9%	841	13.4%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	1615	20.9%	2095	21.2%
Level 2	1898	24.5%	2796	28.3%
Level 3	2765	35.7%	3207	32.5%
Level 4	1462	18.9%	1767	17.9%

	N	Mean	Std. Deviation
2000	1490	2.67	0.92
2001	6279	2.39	0.91
2002	7740	2.53	1.02
2003	9865	2.47	1.02

The percentage of students reaching the highest levels of performance has shown a down-up-down pattern through time; it decreased from 2000 to 2001, increased from 2001 to 2002, and then decreased again between 2002 and 2003.

The decrease in mean performance between 2002 and 2003 is statistically significant and corresponds to an effect size of .06 standard deviation units.

Grade 7

	1999		2000		2001	
	Count	Percentage	Count	Percentage	Count	Percentage
Level 1	1246	22.6%	817	11.5%	2465	22.6%
Level 2	2255	40.9%	3489	49.2%	4083	37.4%
Level 3	1313	23.8%	1884	26.6%	2869	26.3%
Level 4	695	12.6%	900	12.7%	1490	13.7%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	2752	30.8%	3677	31.4%
Level 2	3317	37.1%	4583	39.1%
Level 3	2162	24.2%	1896	16.2%
Level 4	715	8.0%	1559	13.3%

	N	Mean	Std. Deviation
1999	5509	2.26	0.95
2000	7090	2.4	0.85
2001	10907	2.31	0.97
2002	8946	2.09	0.93
2003	11715	2.11	1.00

By inspection, the percentage of students reaching the highest levels of performance remained relatively constant from 1999 to 2002, decreasing slightly in 2002. Although the percentage of students achieving Level 4 increased between 2002 (8.0%) and 2003 (13.3%), the percentage of students reaching Level 3 decreased from 24.2% in 2002 to 16.2% in 2003. Thus, the percentage of students achieving either a Level 3 or Level 4 decreased between 2002 and 2003 although the mean increased slightly.

The increase in mean performance is *not* statistically significant and corresponds to an effect size of .02 standard deviation units.

Grade 8

	2000		2001	
	Count	Percentage	Count	Percentage
Level 1	244	23.4%	951	18.9%
Level 2	410	39.3%	2109	41.9%
Level 3	282	27.0%	1544	30.7%
Level 4	107	10.3%	428	8.5%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	1751	27.1%	2509	30.7%
Level 2	2390	37.0%	2158	26.4%
Level 3	1495	23.1%	1556	19.0%
Level 4	826	12.8%	1955	23.9%

	N	Mean	Std. Deviation
2000	1043	2.24	0.93
2001	5032	2.29	0.87
2002	6462	2.22	0.98
2003	8178	2.36	1.15

No clear patterns emerge in the data, although the combined percentages of students reaching Levels 3 and 4 increased from 35.9% in 2002 to 42.9% in 2003. The increase in mean performance is statistically significant and corresponds to an effect size of .12 standard deviation units.

Course One

	1999		2001	
	Count	Percentage	Count	Percentage
Level 1	639	36.9%	179	22.3%
Level 2	754	43.5%	496	61.8%
Level 3	219	12.6%	121	15.1%
Level 4	122	7.0%	6	0.7%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	109	17.5%	1139	65.6%
Level 2	378	60.8%	501	28.9%
Level 3	122	19.6%	84	4.8%
Level 4	13	2.1%	12	0.7%

	N	Mean	Std. Deviation
2000	1734	1.90	0.88
2001	802	1.94	0.63
2002	622	2.06	0.67
2003	1736	1.41	0.62

No clear patterns emerge in the 1999 to 2002 data. The percentage of students achieving either a Level 3 or Level 4 decreased dramatically from 2002 (21.7%) to 2003 (5.5%). The decrease in mean performance is statistically significant and corresponds to an effect size of 1.06 standard deviation units.

Course Two

	2000		2001	
	Count	Percentage	Count	Percentage
Level 1	910	33.3%	369	31.9%
Level 2	1219	44.6%	603	52.1%
Level 3	480	17.6%	163	14.1%
Level 4	124	4.5%	23	2.0%

	2002		2003	
	Count	Percentage	Count	Percentage
Level 1	109	44.5%	57	24.7%
Level 2	98	40.0%	89	38.5%
Level 3	23	9.4%	53	22.9%
Level 4	15	6.1%	32	13.9%

	N	Mean	Std. Deviation
2000	2733	1.93	0.83
2001	1158	1.86	0.72
2002	245	1.77	0.86
2003	231	2.26	0.98

Far fewer students took the exam in 2002 and 2003 than in 2000 and 2001, so comparisons of the first two test administrations with the second two should be treated with extreme caution. The percentage of students achieving a Level 3 or Level 4 increased between 2002 (15.5%) and 2003 (36.8%).

The increase in mean performance between 2002 and 2003 is statistically significant and corresponds to an effect size of .50 standard deviation units.

Principles and Standards for School Mathematics NCTM

Problem Solving: Instructional programs should enable all students to-

- Build new mathematical knowledge through problem solving;
- Solve problems that arise in mathematics and in other contexts;
- Apply and adapt a variety of appropriate strategies to solve problems;
- Monitor and reflect on the process of mathematical problem solving.

Reasoning and Proof: Instructional programs should enable all students to-

- Recognize reasoning and proof as fundamental aspects of mathematics;
- Make and investigate mathematical conjectures;
- Develop and evaluate mathematical arguments and proofs;
- Select and use various types of reasoning and methods of proof.

Communication: Instructional programs should enable all students to-

- Organize and consolidate their mathematical thinking through communication;
- Communicate their mathematical thinking coherently and clearly to peers, teachers, and others;
- Analyze and evaluate the mathematical thinking and strategies of others;
- Use the language of mathematics to express mathematical ideas precisely.

Connections: Instructional programs should enable all students to-

- Recognize and use connections among mathematical ideas;
- Understand how mathematical ideas interconnect and build on one another to produce a coherent whole;
- Recognize and apply mathematics in contexts outside of mathematics.

Representation: Instructional programs should enable all students to-

- Create and use representations to organize, record, and communicate mathematical ideas;
- Select, apply, and translate among mathematical representations to solve problems;
- Use representations to model and interpret physical, social, and mathematical phenomena.

Number and Operations: Instructional programs should enable all students to-

- Understand numbers, ways of representing numbers, relationships among numbers, and number systems;
- Understand meanings of operations and how they relate to one another;
- Compute fluently and make reasonable estimates.

Algebra: Instructional programs should enable all students to-

- Understand patterns, relations, and functions;
- Represent and analyze mathematical situations and structures using algebraic symbols;
- Use mathematical models to represent and understand quantitative relationships;
- Analyze change in various contexts.

Geometry: Instructional programs should enable all students to-

- Analyze characteristics and properties of two- and three-dimensional geometric shapes and develop mathematical arguments about geometric relationships;
- Specify locations and describe spatial relationships using coordinate geometry and other representational systems;
- Apply transformations and use symmetry to analyze mathematical situations;
- Use visualization, spatial reasoning, and geometric modeling to solve problems.

Measurement: Instructional programs should enable all students to-

- Understand measurable attributes of objects and the units, systems, and processes of measurement;
- Apply appropriate techniques, tools, and formulas to determine measurements.

Data Analysis and Probability Standard: Instructional programs should enable all students to-

- Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them;
- Select and use appropriate statistical methods to analyze data;
- Develop and evaluate inferences and predictions that are based on data;
- Understand and apply basic concepts of probability

**Blue Print for Third Grade MAC
Alignment of Tasks to NCTM Content and Process Standards**

	Even and Odd	Addition Trains	Vending Machines	Cherie's Shapes	Patchwork Quilt
NUM	P	P		S	
ALG				P	
GEO				S	P
MEAS			S		
DATA			P		
PS		*	*		*
REAS	*	*	*	*	
COMM	*			*	*
CONN				*	
REP	*	*	*	*	*

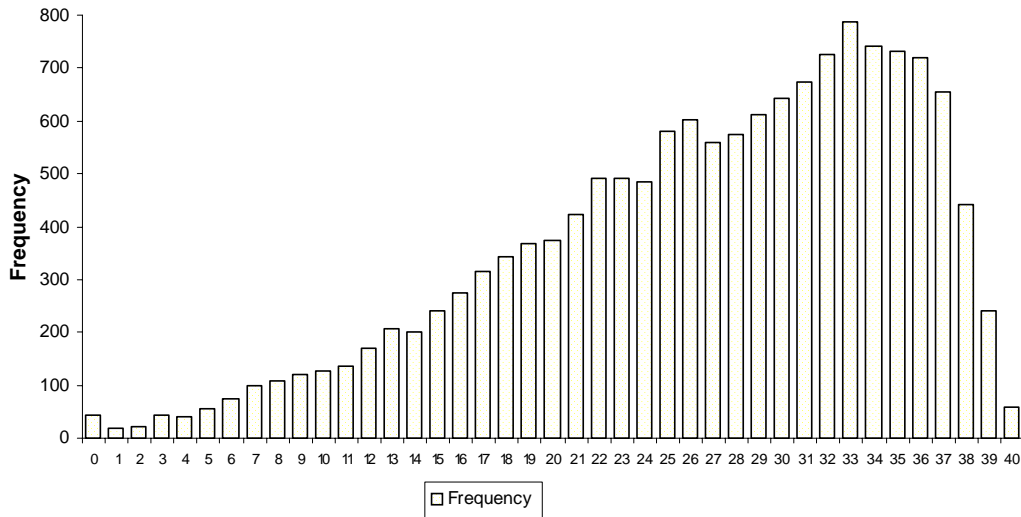
NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

- P denotes Primary NCTM Content Standard
- S denotes Secondary NCTM Content Standard
- denotes NCTM Process Standard

Overall Frequency Distribution by Total Score

Grade 3

Mean=26.60; S.D.=8.45



Level Frequency Distribution Chart and Frequency Distribution
Number of students tested 2003: Grade 3: 14,267

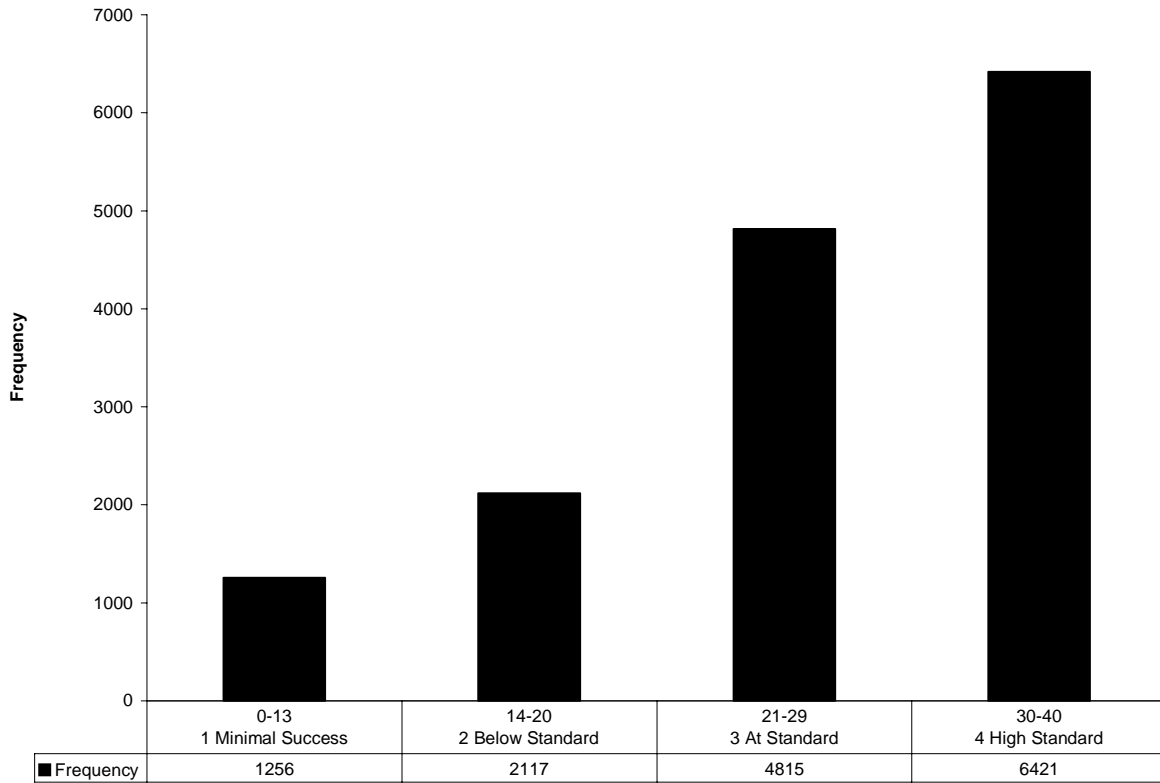
Grade 3 1999 - 2001

Level	% at ('99)	% at least ('99)	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	25%	100%	19%	100%	14%	100%
2	29%	75%	28%	81%	20%	86%
3	34%	46%	35%	53%	41%	66%
4	12%	12%	18%	18%	25%	25%

Grade 3 2002 - 2003

Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	8%	100%	9%	100%
2	20%	92%	14%	91%
3	31%	72%	33%	77%
4	41%	41%	44%	44%

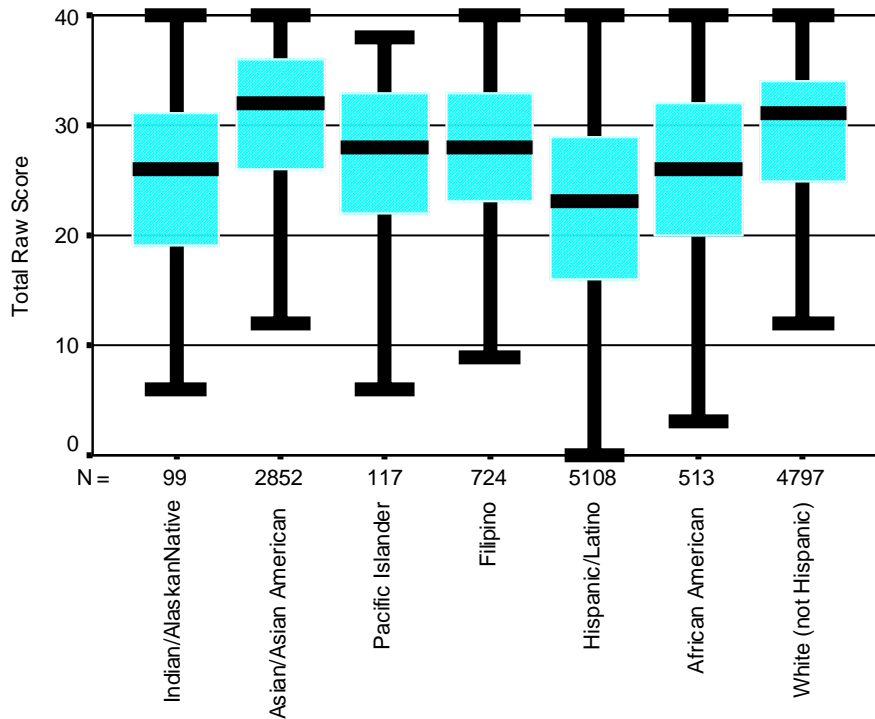
Distribution by Performance Levels



Grade 3 – Ethnicity

The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values* within a category lie between the highest and lowest horizontal bars.

Grade 3 - Ethnicity



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Distribution of sampling means - Grade 3 - Ethnicity

In this section, test scores are compared across different ethnic groups¹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if the scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than group B. When the two scores overlap, then there is no significant difference between the groups.

The scores of Indian/Alaskan Native students are significantly lower than those of Asian/Asian American and White students, and are significantly higher than those of Hispanic/Latino students. There are no significant differences between Indian/Alaskan Native students and any other ethnic group.

The scores of Asian/Asian American students are significantly higher than those of students in all other ethnic groups.

¹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

The scores of Pacific Islander students are significantly lower than those of Asian/Asian American and White students. The scores of Pacific Islander students are significantly higher than those of Hispanic/Latino students. There are no significant differences between the scores of Pacific Islander students and any other ethnic group.

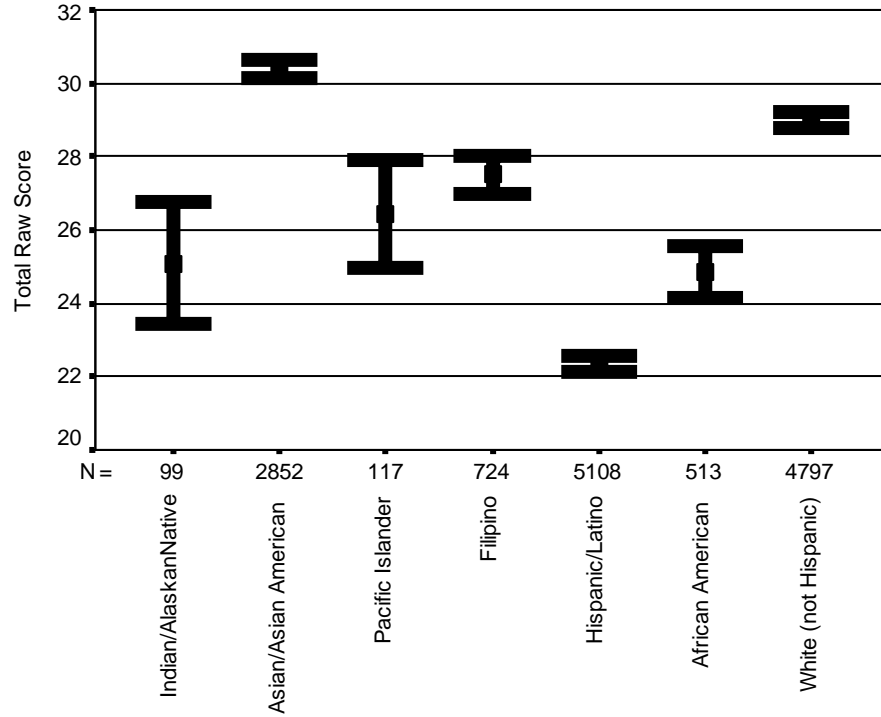
The scores of Filipino students are significantly lower than those of Asian/Asian American and White students, and significantly higher than those of Hispanic/Latino and African American students. There are no significant differences between the scores of Filipino students and Indian/Alaskan Native students, or between the scores of Filipino students and Pacific Islanders.

The scores of Hispanic/Latino students are significantly lower than the scores of students from any of the other ethnic groups.

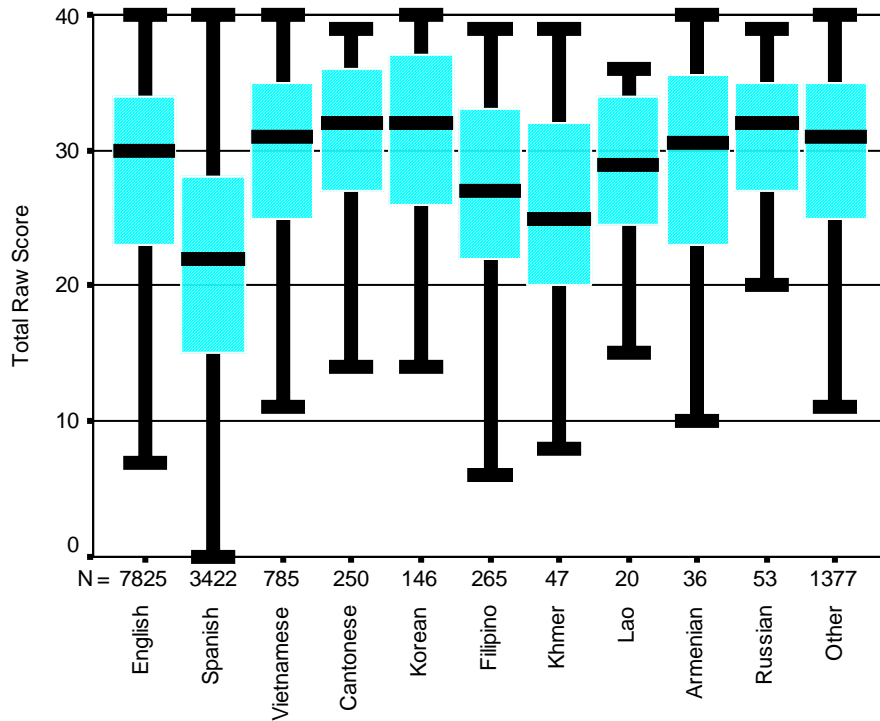
The scores of African American students are significantly lower than those of Asian/Asian American, Filipino, and White students. The scores of African American students were significantly higher than those of the Hispanic/Latino students. There are no significant differences between the scores of African American students and any other ethnic group.

The scores of White students are significantly higher than those of Indian/Alaskan Native, Pacific Islander, Filipino, Hispanic/Latino, and African American students. The scores of White students were significantly lower than those of the Asian/Asian American students.

Grade 3 – Ethnicity - Distribution of sampling means



Grade 3 - Home Language



Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian,

Japanese, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 3 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home². One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with a home language of Vietnamese, Cantonese, Korean, or “Other”. The scores of students with English as a home language are significantly higher than those of students with a home language of Spanish. There is no significant difference between scores of students with a home language of English, and those with a home language of Filipino, Khmer, Lao, Armenian, or Russian.

The scores of students with Spanish as a home language are significantly lower than the scores of students with any other home language.

The scores of students with Vietnamese as a home language are significantly higher than those of students with English, Spanish, Filipino, and Khmer as a home language. There are no significant differences between the scores of students with Vietnamese as a home language and any other group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, Filipino, and Khmer as a home language. There are no significant differences between the scores of students with Cantonese as a home language and any other group.

The scores of students with Korean as a home language are significantly higher than those of students with English, Spanish, Filipino, and Khmer as a home language. There are no significant differences between the scores of students with Korean as a home language and any other group.

The scores of students with Filipino as a home language are significantly lower than those of students with Vietnamese, Cantonese, Korean, and “Other” as a home language, and significantly higher than those of students with Spanish as a home language. There are no significant differences between the scores of students with Filipino as a home language and any other group.

² Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

The scores of students with Khmer as a home language are significantly lower than those of students with Vietnamese, Cantonese, Korean, and “Other” as a home language. The scores of students with Khmer as a home language are significantly higher than those of students with Spanish as their home language. There are no significant differences between the scores of students with Khmer as a home language and any other group.

The scores of students with Lao as a home language are significantly higher than the scores of students with Spanish as their home language, but there are no other significant differences with any other home language category.

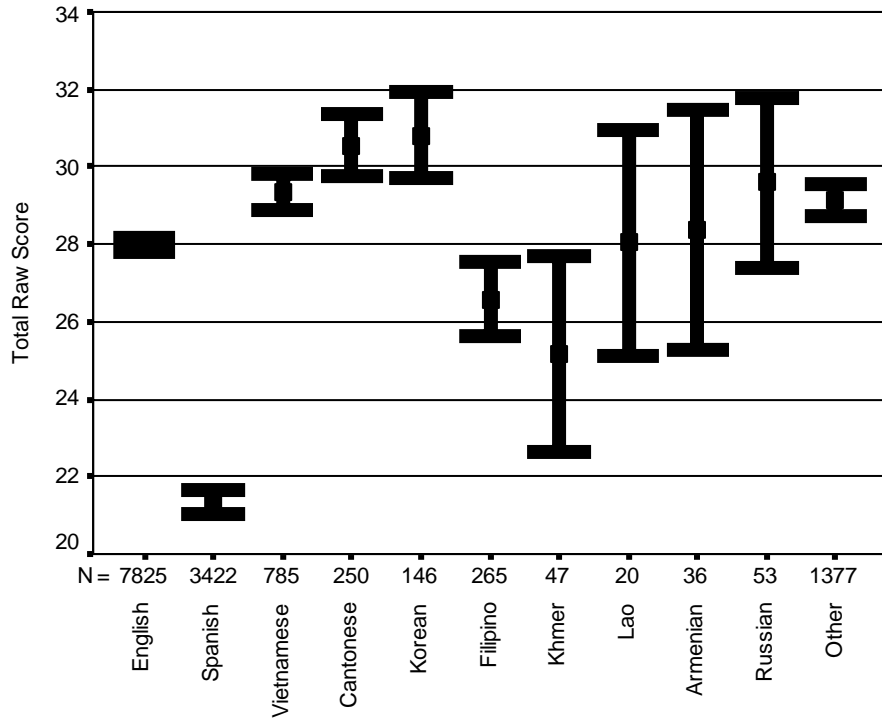
The scores of students with Armenian as a home language are significantly higher than the scores of students with Spanish as their home language, but there are no other significant differences with any other language group.

The scores of students with Russian as a home language are significantly higher than those of students with Spanish as a home language. There are no significant differences between the scores of students with Russian as a home language and any other language group.

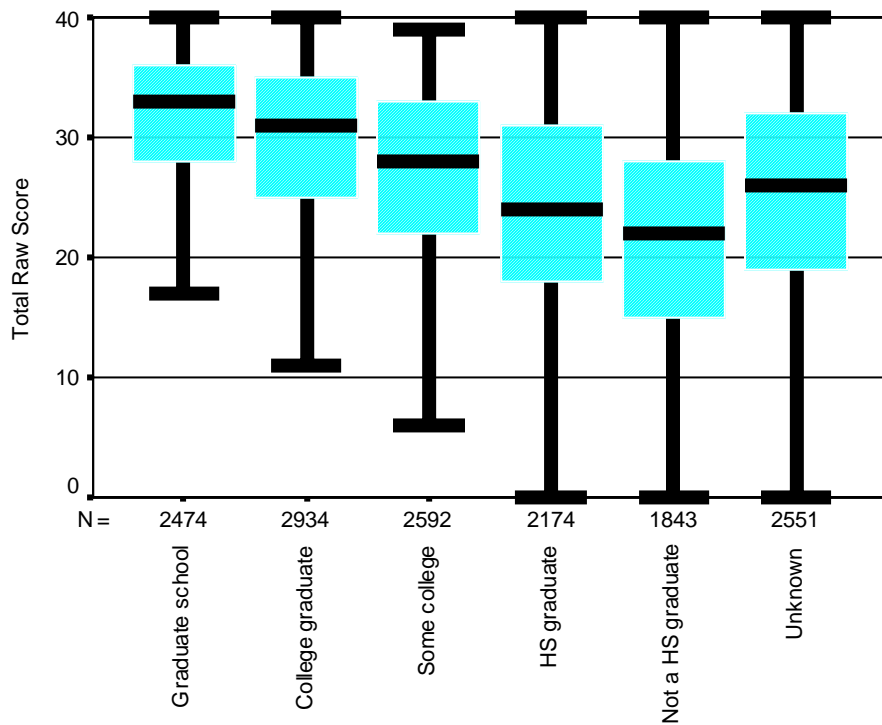
The scores of students with “Other” as a home language are significantly higher than those of students with English, Spanish, Filipino, and Khmer as a home language. There are no other significant differences between the scores of students with “Other” as a home language and any other language group.

Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 3 -Home Language



Grade 3 - Parent Education



Distribution of sampling means -Grade 3 -Parent Education

In this section, test scores are compared across groups of different levels of parent education³. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate school education are significantly higher than those of students in all other Parent Education categories.

The scores of students whose parents have a college education are significantly lower than those whose parents have a graduate education, and significantly higher than all other groups.

The scores of students whose parents have some college education are significantly higher than those of students whose parents are not High School graduates and those whose parents are High School graduates, and significantly lower than those whose parents have a college education or a graduate education.

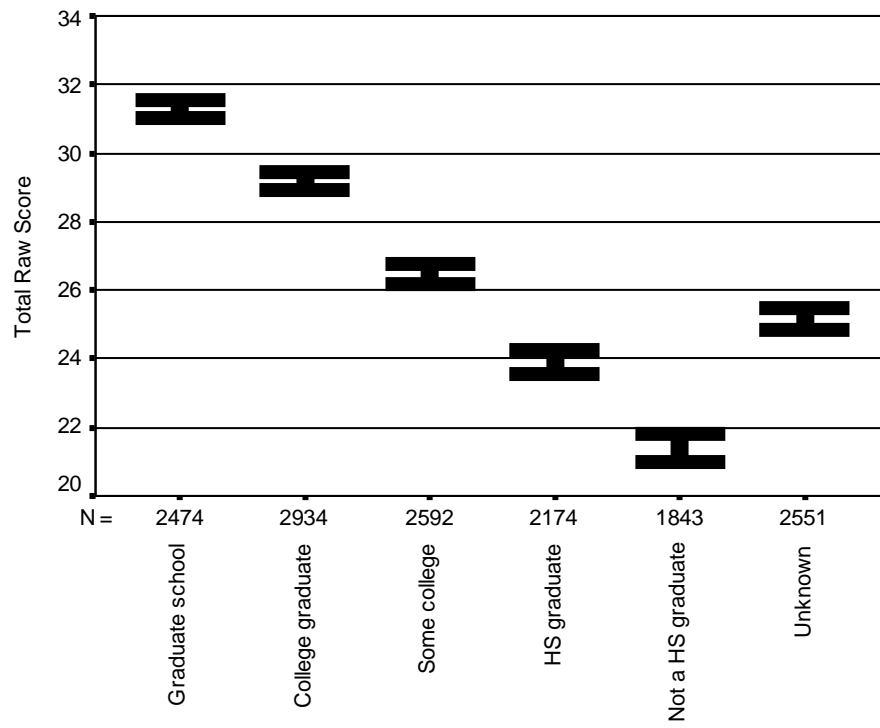
The scores of students whose parents are High School graduates are significantly higher than those of students whose parents are not High School graduates, and significantly lower than all other Parent Education categories.

The scores of students whose parents are not High School graduates are significantly lower than the scores of students in all other Parent Education categories.

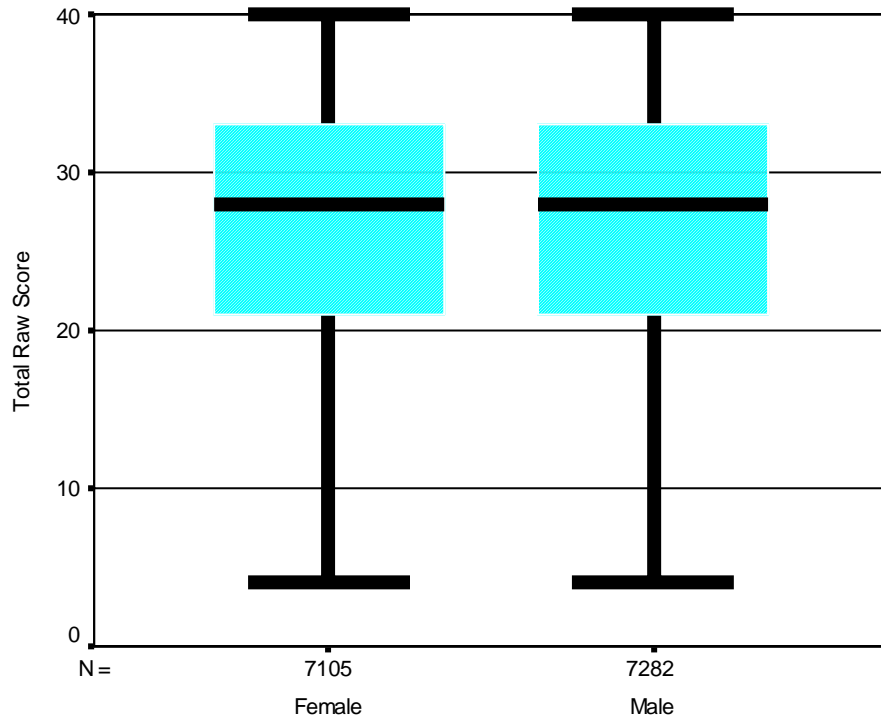
The scores of students whose parents' education is unknown are significantly higher than those of students whose parents are not High School graduates and those whose parents are High School graduates, and significantly lower than those whose parents have a college education or a graduate education.

³ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 3 - Parent Education



Grade 3 - Gender



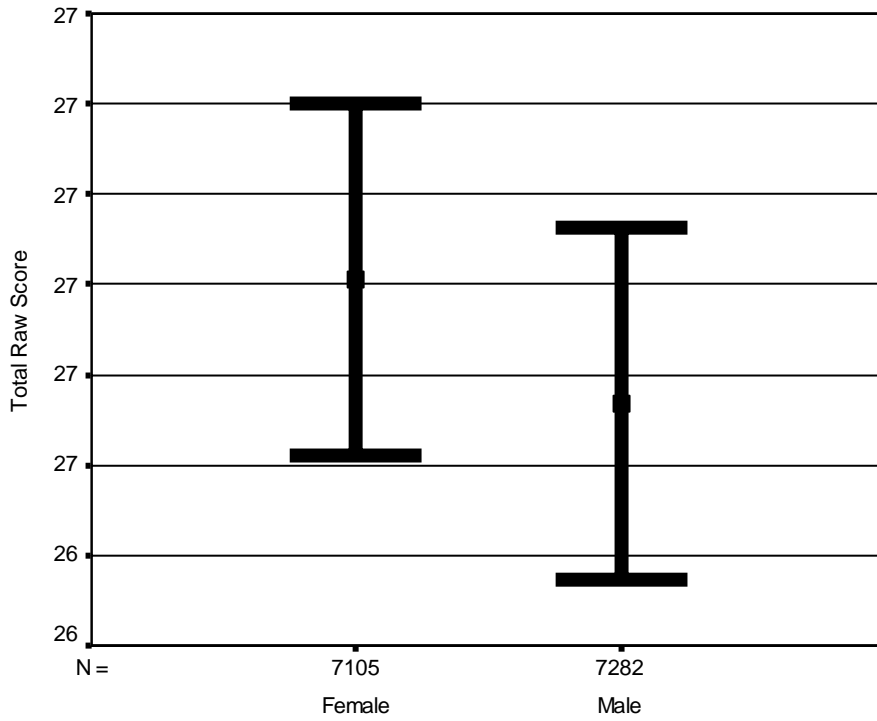
Distribution of sampling means - Grade 3 - Gender

In this section, test scores are compared across gender⁴. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

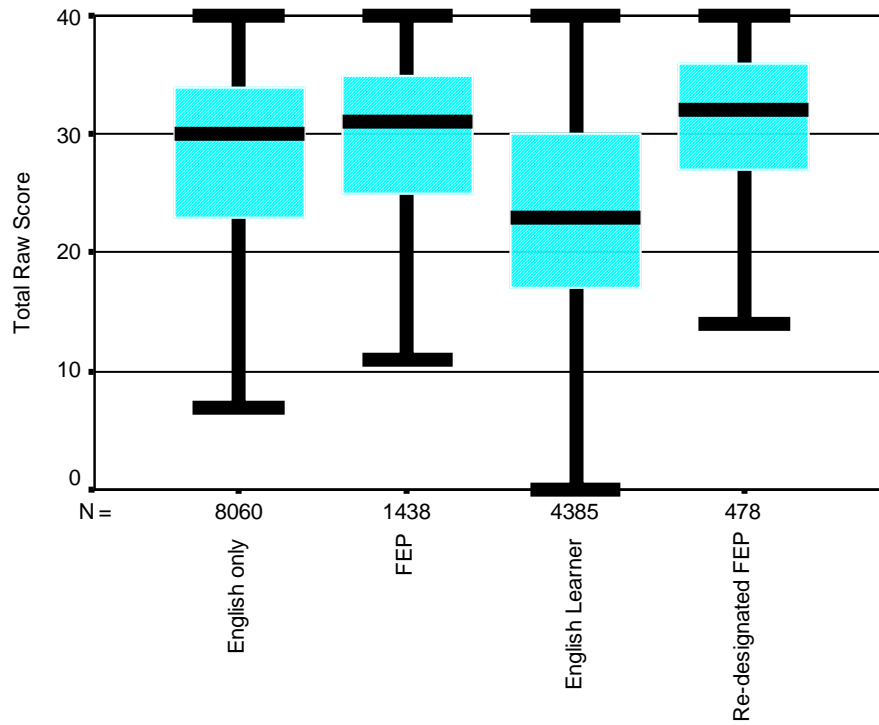
There is no statistically significant difference in the scores of females and males.

⁴ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 3 - Gender



Grade 3 - Language Fluency



Distribution of sampling means -Grade 3 - Language Fluency

In this section, test scores are compared across groups of different language fluency⁵. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly lower than those of students described as Full English Proficiency (FEP) and those described as Re-designated FEP, and significantly higher than those students described as English Learners.

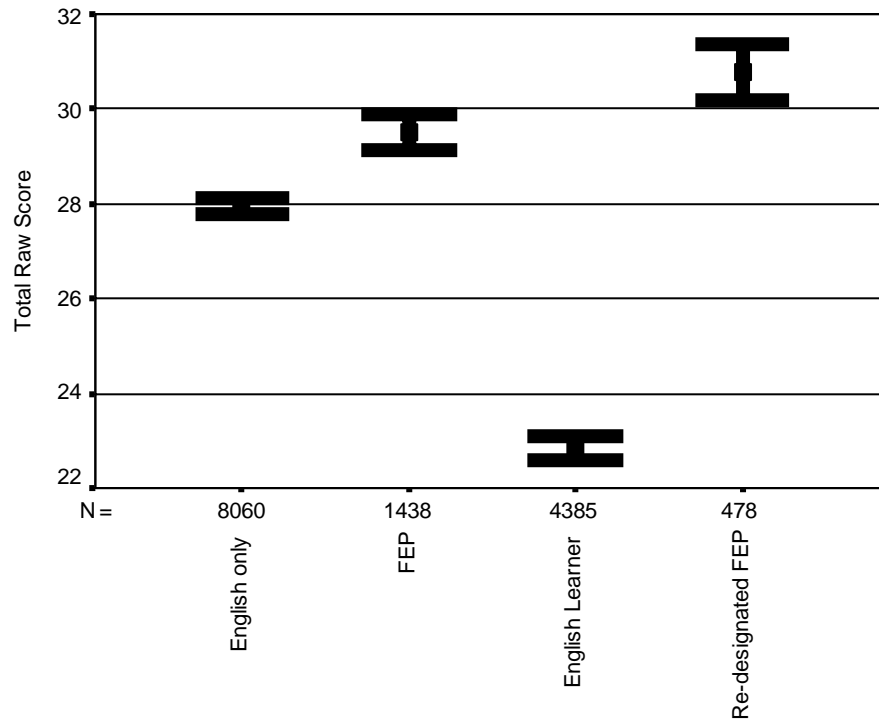
The scores of students in the FEP category are significantly higher than those of students in the English only and English Learner categories. The scores of students in the FEP category are significantly lower than the scores of students in the re-designated FEP group.

The scores of students in the English Learner category are significantly lower than those of the students in all other Language Fluency categories.

The scores of students in the Re-designated FEP category are significantly higher than those of students in all other categories of Language Fluency.

⁵ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

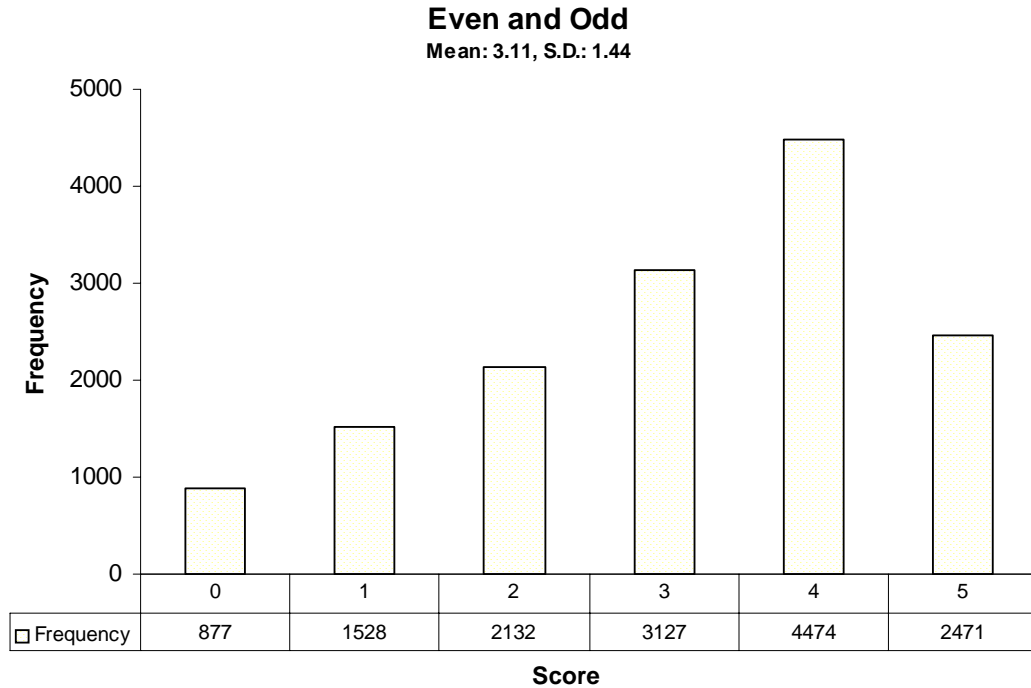
Distribution of sampling mean -Grade 3 -Language Fluency



3rd grade Task 1 Even and Odd Numbers

Student Task	Solve to complete given number sentences using even and odd addends. Explain effects of adding odd numbers.
Core Idea 1 Number Properties	Understand numbers, ways of representing numbers, relationships among numbers, and number systems. <ul style="list-style-type: none"> Develop a sense of whole numbers and represent and use them in flexible ways, including relating, composing, and decomposing numbers
Core Idea 2 Number Operations	Understand the meanings of operations and how they relate to each other, make reasonable estimates and compute fluently. <ul style="list-style-type: none"> Understand the effects of adding and subtracting whole numbers Develop fluency in adding and subtracting whole numbers

Grade 3 – Even and Odd



Score:	0	1	2	3	4	5
% < =	6.0%	16.5%	31.1%	52.5%	83.1%	100.0%
% > =	100.0%	94.0%	83.5%	68.9%	47.5%	16.9%

The maximum score available for this task is 5 points.

The cut score for a level 3 response is 3 points.

Most students (about 94%) could add even numbers to get a sum of 8. Many students (about 70%) could add even numbers to get 8 and find at least one combination of odd numbers to sum to 15. More than half the students could find multiple combinations for making 8 and 15, while paying attention to the constraints in the task. More than 15% could also give a justification about adding odd numbers. About 6% of the students scored no points on this task.

Based on teacher observations, this is what third grade students seemed to know and be able to do:

- Add accurately.
- Find multiple combinations to get a specific sum.

Areas of difficulty for third graders, third grade students struggled with:

- Using constraints, such as “only even numbers” or “only odd numbers”
- Making a generalization about adding odd and even numbers
- Writing a mathematical justification
- Noting which combinations had already been used

Implications for Instruction:

While many students were very successful with this problem, some students struggled with the directions. It is important for students to pick out for themselves the constraints of the problems, to develop the skill of finding key ideas. Students need to have a variety of rich problems that require them to pay attention to several demands such as equals 15, use only addition, and use only odd numbers. Students need practice with solving problems, where the numbers used are limited in some way (greater than 5, odd, multiples of 3, etc.) to develop their problem solving skills as well as computational practice. Some students are still unclear about whether “0” is odd or even. This is an interesting question to ask students and have them develop convincing arguments about. Students also need opportunities to explore how numbers work and explain the patterns made by using certain kinds of numbers. The task gives students the opportunity not just to repeat information like an odd plus an odd equals an even, but to construct a convincing argument connecting that information with why 3 numbers are needed to get an odd answer. However students need many different kinds of opportunities to stretch their thinking and ability to make justifications and should not be limited to just problems about odd and even. To help them develop the ability to make convincing arguments, teachers might pose questions like “Will that always work? Why? How can we show why?...”

3rd grade

Task 2

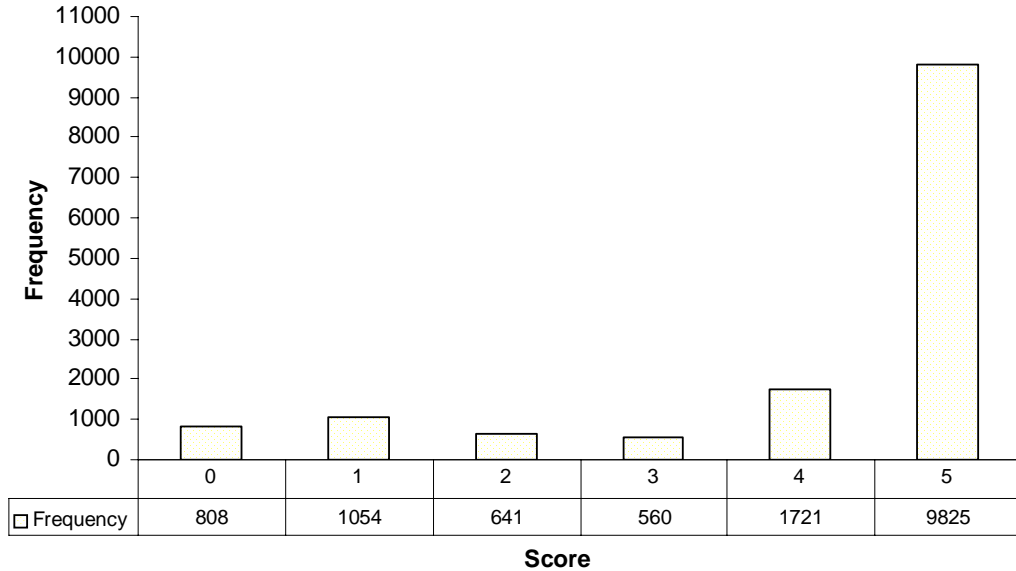
Addition Trains

Student Task	Given a sum and specific numbers, find the addends to correctly complete a number sentence.
Core Idea 2 Number Operations	Understand the meanings of operations and how they relate to each other, make reasonable estimates and compute fluently. <ul style="list-style-type: none">• Understand the effects of adding and subtracting whole numbers• Develop fluency in adding and subtracting whole numbers• Develop a sense of whole numbers and represent and use them in flexible ways, including relating, composing, and decomposing numbers.
Core Idea 3 Patterns Functions And Algebra	Understand patterns and use mathematical models to represent and to understand qualitative and quantitative relationships. <ul style="list-style-type: none">• Illustrate general principles and properties of operations using specific numbers

Grade 3 – Addition Trains

Addition Trains

Mean: 4.11, S.D.: 1.55



Score:	0	1	2	3	4	5
% <=	5.5%	12.7%	17.1%	21.0%	32.7%	100.0%
% >=	100.0%	94.5%	87.3%	82.9%	79.0%	67.3%

The maximum score available on this task is 5 points.

The cut score for a level 3 response is 3 points.

Most students (about 95%) could use both constraints to find a sum equal to eleven and a sum equal to 18. Students meeting standards (83%) could find addition trains for three of the given target numbers. More than 65% of the students could meet all the demands of the task. A little more than 5% of the students scored no points on this task.

Based on teacher observations, this is what third grade students seemed to be able to do:

- Students were most successful on using a combination of 3’s and 5’s to make 18.
- Students were good at using addition to make the correct sums, even if they did not use the constraint.
- Many students recognized that problems 1 through 3 were multiples of 5 plus one.

Areas of difficulty for third graders, third grade students struggled with:

- Many students had difficulty using the constraints of “either 3 or 5” as the addends.

Implications for Instruction:

Students need practice with interpreting problems and their constraints. Teachers might ask questions like, “How many different sums less than 20 can you make using two addends?” Students need to have fluency in addition of more than two whole numbers. Students would often use numbers of convenience, rather than trying to get the answer using only the choices given. Students need practice with solving problems, where the numbers used are limited in some way (greater than 5, odd, multiples of 3, etc.) to develop their flexibility with numbers. It is important for them not only to be able to add, but also to be able to work backwards from the total to finding a variety of ways to get that total. Students need to be able to reason about the number system and explore how different numbers can combine to make a given total. Part of problem solving is the ability to find solutions with limited choices.

3rd grade

Task 3

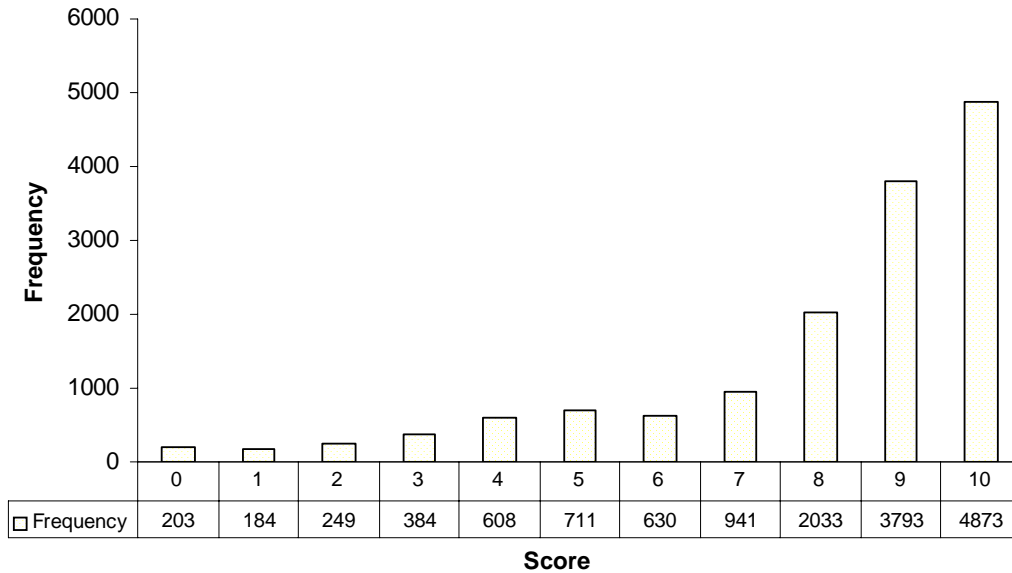
Vending Machines

Student Task	Interpret a pictograph and add additional information. Complete a chart with information provided.
Core Idea 5 Data Analysis	Collect, organize, display, and interpret data about themselves and their surroundings. <ul style="list-style-type: none">• Represent and interpret data using pictograph• Describe important features of a set of data
Core Idea 3 Patterns, Functions, and Algebra	Understand patterns and use mathematical models to represent and to understand qualitative and quantitative relationships. <ul style="list-style-type: none">• Show quantitative change• Use concrete, pictorial, and verbal representations to develop an understanding of invented and conventional symbolic notations• Model problem situations using representations such as graphs and tables to draw conclusions

Grade 3 – Vending Machines

Vending Machines

Mean: 8.03, S.D.: 2.39



Score:	0	1	2	3	4	5	6	7	8	9	10
% <=	1.4%	2.6%	4.4%	7.0%	11.1%	16.0%	20.3%	26.8%	40.7%	66.6%	100.0%
% >=	100.0%	98.6%	97.4%	95.6%	93.0%	88.9%	84.0%	79.7%	73.2%	59.3%	33.4%

The maximum score available on this task is 10 points.
 The cut score for a level 3 response is 6 points.

Most students (about 89%) could answer simple questions about the drink graph and find the total number of drinks. More than half the students (about 60%) could answer questions about the drinks graph, explain how they found the total number of drinks, and correctly fill out most of the graph for chips. More than 30% could also fill in an additional apple juice on the drink graph to meet all the demands of the task.

Approximately 1% of the students scored no points on this task.

Based on teacher observations, this is what third grade students seemed to know and be able to do:

- Students could correctly answer questions from a graph (e.g. “How many cola drinks?”, “What kind of drink sold the most?”, and “How many drinks were sold in all?”)
- Most students could do comparison subtraction.
- Most students could successfully either add data to the drink graph or do some correct portion of making their own graph.

Areas of difficulty for third graders, third grade students struggled with:

- Explaining their thinking for finding the total number of drinks.
- Matching symbols with quantity to represent data on a pictograph.
- Adding data to an existing graph.

Implications for Instruction:

Students need experience interpreting and completing pictographs and bar graphs. They need more practice using symbols that represent different quantities of the same item to complete a graph. If students are using bar graphs, then it is important that the scale on the vertical axis not always be one. Students need to have a variety of experiences with data, learning to make their own graphs as well as completing and interpreting the graphs of others. Asking students a variety of questions about a graph, such as, “How many more. . . ?” or “How many total . . . are represented on the graph?” are important to help them understand and analyze what information the graph is representing. Students also need to have frequent opportunities to explain how they got their answers, learning to describe which numbers they chose and also why they picked a particular operation. Hearing, evaluating, and comparing the explanations of others allows them to learn what is valued in a good response and to self-correct and to improve the quality of answers they make for themselves.

3rd grade

Task 4

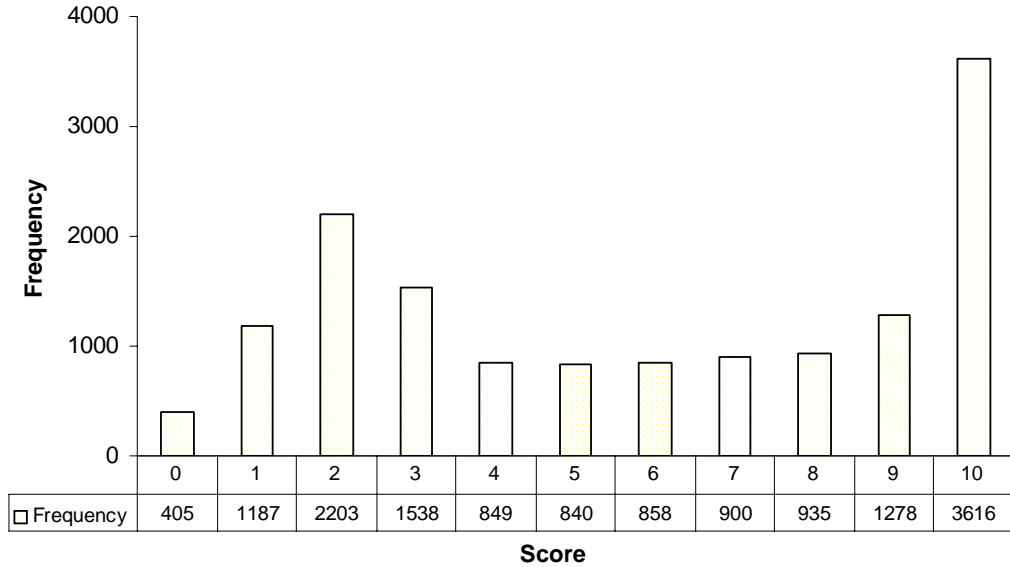
Cherie’s Shapes

Student Task	Name some simple geometric shapes. Use multiplication and/or division to complete given patterns and design their own repeating patterns.
Core Idea 3 Patterns, Functions, and Algebra	Understand patterns and use mathematical models to represent and to understand qualitative and quantitative relationships. <ul style="list-style-type: none"> • Describe and extend geometric patterns • Represent and analyze patterns using words
Core Idea 2 Number Operations	Understand the meanings of operations and how they relate to each other, make reasonable estimates, and compute fluently. <ul style="list-style-type: none"> • Understand multiplication as counting by multiples and division as forming equal groups
Core Idea 4 Geometry and Measurement	Recognize and use characteristics, properties and relationships of two-dimensional geometric shapes <ul style="list-style-type: none"> • Identify and compare attributes of two-dimensional shapes and develop vocabulary to describe the attributes

Grade 3 – Cherie’s Shapes

Cherie's Shapes

Mean: 5.78, 3.42



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	2.8%	10.9%	26.0%	36.5%	42.3%	48.1%	53.9%	60.1%	66.5%	75.2%	100.0%
% > =	100.0%	97.2%	89.1%	74.0%	63.5%	57.7%	51.9%	46.1%	57.7%	33.5%	24.8%

The maximum score available on this task is 10 points.

The cut score for a level 3 response is 4 points.

Most students (approximately 90%) could give all the correct shape names or name all the shapes but hexagon and draw their own pattern. More than half the students could find how many times all the patterns were needed, including their own design. Many of the students with scores of 7,8, or 9 struggled with explaining their thinking in part 2 and 3. Almost 25% of the students could meet all the demands of the task. Less than 3% of the students scored no points on this task.

Based on teacher observations, this is what third grade students seemed to know and be able to do:

- Identify and name geometric shapes.
- Use counting, repeated addition, or multiplication to find the number of times equal groups will fit into a larger number.
- Design a pattern that will fit evenly into 30.

Areas of difficulty for third graders, third grade students struggled with:

- Naming hexagons
- Matching a number sentence to the situation of how many groups will fit into a larger shape, understanding number operation of multiplication and its inverse.
- Many students have not made the transition from thinking additively to thinking multiplicatively.
- Students are still having trouble with part/whole relationships. If they get the correct number sentence, like $5 \times 6 = 30$, they often think the answer is it will fit 30 times.
- Recognition of shapes with nonstandard orientations.

Instructional Implications:

Students need to be able to create repeating patterns being mindful of multiple constraints. At this grade level students are beginning to use simple division or multiplication to solve problems. It is important for them to understand the meaning behind the operation and not just the computation. There is a fundamental idea here about working with equal groups and understanding part/whole relationships. A weakness here leads to huge problems when trying to understand fractions, proportional reasoning and scaling, probability and other mathematical topics. Students need to have a deep understanding of the operation and multiplicative relationships to successfully move forward in their mathematical thinking. Students at this grade level should be able to work flexibly with number fact families. Being able to see the relationship between multiplication facts and using that relationship to solve problems is a big idea for students.

3rd grade

Task 5

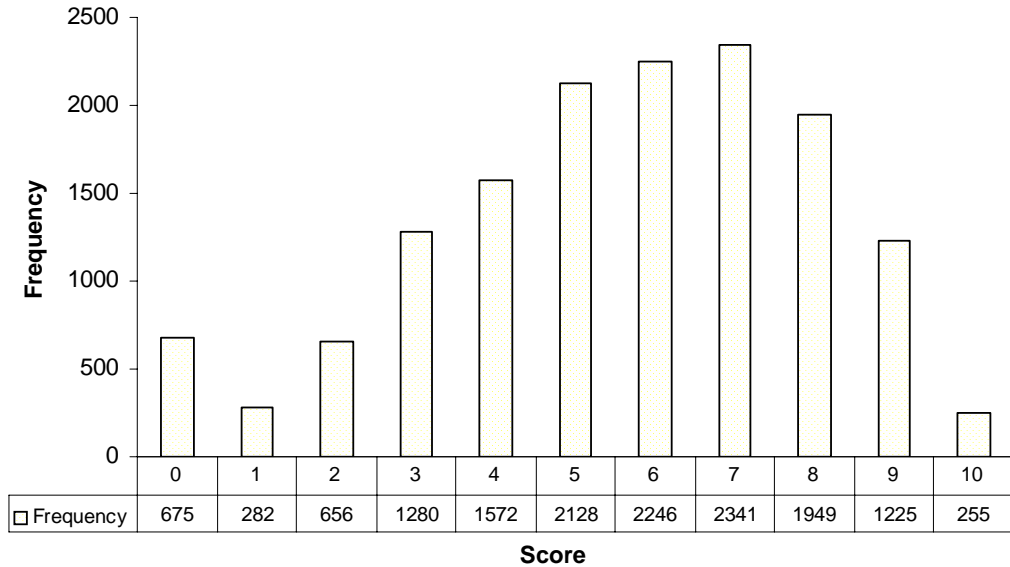
Patchwork Quilt

Student Task	Recognize and describe properties of simple shapes. Mark two lines of symmetry on a given design. Create a design with two lines of symmetry.
Core Idea 4 Geometry and Measurement	<p>Recognize and use characteristics, properties, ad relationships of two-dimensional geometric shapes</p> <ul style="list-style-type: none"> • Identify and compare attributes of two-dimensional shapes and develop vocabulary to describe the attributes • Recognize geometric ideas and relationships and apply them to problems • Use visualization, spatial reasoning, and geometric modeling to solve problems.

Grade 3 – Patchwork Quilt

Patchwork Quilt

Mean: 5.57, S.D.: 2.40



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	4.6%	6.6%	11.0%	19.8%	30.6%	45.1%	60.5%	76.5%	89.9%	98.3%	100.0%
% > =	100.0%	95.4%	93.4%	89.0%	80.2%	69.4%	54.9%	39.5%	69.4%	10.1%	1.7%

The maximum score available on this task is 10 points.
The cut score for a level 3 response is 5 points.

Most students (about 90%) could name the three shapes. More than half the students (about 70%) could name the shapes and draw lines of symmetry. A little less than half of the students could name two attributes of a rectangle. Only 2% of the students could meet all the demands of the task. 4% of the students scored no points on this task.

Based on teacher observations, this is what third grade students seemed to know and be able to do:

- Name geometric shapes
- Draw lines of symmetry
- Know that a rectangle has 4 sides or 4 angles

Areas of difficulty for third graders, third grade students struggled with:

- Understanding that squares are rectangles
- Precisely identifying distinguishing attributes of shapes
- Drawing designs with 2 lines of symmetry
- Knowing that the angles in the rectangle were right angles or 90 degrees

Implications for Instruction:

Students need more practice identifying specific characteristics of simple geometric shapes and experience with classifying and sorting. Students need to work with the logic of sorting to find mutually exclusive descriptors, understand the intersections of two groups, or the concept of subsets. Students also need to demonstrate an understanding of symmetry by showing lines of symmetry on a given design, and create a simple design with two lines of symmetry. Students need a variety of experiences looking at geometric properties, including sorting, paper folding, drawing designs, and making patterns with blocks. Students at this age need to work with a variety of symmetrical shapes, including shapes whose lines of symmetry are not vertical or horizontal.

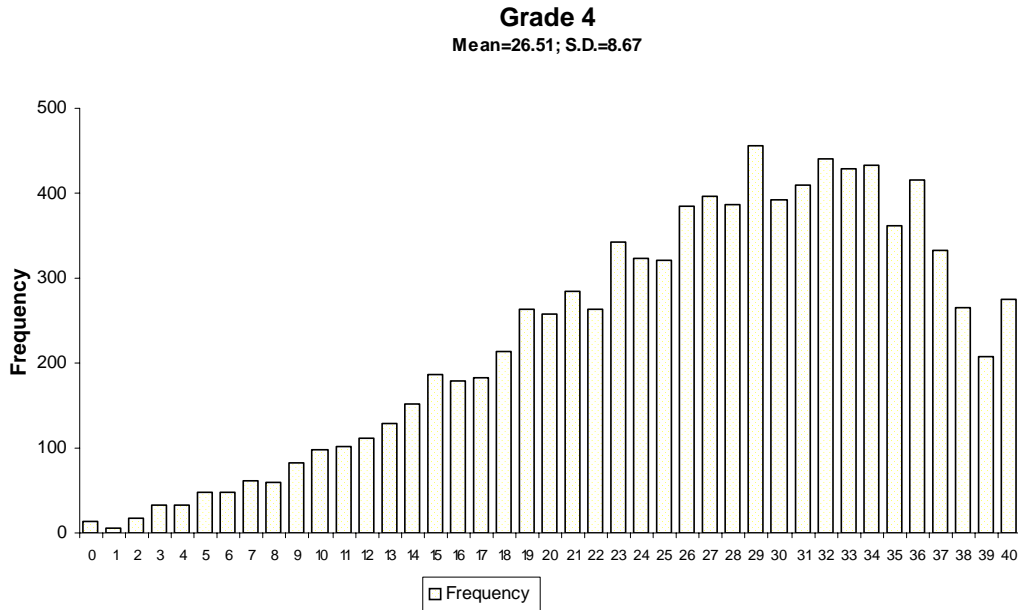
**Blue Print for Fourth Grade MAC
Alignment of Tasks to NCTM Content and Process Standards**

	Shapes with Straws	Number Trains	Hexagon Desks	Flower Arranging	Traveling to School
NUM	S	P		P	P
ALG			P	S	
GEO	P				
MEAS			S		S
DATA			S		P
PS		*	*	*	*
REAS	*	*	*	*	*
COMM	*		*	*	*
CONN	*		*		*
REP		*	*		

NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

P denotes Primary NCTM Content Standard
 S denotes Secondary NCTM Content Standard
 • denotes NCTM Process Standard

Overall Frequency Distribution by Total Score



Level Frequency Distribution Chart and Frequency Distribution

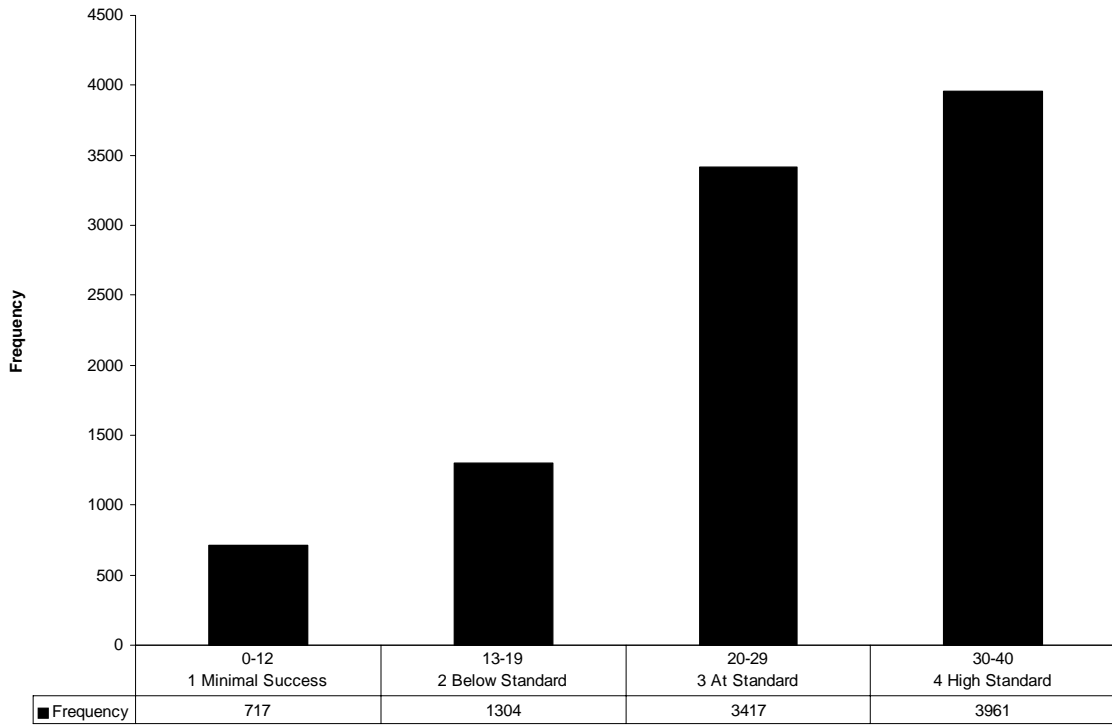
2003 - Numbers of students Grade 4:
tested: 9399

Grade 4 2000 - 2001

Level	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	30%	100%	18%	100%
2	37%	77%	23%	82%
3	26%	33%	39%	58%
4	7%	7%	19%	19%

Grade 4 2002 - 2003

Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	8%	100%	8%	100%
2	27%	92%	14%	92%
3	37%	66%	36%	78%
4	28%	28%	42%	42%

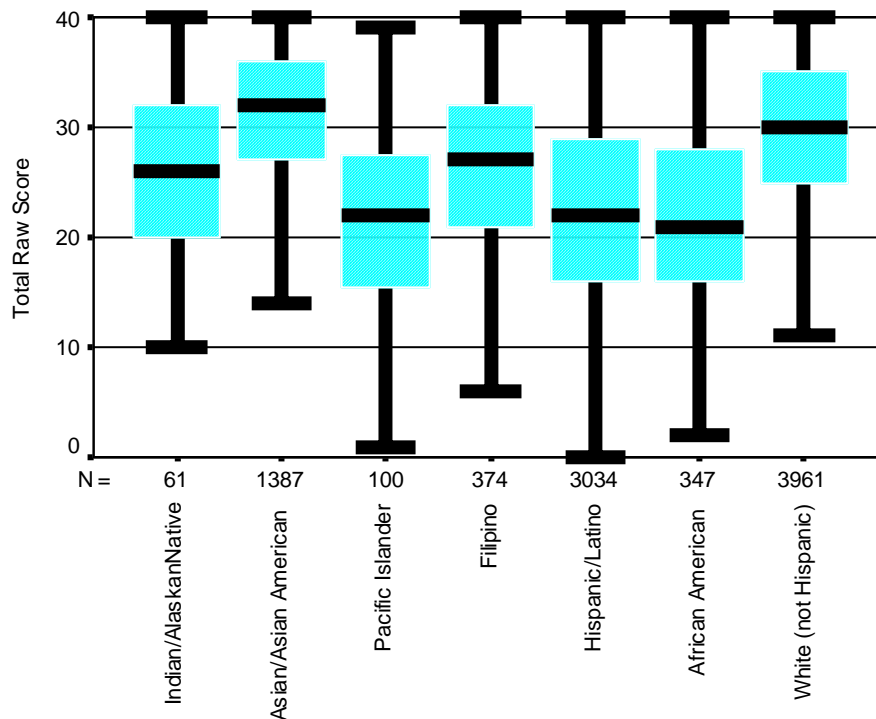


Grade 4

Distribution by Performance Levels

Grade 4 - Ethnicity

The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values* within a category lie between the highest and lowest horizontal bars.



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Distribution of sampling means -Grade 4 - Ethnicity

In this section, test scores are compared across different ethnic groups⁶. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if the scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than group B. When the two scores overlap, then there is not significant difference between the groups.

The scores of Indian/Alaskan Native students are significantly lower than those of Asian/Asian American and White students, and are significantly higher than those of Pacific Islander, Hispanic/Latino, and African American students. There are no

⁶ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

significant differences between the scores of Indian/Alaskan Native students and Filipino students.

The scores of Asian/Asian American students are significantly higher than the scores of students of all other ethnic groups.

The scores of Pacific Islander students are significantly lower than those of the Indian/Alaskan Native, Asian/Asian American, Filipino, and White students. There are no significant differences between the scores of Pacific Islander students and any other ethnic group.

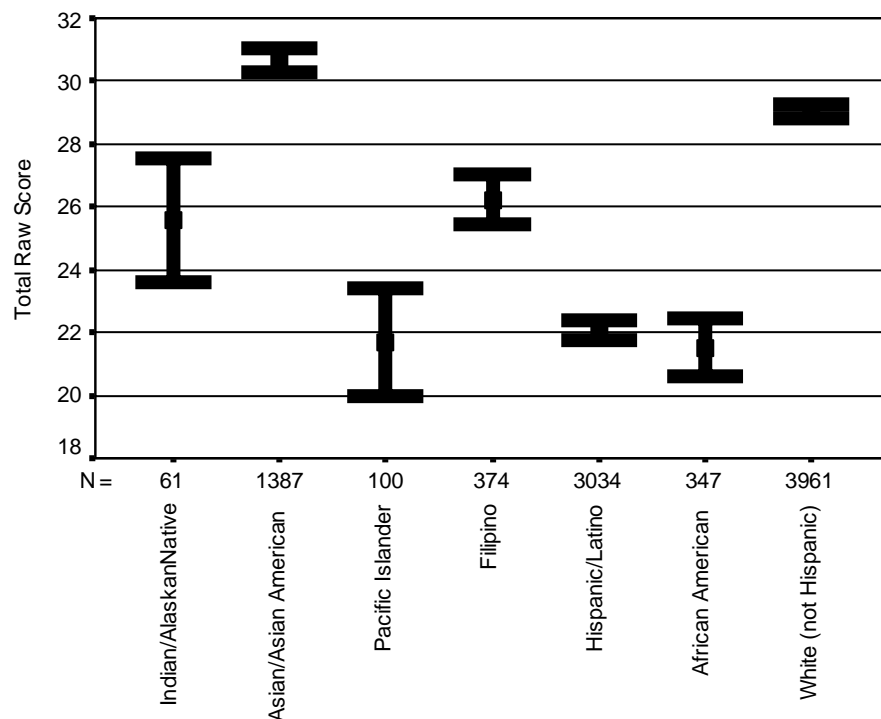
The scores of Filipino students are significantly lower than those of the Asian/Asian American and White students, and significantly higher than those of Pacific Islander, Hispanic/Latino and African American students. There are no significant differences between the scores of Filipino students and Indian/Alaskan Native students.

The scores of Hispanic/Latino students are significantly lower than those of the Indian/Alaskan Native, Asian/Asian American, Filipino, and White students. The scores of the Hispanic/Latino students are not significantly different from those of the Pacific Islanders and the African American students.

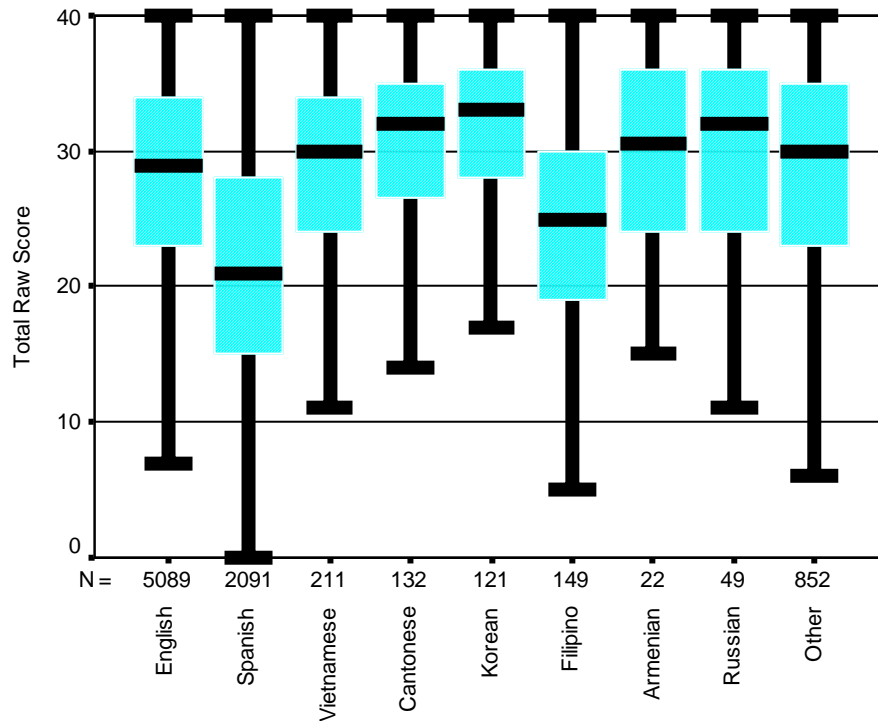
The scores of African American students are significantly lower than those of the Indian/Alaskan Native, Asian/Asian American, Filipino, and White students. The scores of African American students were not significantly different from those of the Hispanic/Latino students or the Pacific Islanders.

The scores of White students are significantly higher than those of Indian/Alaskan Native, Pacific Islander, Filipino, Hispanic/Latino, and African American students. The scores of White students were significantly lower than those of the Asian/Asian American students.

Distribution of sampling means - Grade 4 - Ethnicity



Grade 4 - Home Language



Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Lahu, Lao, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means -Grade 4 -Home Language

In this section, test scores are compared across groups of students who speak different languages at home⁷. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with Cantonese or Korean as a home language, and significantly higher than those of students with Spanish or Filipino as a home language. There are no significant differences between the scores of students with English as a home language and any other group.

⁷ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

The scores of students with Spanish as a home language are significantly lower than those of students in any other home language group.

The scores of students with Vietnamese as a home language are significantly higher than those of students with Spanish or Filipino as a home language. There are no significant differences between the scores of students with Vietnamese as a home language and any other group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, or Filipino as a home language. There are no significant differences between the scores of students with Cantonese as a home language and any other group.

The scores of students with Korean as a home language are significantly higher than those of students with English, Spanish, or Filipino as a home language. There are no significant differences between the scores of students with Korean as a home language and any other group.

The scores of students with Filipino as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a home language, and significantly higher than those of students with Spanish as a home language.

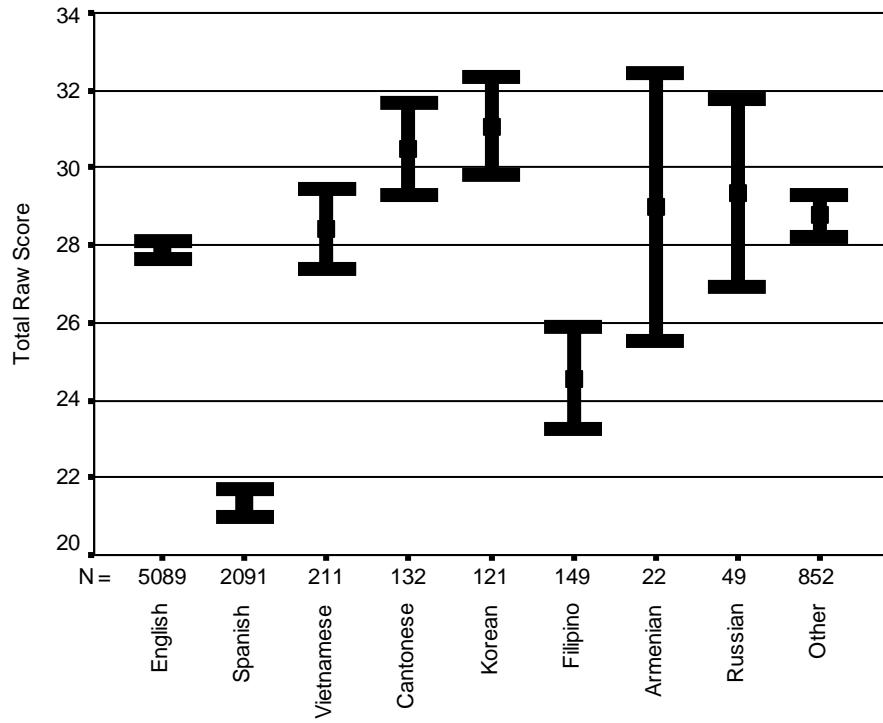
The scores of students with Armenian as a home language are significantly higher than the scores of students with Spanish as their home language, but there are no other significant differences with any other language group.

The scores of students with Russian as a home language are significantly higher than those of students with Spanish or Filipino as a home language. There are no significant differences between the scores of students with Russian as a home language and any other language group.

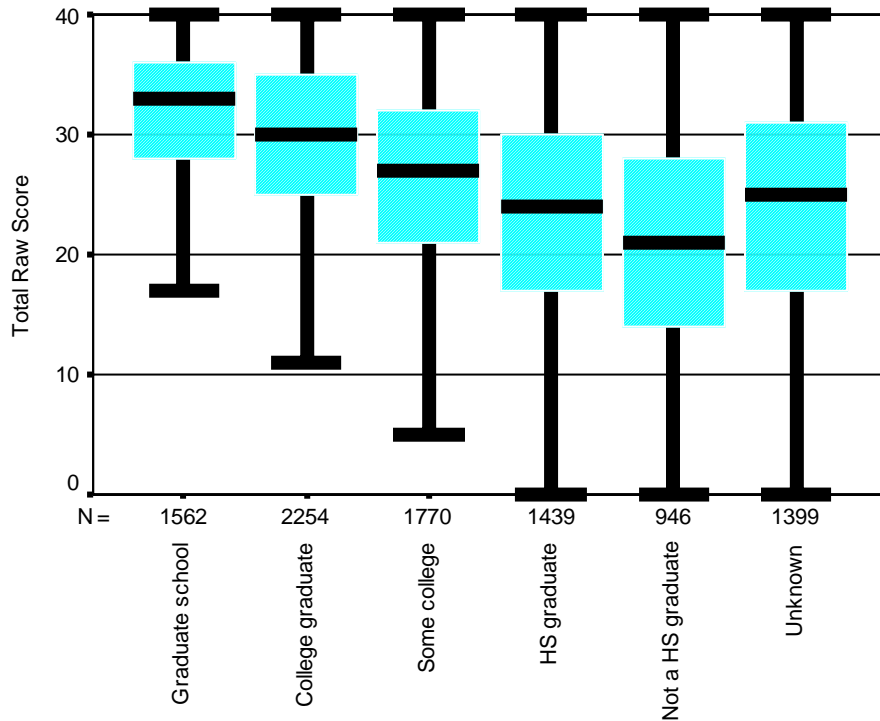
The scores of students with “Other” as a home language are significantly higher than those of students with Spanish or Filipino as a home language. There are no significant differences between the scores of students with “Other” as a home language and any other language group.

Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Lahu, Lao, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means- Grade 4 -Home Language



Grade 4 - Parent Education



Distribution of sampling means -Grade 4 -Parent Education

In this section, test scores are compared across groups of different levels of parent education⁸. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate school education are significantly higher than those of students in all other Parent Education categories.

The scores of students whose parents have a college education are significantly lower than those whose parents have a graduate education, and significantly higher than all other groups.

The scores of students whose parents have some college education are significantly lower than those whose parents have a college education or a graduate education. The scores of students whose parents have some college education are significantly higher than those of students whose parents are not High School graduates, those whose parents are High School graduates, and those whose parent education level is unknown.

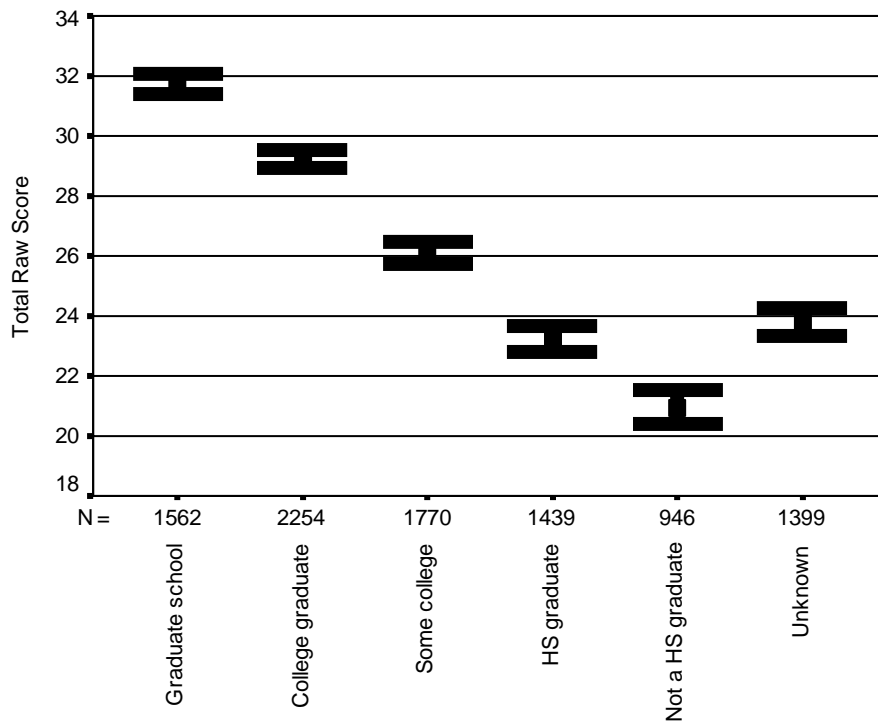
⁸ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

The scores of students whose parents are High School graduates are significantly higher than those of students whose parents are not High School graduates, and significantly lower than those whose parents have a college education or a graduate education. There is no statistically significant difference between the scores of students whose parents are High School graduates and those in the “Unknown” category.

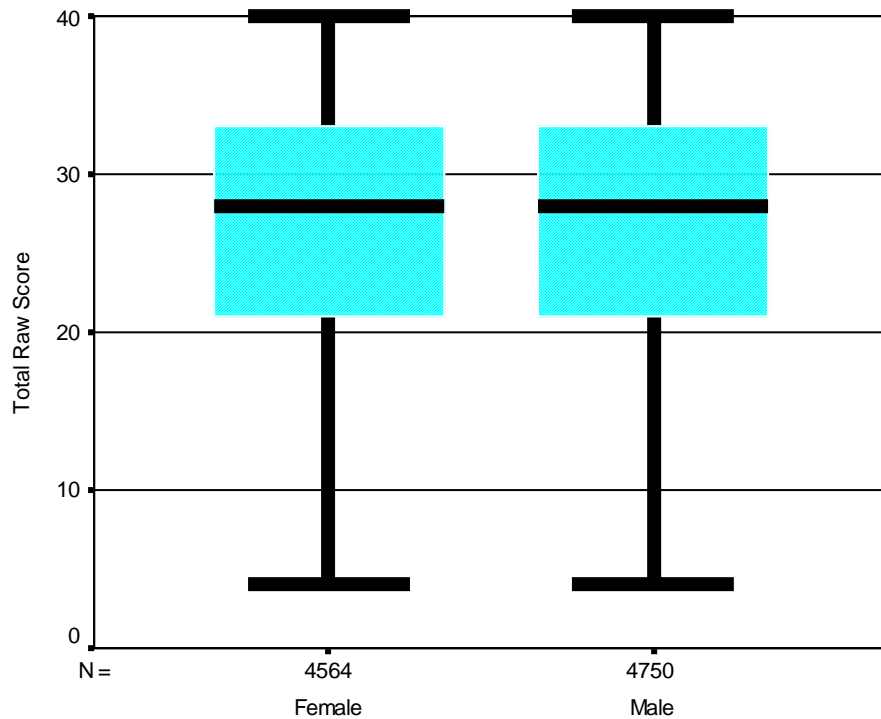
The scores of students whose parents are not High School graduates are significantly lower than the scores of students in all other Parent Education categories.

The scores of students whose parents’ education is unknown are significantly higher than those of students whose parents are not High School graduates, but lower than those whose parents had some college, graduated from college, or who had a graduate education. There was no statistically significant difference between students whose parents’ education is unknown and those whose parents are High School graduates.

Distribution of sampling means - Grade 4 - Parent Education



Grade 4 - Gender



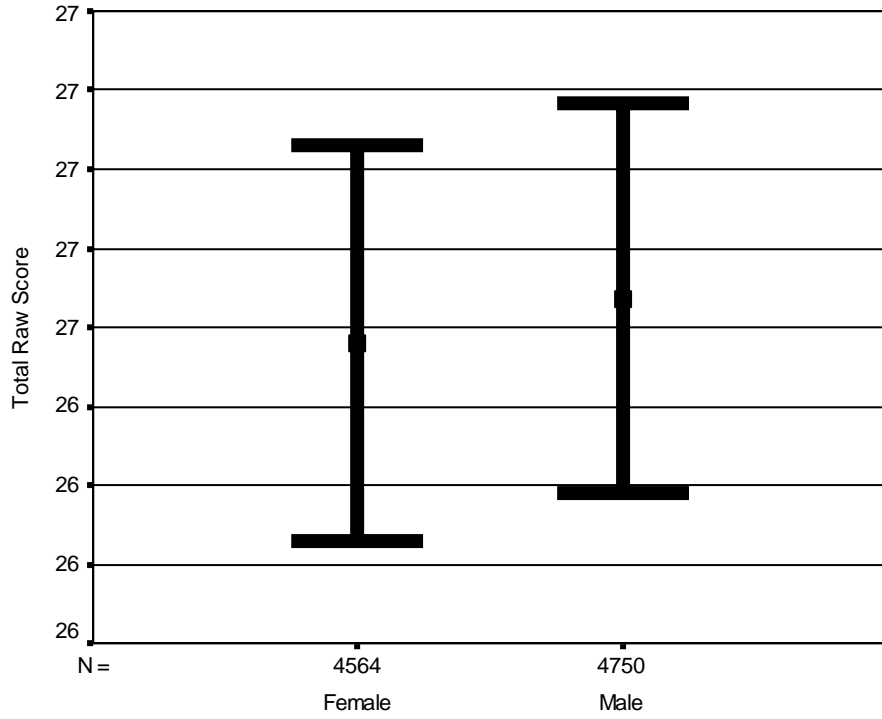
Distribution of sampling means - Grade 4 - Gender

In this section, test scores are compared across gender⁹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

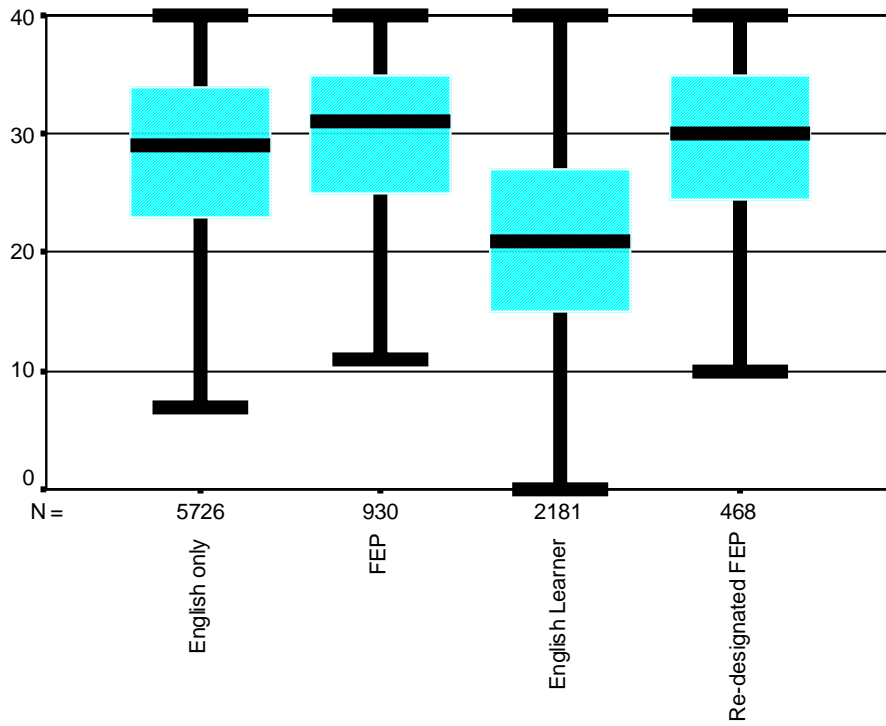
There is no statistically significant difference between the scores of females and males.

⁹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 4 – Gender



Grade 4 - Language Fluency



Distribution of sampling means - Grade 4 - Language Fluency

In this section, test scores are compared across groups of different language fluency¹⁰. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly lower than those of students described as Full English Proficiency (FEP) and those described as Re-designated FEP, and significantly higher than those students described as English Learners.

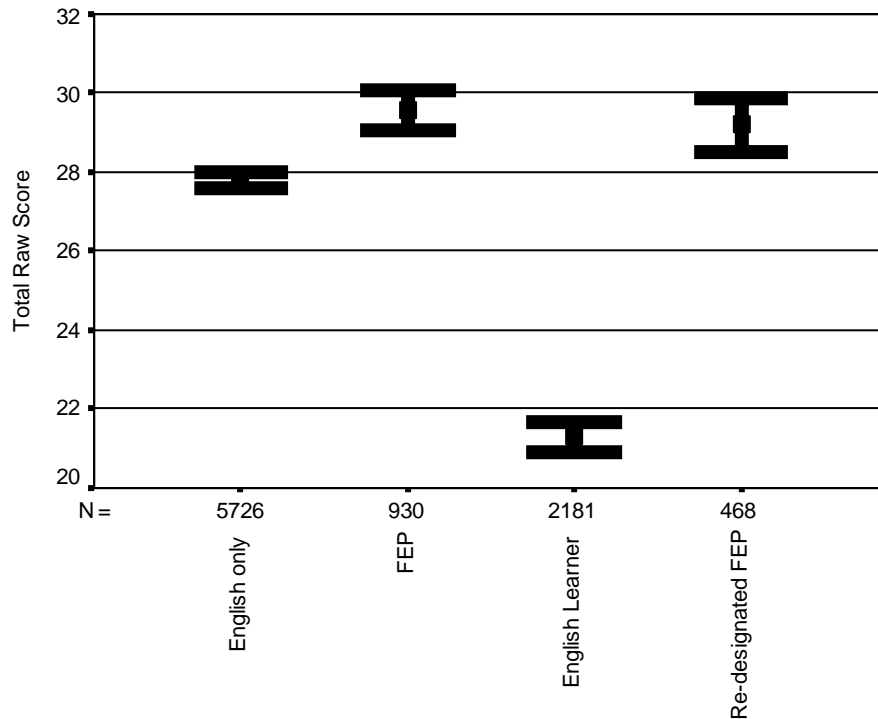
The scores of students in the FEP category are significantly higher than those of students in the English only and English Learner categories. The scores of students in the FEP category are not significantly different than the scores of students in the re-designated FEP group.

The scores of students in the English Learner category are significantly lower than those of the students in all other Language Fluency categories.

The scores of students in the Re-designated FEP category are significantly higher than those of students in all the English Only and English Learner categories, but not significantly different from those in the FEP category of Language Fluency.

¹⁰ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 4 - Language Fluency



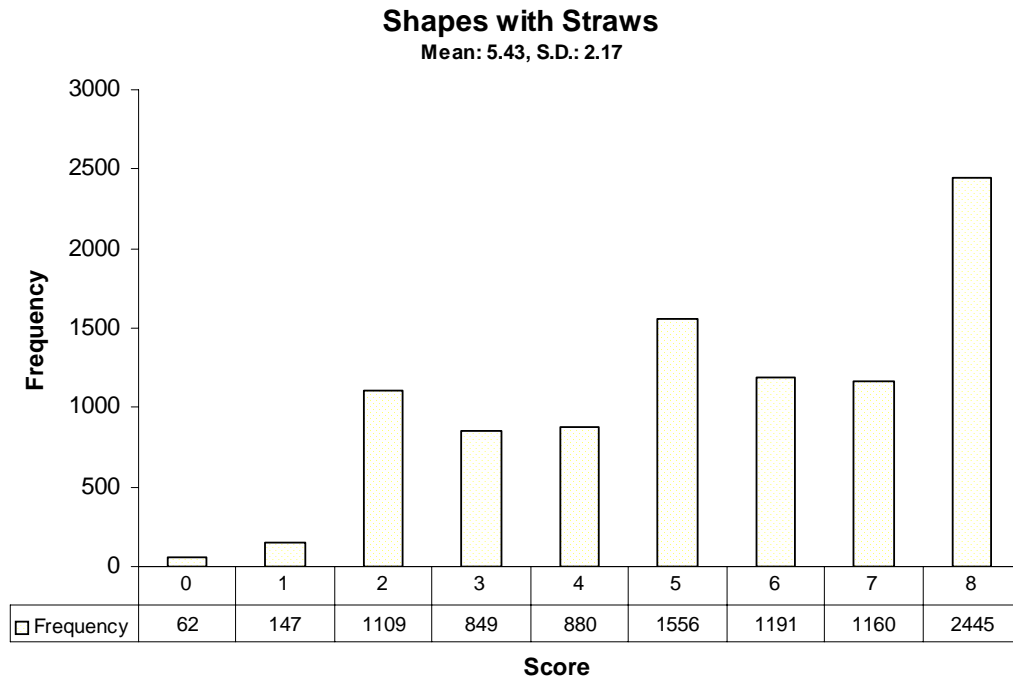
4th grade

Task 1

Shapes with Straws

Student Task	Solve problems about divisors and multiples in a geometric context.
Core Idea 2 Number Operations	<p>Understand the meanings of operations and how they relate to each other, make reasonable estimates, and compute fluently.</p> <ul style="list-style-type: none"> Develop fluency with basic number combinations for multiplication and division and use these combinations to mentally compute related problems
Core Idea 3 Geometry and Measurement	<p>Use characteristics, properties, and relationships of two-dimensional geometric shapes and apply appropriate techniques to determine measurements.</p> <ul style="list-style-type: none"> Recognize geometric ideas and relationships and apply them to problems. (3rd grade) Identify and compare attributes of two-dimensional shape and develop vocabulary to describe the attributes.

Frequency Distribution for each Task – Grade 4
Grade 4 – Shapes with Straws



Score:	0	1	2	3	4	5	6	7	8
% < =	0.7%	2.2%	14.0%	23.1%	32.4%	49.0%	61.6%	74.0%	100.0%
% > =	100.0%	99.3%	97.8%	86.0%	76.9%	67.6%	51.0%	38.4%	26.0%

The maximum score available on this task is 8 points.
The cut score for a level 3 response is 4 points.

Most students (more than 98%) could identify the number of straws to make each of the geometric shapes. About 77% of the students met standards. They could generally name the shapes, find the number of straws needed to make squares and triangles, and do some of thinking for finding straws for 9 triangles. Many students (about 67%) could identify the number of straws to make the shapes, find the number of straws to make 9 triangles and 5 hexagons (but not necessarily the total), and answer some part of question 2. A little less than half the students (about 40%) could find the straws needed for each shape, find the number of shapes that could be made with 24 straws and find the number of straws to make 9 triangles and 5 hexagons (but not necessarily the total). More than 25% could complete all the demands of the task. Less than 1% of the students scored no points on this task.

Based on teacher observations, this is what fourth grade students seemed to know and be able to do:

- Find the number of straws needed to make squares and triangles.
- Use counting, drawing, repeated addition, multiplication or division to find the number of shapes that could be made with 24 straws.
- Use multiplication to find the number of straws to make 9 triangles.

Areas of difficulty for fourth graders, fourth grade students struggled with:

- Finding the number of straws needed to make hexagons and pentagons.
- Finding the number of hexagons or pentagons that could be made with 24 straws.
- Finding the total number of straws to make 9 triangles and 5 hexagons. They could generally find the amount of each, but missed the logic of combining the answers to find the total.

Implications for Instruction:

Students need to interpret situations that require multiplication and/or division to solve. Students also need experience with demonstrating mathematical reasoning by explaining how they determined their answers. Students at this age should have many experiences solving word problems using multiplication and division. Students need more experiences with word problems in which the numbers are embedded in a context where the needed information is not laid out for them, but rather where they have to figure out what information they will need and how to use it.

4th grade

Task 2

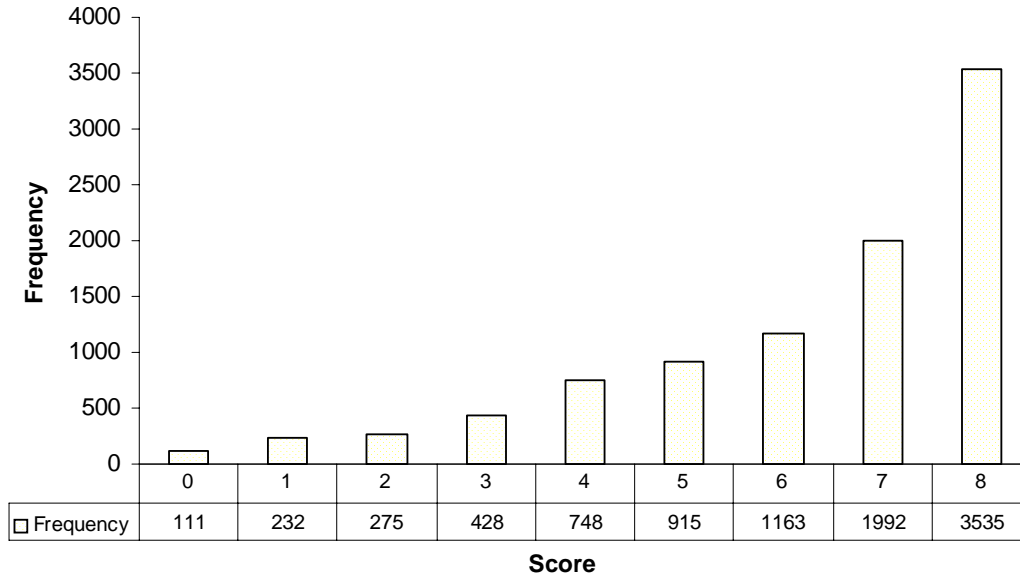
Number Trains

Student Task	Solve problems about factors and multiples in a toy train context.
Core Idea 3 Patterns, Functions, and Algebra	<p>Understand patterns and use mathematical models to represent and to understand qualitative and quantitative relationships.</p> <ul style="list-style-type: none"> • Use inverse operations to solve multi-step problems • Understand and use the concept of equality
Core Idea 2 Number Operations	<p>Understand the meanings of operations and how they relate to each other, make reasonable estimates, and compute fluently.</p> <ul style="list-style-type: none"> • Develop fluency with basic number combinations for multiplication and division and use these combinations to mentally compute related problems

Grade 4 – Number Trains

Number Trains

Mean: 6.26, S.D.: 2.00



Score:	0	1	2	3	4	5	6	7	8
% < =	1.2%	3.6%	6.6%	11.1%	19.1%	28.8%	41.2%	62.4%	100.0%
% > =	100.0%	98.8%	96.4%	93.4%	88.9%	80.9%	71.2%	58.8%	37.6%

The maximum score available on this task is 8 points.

The cut score for a level 3 response is 4 points.

Most students (about 90%) could find the factors for 18 and 24 as well as multiply 2×18 to get 36. Many students (about 71%) could find all the factor pairs for 12, 18, and 24. They had difficulty finding all the factor pairs for 36. Almost 40% of the students could meet all the demands of the task. Only 1% of the students scored no points on this task.

Based on teacher observations, this is what fourth grade students seem to know and be able to do:

- Find factor pairs to equal various products.
- Multiply 2×18 .

Areas of difficulty for four graders, fourth grade students struggled with:

- Making an exhaustive list of all the factor pairs for a given product.
- Making an organized list.
- Remembering that $1 \times$ itself is a factor pair for any number, except 0.

Implications for Instruction:

Students need practice with interpreting problems and their constraints. At this grade level, students need to be fluent with, and recognize, multiple factors of composite numbers. Some students had difficulty recognizing when they had already used a given solution. Students need to be able to develop the logical reasoning to check for differences and compare answers. Students need more experience with the problem solving strategy: make an organized list. This strategy will help them to notice or eliminate duplicates. The Lane County Problem Solving in Mathematics by Dale Seymour Publications is a good resource for teaching problem solving strategies.

4th grade

Task 3

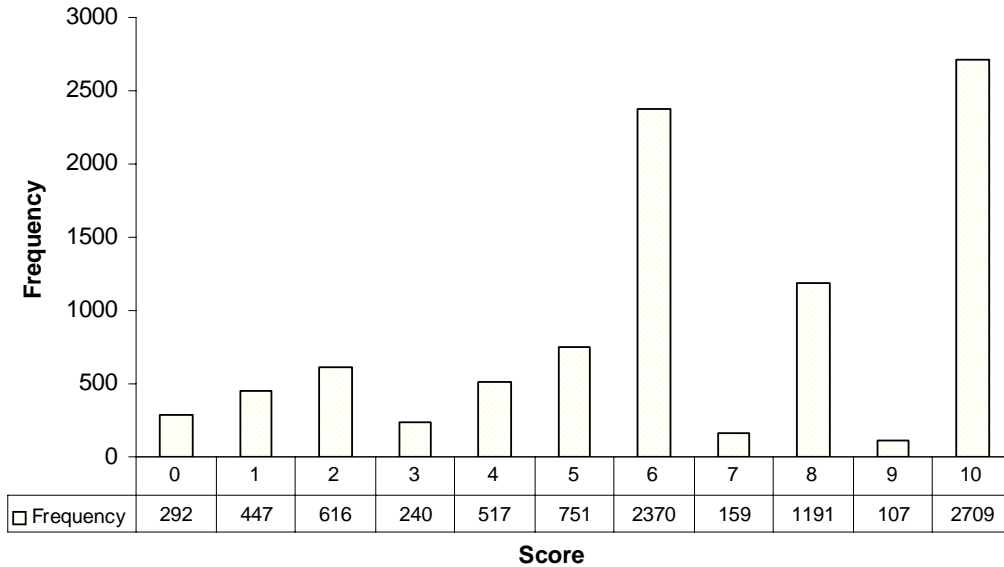
Hexagon Desks

Student Task	Find, extend, and analyze a number pattern involving chairs placed around hexagonal desks. Plot and use a graph to display the number pattern information.
Core Idea 3 Patterns, Functions, and Algebra	Understand patterns and use mathematical models to represent and to understand qualitative and quantitative relationships. <ul style="list-style-type: none">• Represent and analyze patterns and functions using words, tables, and graphs• Find the results of a rule for a specific value• Use concrete, pictorial, and verbal representations to solve problems involving unknowns.

Grade 4 – Hexagon Desks

Hexagon Desks

Mean: 6.50, S.D.: 2.98



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	3.1%	7.9%	14.4%	17.0%	22.5%	30.5%	55.7%	57.4%	70.0%	71.2%	100.0%
% > =	100.0%	96.9%	92.1%	85.6%	83.0%	77.5%	69.5%	44.3%	42.6%	30.0%	28.8%

The maximum score available on this task is 10 points.

The cut score for a level 3 response is 5 points.

Most students (about 85%) could find and continue the pattern correctly for at least 2 points and earn partial credit for graphing. Many students (about 70%) could find all the numbers in the table and graph them correctly. 30% of the students could meet all the demands of the task including making a mathematical justification for why 47 students could not fit at 11 desks. About 3% of the students scored no points on this task.

Based on teacher observations, this is what fourth grade students seemed to know and be able to do:

- Continue a pattern in a table
- Graph points on a grid
- Understand the relationship between repeated addition and multiplication

Areas of difficulty for fourth graders, fourth grade students struggled with:

- Relating the numbers in the tables to the points on the grid (many students seemed to be relying on a rule about putting points in a straight line and sometimes thinking the straight line needed to go through zero)
- Connecting the table and the graph as having the same information and being related rather than two separate tasks
- Making sense of the graph, only 14% used the graph to help them make a mathematical argument in part 3
- Making a mathematical argument to explain why 47 students can't fit at the desks

Instructional Implications:

Students at this grade level need to experience analyzing growing patterns. They also need to be able to correctly plot the information from a table onto a graph with pre-established scales. While many students seemed to know that the plots for growing patterns should be on a straight line, the ordered numbers on their graphs did not match the numbers in their tables. Students need more experiences with linear relationships that are not proportional: i.e., the numbers are not multiples. For example, in this problem the number of students increases by four with each additional desk, but the total number of students is not a multiple of four, so this is not a proportional relationship. Students also need to be able to work backward from a solution. Students might do this by extending the table, extending their graph, or by using repeated addition. An important part of patterning is to recognize the relationship between the two variables. Learning to think and describe what is changing and what is staying the same helps students to find rules that will work for all cases.

4th grade

Task 4

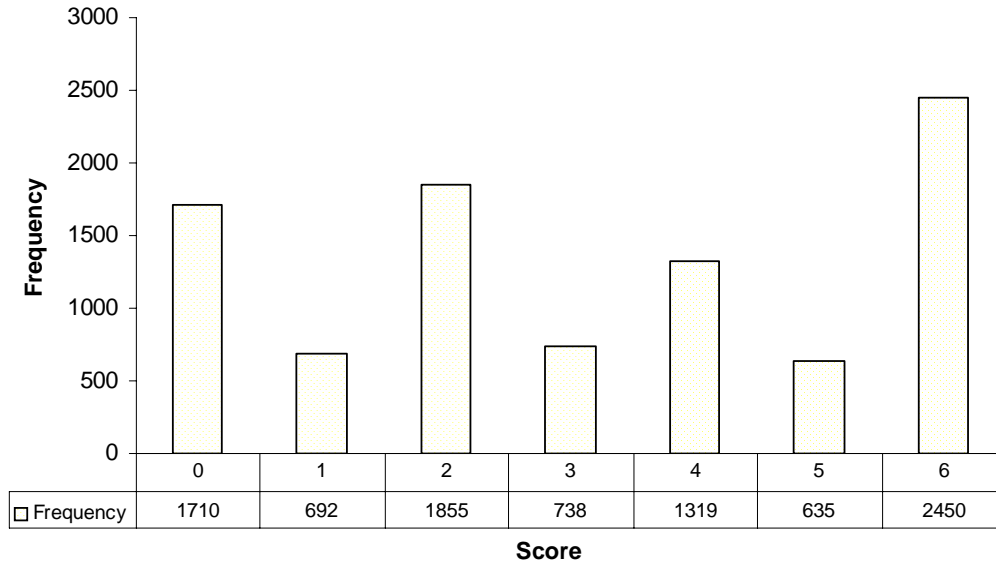
Flower Arranging

Student Task	Divide a number into parts in order to satisfy given conditions with regard to arranging flowers.
Core Idea 1 Number Properties	<p>Understand numbers, ways of representing numbers, relationships among numbers, and number systems.</p> <ul style="list-style-type: none"> • Develop a sense of whole numbers and represent and use them in flexible ways, including relating, composing, and decomposing numbers (3rd grade) • Develop understanding of the relative magnitude of whole numbers and the concepts of sequence, quantity, and the relative positions of numbers.

Grade 4 – Flower Arranging

Flower Arranging

Mean: 3.17, S.D.: 2.20



Score:	0	1	2	3	4	5	6
% < =	18.2%	25.6%	45.3%	53.1%	67.2%	73.9%	100.0%
% > =	100.0%	81.8%	74.4%	54.7%	46.9%	32.8%	26.1%

The maximum score available on this task is 6 points.

The cut score for a level 3 response is 3 points.

Most students (about 82%) could find three numbers to equal the correct total. A little less than half the students (about 47%) could find three numbers for each part that matched all the constraints in the problem. About 28% of the students could also explain how the solution fit the constraints. About 18% of the students scored no points on this task. Most of these students attempted the problem.

Based on teacher observations, this is what fourth grade students seemed to know and be able to do:

- Find 3 numbers to add to a given total
- Arrange numbers from large to small by correctly interpreting the meaning of the word “more”

Areas of difficulty for fourth graders, fourth grade students struggled with:

- Using only odd numbers to make a total
- Using 3 constraints in one problem
- Explaining how they used the constraints to check their solution

Instructional Implications:

Students need experience solving problems and satisfying multiple constraints. Students needed to find three different odd numbers to add to a particular sum. The students were also given information regarding the order of the numbers, which type of flower was the greatest, etc. Solving number puzzles with different clues might be one way to help students develop their logical reasoning, as well as computational fluency. Students also need to be able to explain mathematical reasoning and thinking in a way that is clear to others. Learning techniques like making an organized list and eliminating possibilities seemed to be helpful for some students. Students need rich tasks with lots of information and opportunities to struggle and organize that information in ways that make sense to them. Giving students opportunities to compare strategies with others helps them clarify why a strategy is useful and what makes a strategy useful.

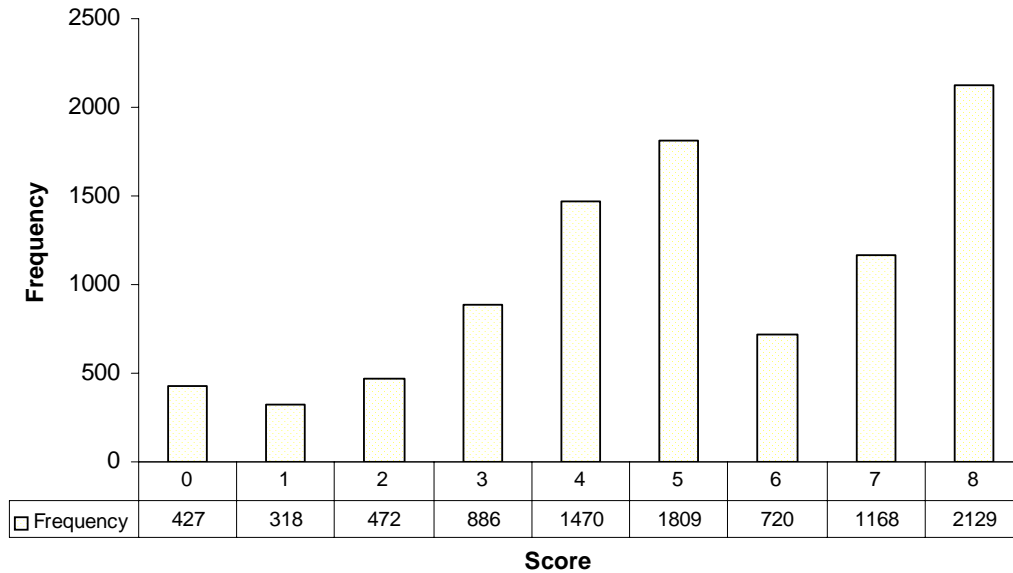
4th grade Task 5 Traveling To School

Student Task	Interpret a table of travel times. Use the information from this table to solve problems.
Core Idea 5 Data Analysis	Collect, organize, represent and interpret numerical and categorical data, and clearly communicate their findings. <ul style="list-style-type: none"> • Interpret data to answer questions about a situation
Core Idea 2 Number Operations	Reason about and solve problem situations that involve more than one operation in multi-step problems.
Core Idea 1 Number Properties	Understand numbers, ways of representing numbers, relationships among numbers, and number systems. <ul style="list-style-type: none"> • Use models, benchmarks, and equivalent forms to judge the size of friendly fractions

Grade 4 – Traveling to School

Traveling to School

Mean: 5.15, S.D.: 2.29



Score:	0	1	2	3	4	5	6	7	8
% < =	4.5%	7.9%	12.9%	22.4%	38.0%	57.3%	64.9%	77.3%	100.0%
% > =	100.0%	95.5%	92.1%	87.1%	77.6%	62.0%	42.7%	35.1%	62.0%

The maximum score available on this task is 8 points.

The cut score for a level 3 response is 4 points.

Most students (about 87%) could answer questions about a table and do comparison subtraction. Many students (about 77%) could answer questions about the graph, do comparison subtraction, and estimate the time to run to school. More than half the students (62%) could answer questions about the table, find the time it took to travel partway to school, return home, and then make it to school, and make an estimate and explain the reasoning behind it for the time it would take to run to school. About 22% of the students could meet all the demands of the task. About 4.5% of the students scored no points on this task.

Based on teacher observations, this is what fourth grade students seemed to know and be able to do:

- Answer questions from a table
- Comparison subtraction
- Estimate a reasonable rate

Areas of difficulty for fourth graders, fourth grade students struggled with:

- Giving a justification for an estimate by relating it to other information provided in the problem
- Making diagrams to track information in a word problem
- Finding the time for a trip with multiple stops

Instructional Implications:

Students need practice reading and interpreting information from a variety of tables with different formats. They also need to be able to choose from a variety of problem-solving strategies to correctly solve a word problem. Students need experience with problems that require them to draw a picture or diagram. Students need more experience with analyzing and solving multi-step word problems with units of time, elapsed time, etc. Students also need frequent opportunities to make estimates and justify their estimates. Both the teacher and the students need to articulate what is valued in a good explanation. In this particular case, the estimate should have compared or showed a relationship between running and some other rate in the table. Good mathematicians find connections between different pieces of the information.

Blue Print for Fifth Grade MAC

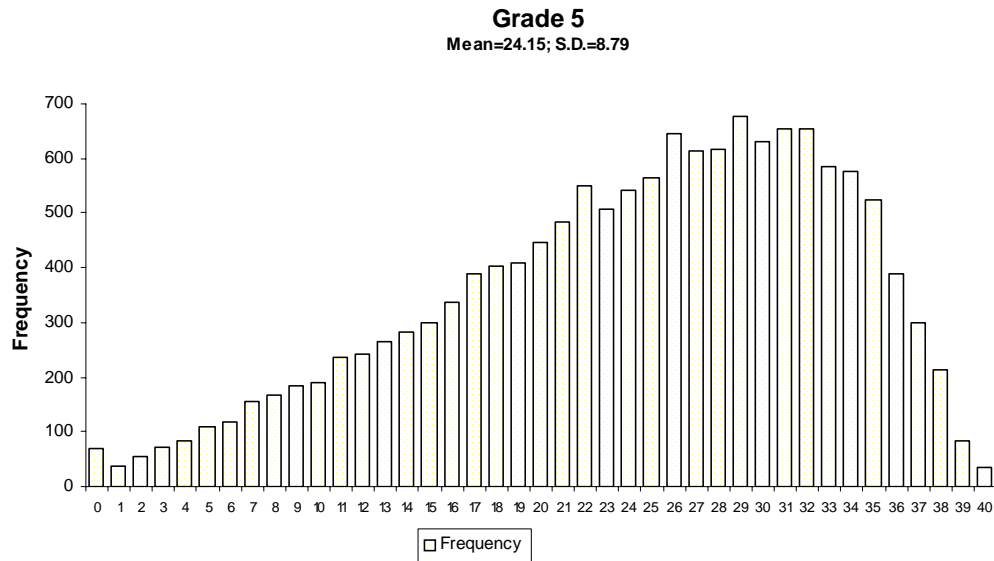
Alignment of Tasks to NCTM Content and Process Standards

	Number Story	Raspberry Cake	Buttons	Winter Sports	Juan's Shapes
NUM	P	S		S	
ALG	S		P	S	
GEO					S
MEAS		P			P
DATA				P	
PS	*	*		*	*
REAS	*	*	*	*	*
COMM	*		*	*	*
CONN		*	*		
REP		*	*		*

NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

P denotes Primary NCTM Content Standard
 S denotes Secondary NCTM Content Standard
 • denotes NCTM Process Standard

Overall Frequency Distribution by Total Score



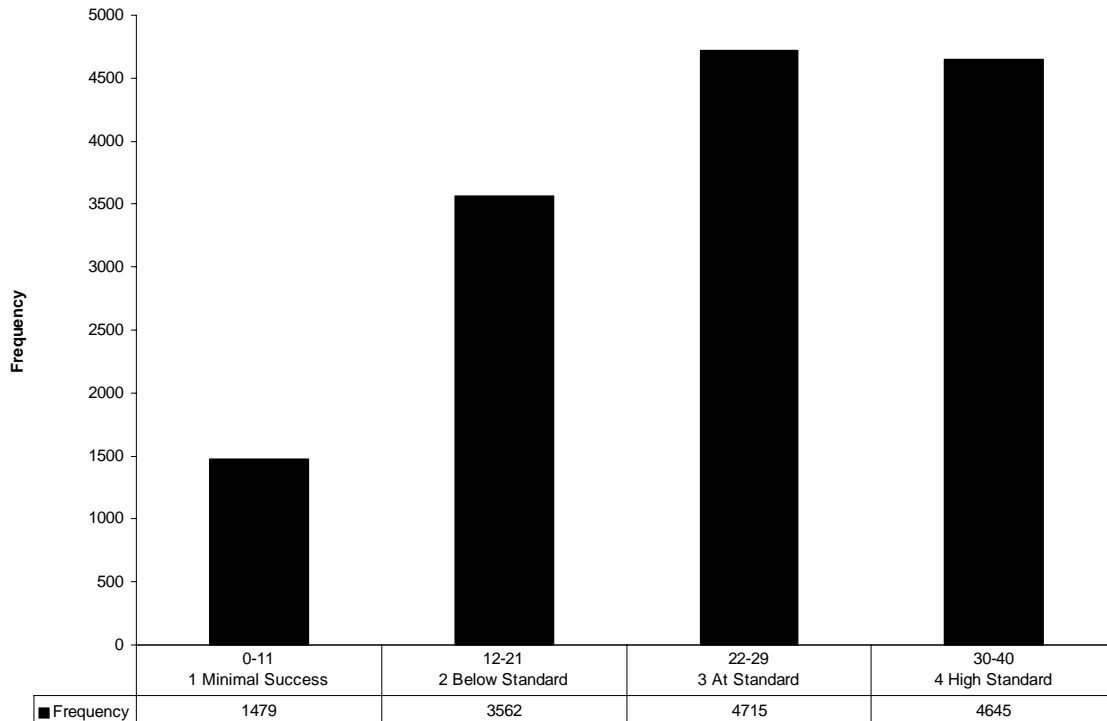
Level Frequency Distribution Chart and Frequency Distribution

2003 - Numbers of students Grade 5: 14401
tested:

Level	% at ('99)	% at least ('99)	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	25%	100%	12%	100%	11%	100%
2	43%	75%	36%	88%	30%	89%
3	22%	32%	34%	52%	36%	59%
4	10%	10%	18%	18%	23%	23%

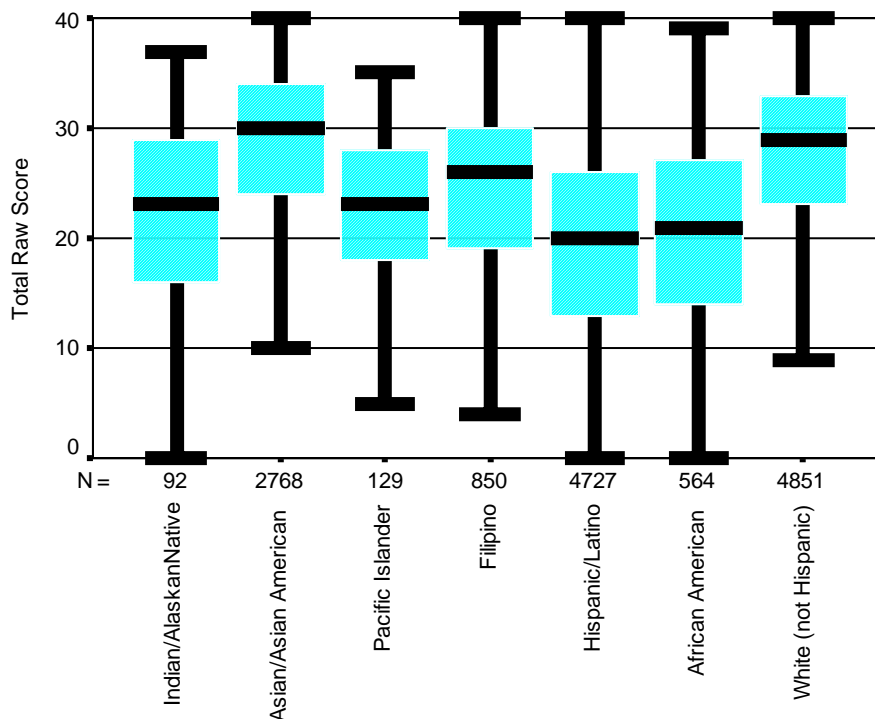
Grade 5 2002 - 2003

Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	22%	100%	10%	100%
2	31%	78%	25%	90%
3	25%	47%	33%	65%
4	22%	22%	32%	32%



Grade 5 – Ethnicity –

The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values* within a category lie between the highest and lowest horizontal bars.



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Distribution of sampling means - Grade 5 - Ethnicity

In this section, test scores are compared across different ethnic groups¹¹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if the scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than group B. When the two scores overlap, then there is not significant difference between the groups.

The scores of Indian/Alaskan Native students are significantly lower than those of Asian/Asian American and White students, and are significantly higher than those of Hispanic/Latino students. There are no significant differences between the scores of Indian/Alaskan Native students and any other ethnic group.

The scores of Asian/Asian American students are significantly higher than those of students of all other ethnic groups.

The scores of Pacific Islander students are significantly lower than those of Asian/Asian American and White students. The scores of Pacific Islander students are significantly higher than those of Hispanic/Latino students. There are no significant differences between the scores of Pacific Islander students and any other ethnic group.

The scores of Filipino students are significantly lower than those of Asian/Asian American and White students, and significantly higher than those of Hispanic/Latino and African American students. There are no significant differences between the scores of Filipino students and Indian/Alaskan Native students, or between Filipino students and Pacific Islanders.

The scores of Hispanic/Latino students are significantly lower than those of all of the other ethnic groups.

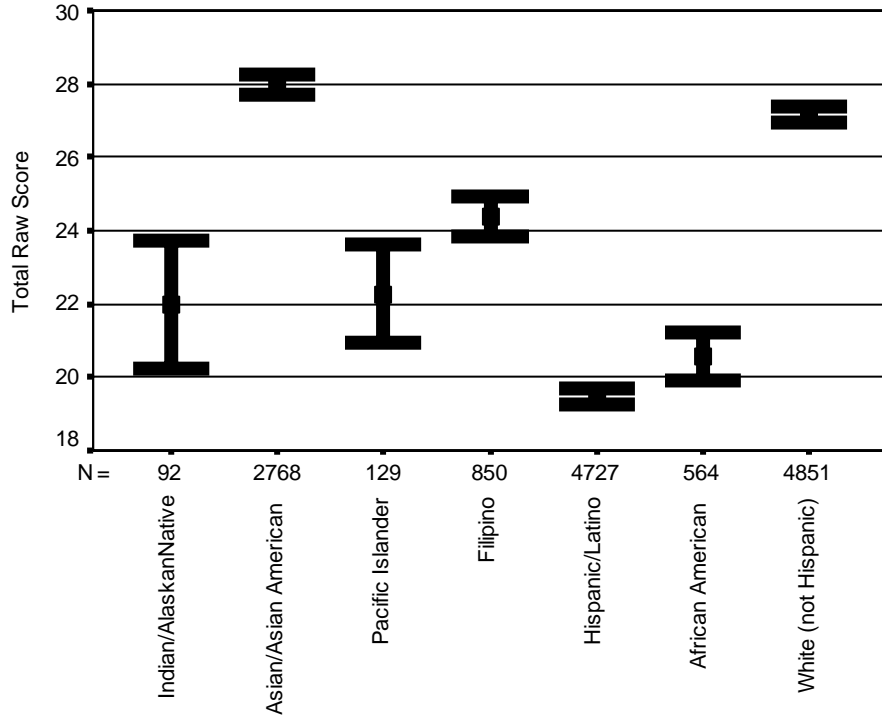
The scores of African American students are significantly lower than those of Asian/Asian American, Filipino, and White students. The scores of African American students were significantly higher than those of the Hispanic/Latino students. There are no significant differences between the scores of African American students and any other ethnic group.

The scores of White students are significantly higher than those of Indian/Alaskan Native, Pacific Islander, Filipino, Hispanic/Latino, and African American students. The

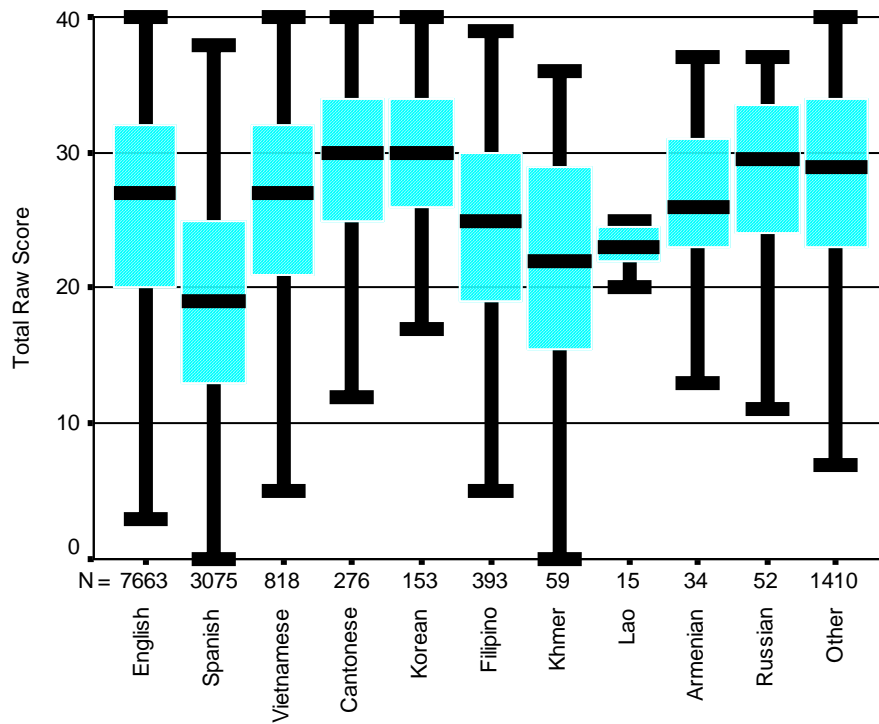
¹¹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

scores of White students were significantly lower than those of the Asian/Asian American students.

Distribution of sampling means - Grade 5 -Ethnicity



Grade 5 - Home Language



Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 5 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home¹². One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with a home language of Cantonese, Korean, or “Other”. The scores of students with English as a home language are significantly higher than those of students with a home language of Spanish, Filipino, or Khmer. There is no significant difference between scores of students with a home language of English and those with a home language of Lao, Armenian, or Russian.

The scores of students with Spanish as a home language are significantly lower than those of students with any other home language except Khmer and Lao. There is no significant difference between scores for students with Spanish as a home language, and students with Khmer or Lao as a home language.

The scores of students with Vietnamese as a home language are significantly higher than those of students with Spanish or Khmer as a home language, and lower than those of students with Cantonese, Korean, or “Other” as a home language. There are no significant differences between the scores of students with Vietnamese as a home language and any other group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or Khmer as a home language. There are no significant differences between the scores of students with Cantonese as a home language and any other group.

The scores of students with Korean as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, Khmer or “Other” as a home language. There are no significant differences between the scores of students with Korean as a home language and any other group.

¹² Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

The scores of students with Filipino as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a home language, and significantly higher than those of students with Spanish as a home language. There are no significant differences between the scores of students with Filipino as a home language and any other group.

The scores of students with Khmer as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a home language. The scores of students with Khmer as a home language are significantly higher than those of students with Spanish as their home language. There are no other significant differences between the scores of students with Khmer as a home language and any other group.

The scores of students with Lao as a home language are not significantly different from the scores of students in any other home language category.

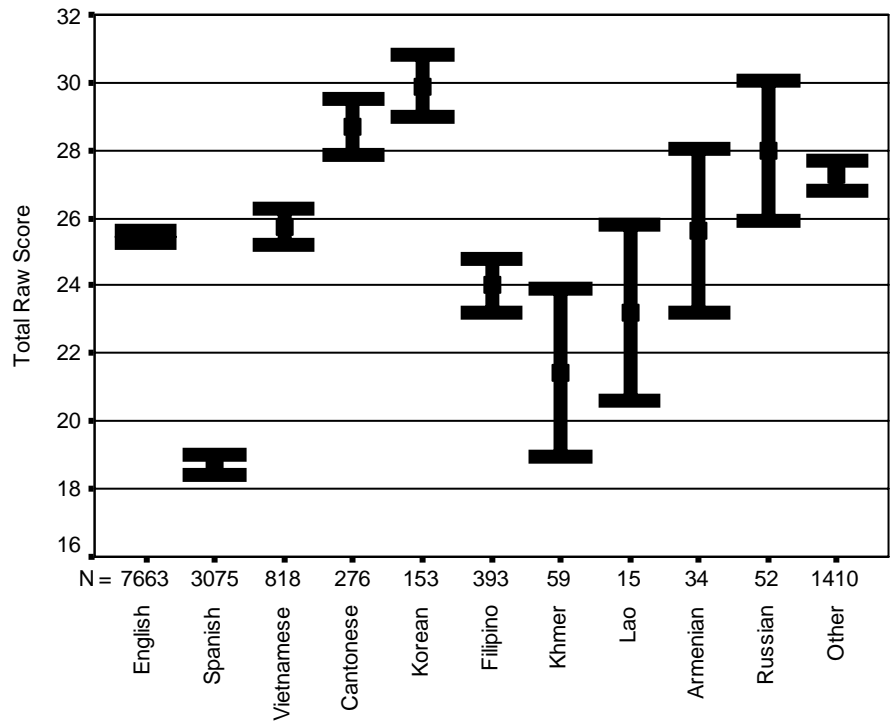
The scores of students with Armenian as a home language are significantly higher than the scores of students with Spanish as their home language, but there are no other significant differences with any other language group.

The scores of students with Russian as a home language are significantly higher than those of students with Spanish, Filipino, or Khmer as a home language. There are no other significant differences between students with Russian as a home language and any other language group.

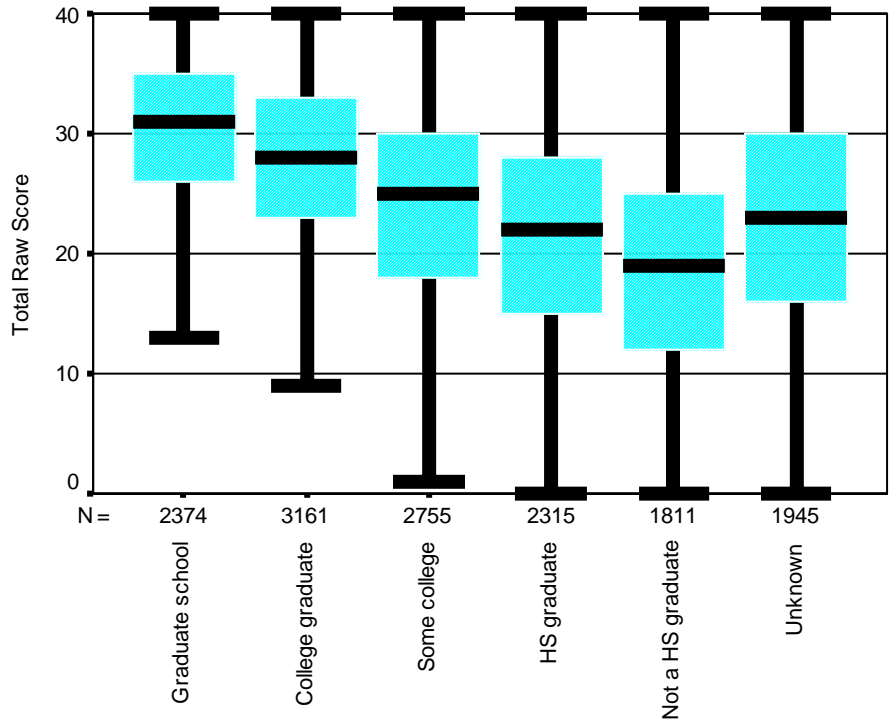
The scores of students with “Other” as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or Khmer as a home language, and significantly lower than students with Korean as a home language. There are no other significant differences between the scores of students with “Other” as a home language and any other language group.

Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means -Grade 5 -Home Language



Grade 5 -Parent Education



Distribution of sampling means - Grade 5 - Parent Education

In this section, test scores are compared across groups of different levels of parent education¹³. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate school education are significantly higher than those of students in all other Parent Education categories.

The scores of students whose parents have a college education are significantly lower than those whose parents have a graduate education, and significantly higher than all other groups.

The scores of students whose parents have some college education are significantly lower than those whose parents completed college or a graduate education, but higher than those of students in all other groups.

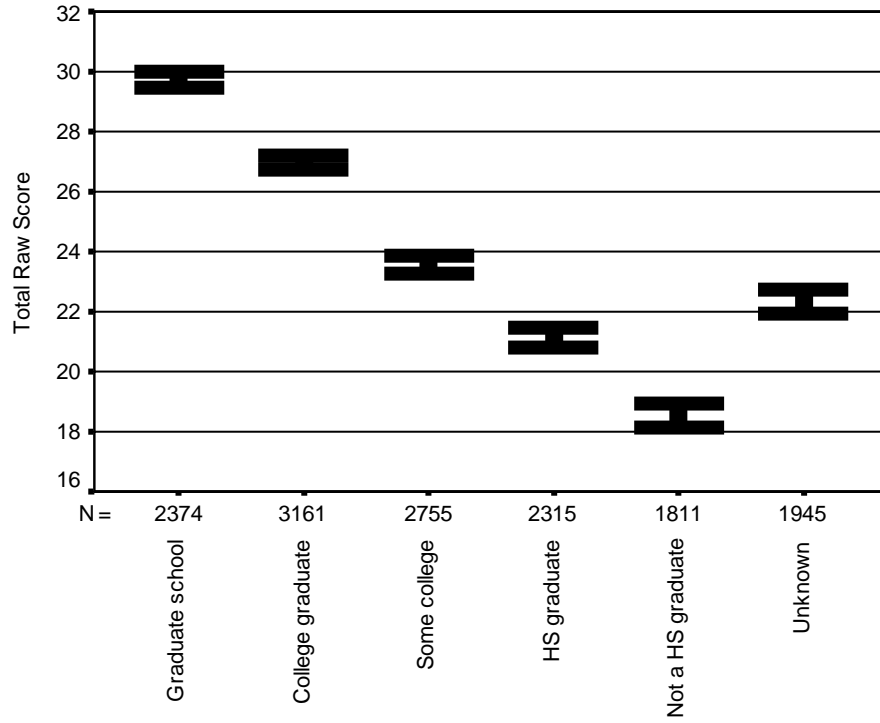
The scores of students whose parents are High School graduates are significantly higher than those of students whose parents are not High School graduates, and significantly lower than those whose parents have a college education, a graduate education, or those with parents in the “Unknown” category.

The scores of students whose parents are not High School graduates are significantly lower than the scores of students in all other Parent Education categories.

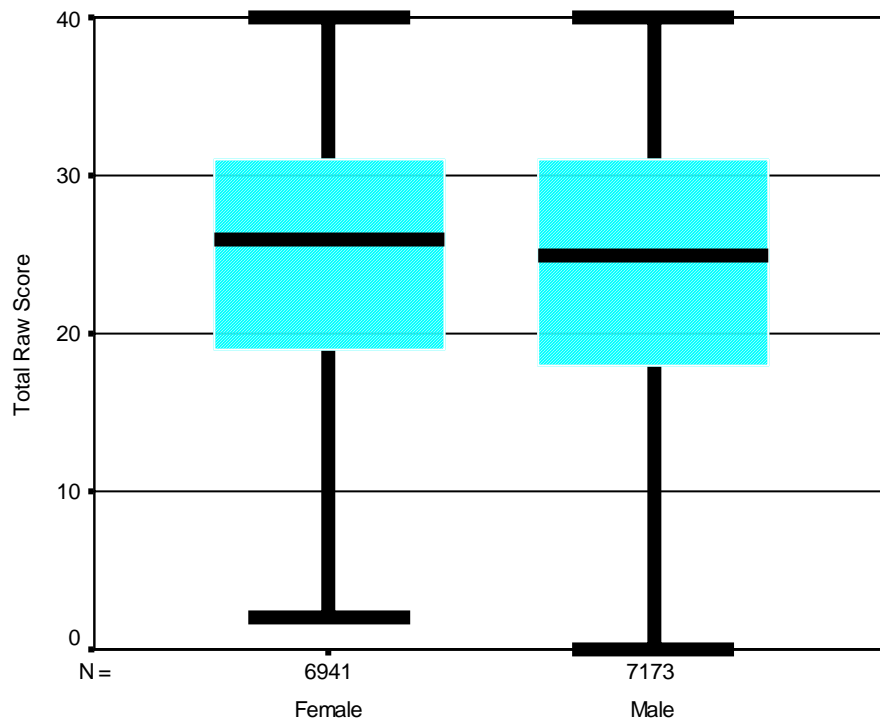
The scores of students whose parents' education is unknown are significantly higher than those of students whose parents are or are not High School graduates, but lower than those whose parents had some college, graduated from college, or who had a graduate education.

¹³ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 5 - Parent Education



Grade 5 - Gender

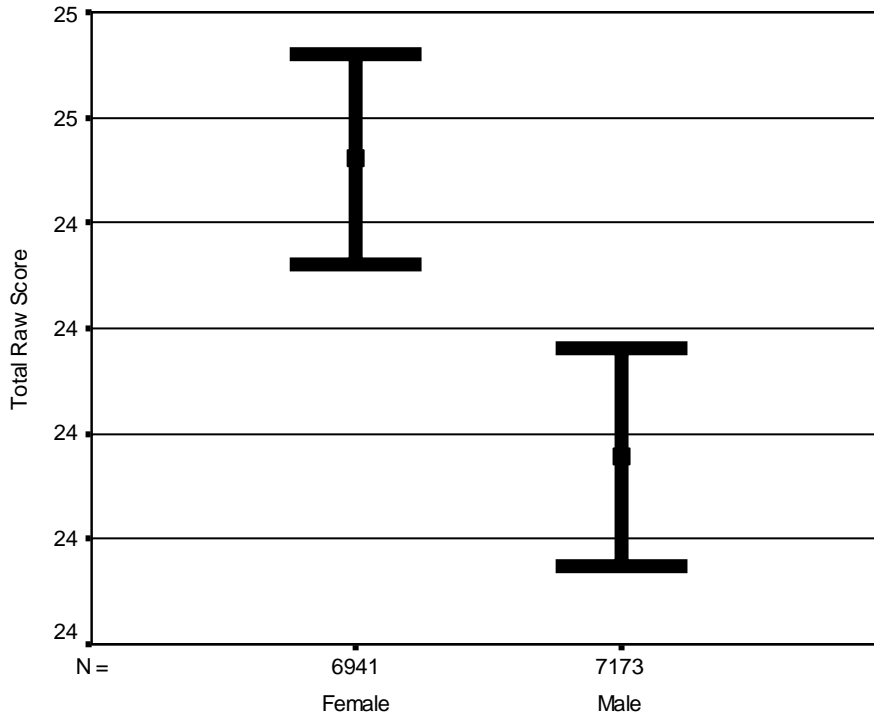


Distribution of sampling means - Grade 5 - Gender

In this section, test scores are compared across gender¹⁴. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

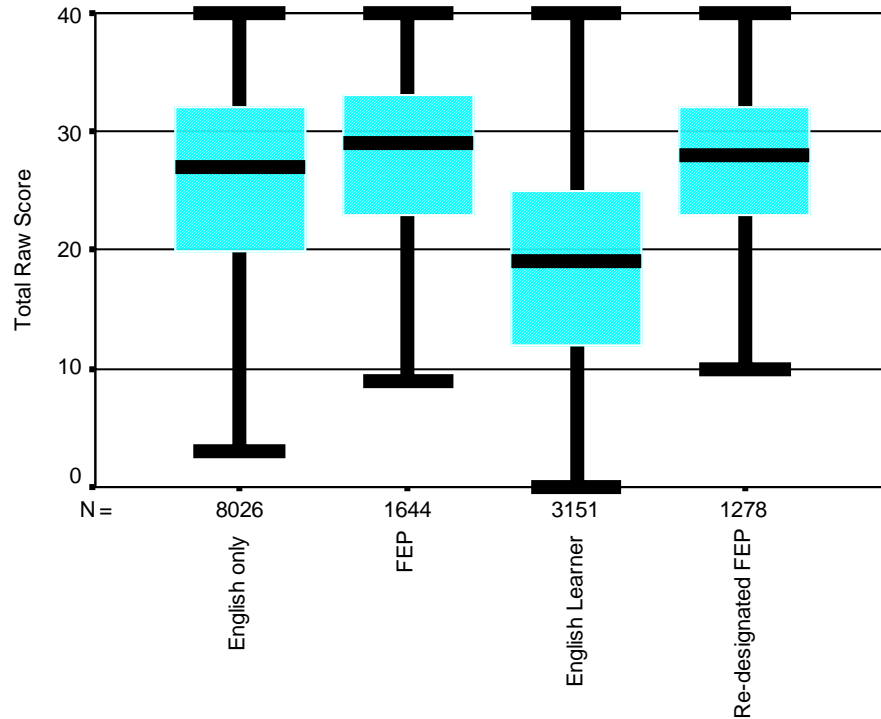
The scores of the females were significantly higher than the scores of the males.

Distribution of sampling means - Grade 5 - Gender



¹⁴ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 5 - Language Fluency



Distribution of sampling means - Grade 5 - Language Fluency

In this section, test scores are compared across groups of different language fluency¹⁵. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly lower than those of students described as Full English Proficiency (FEP) and those described as Re-designated FEP, and significantly higher than those students described as English Learners.

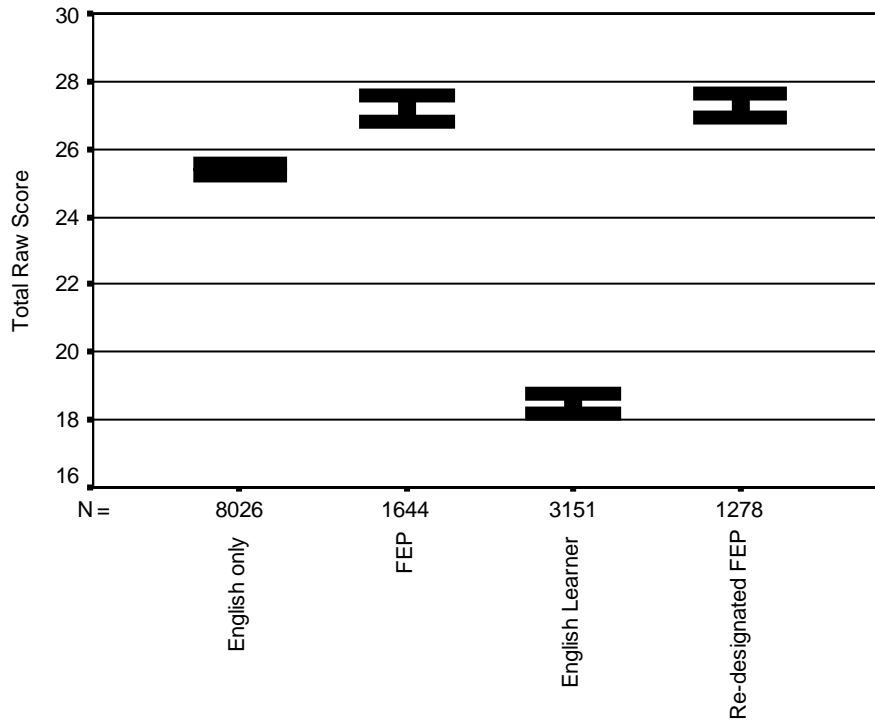
The scores of students in the FEP category are significantly higher than those of students in the English only and English Learner categories. The scores of students in the FEP category are not significantly different than the scores of students in the re-designated FEP group.

The scores of students in the English Learner category are significantly lower than those of the students in all other Language Fluency categories.

¹⁵ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

The scores of students in the Re-designated FEP category are significantly higher than those of students in all the English Only and English Learner categories, but not significantly different from those in the FEP category of Language Fluency.

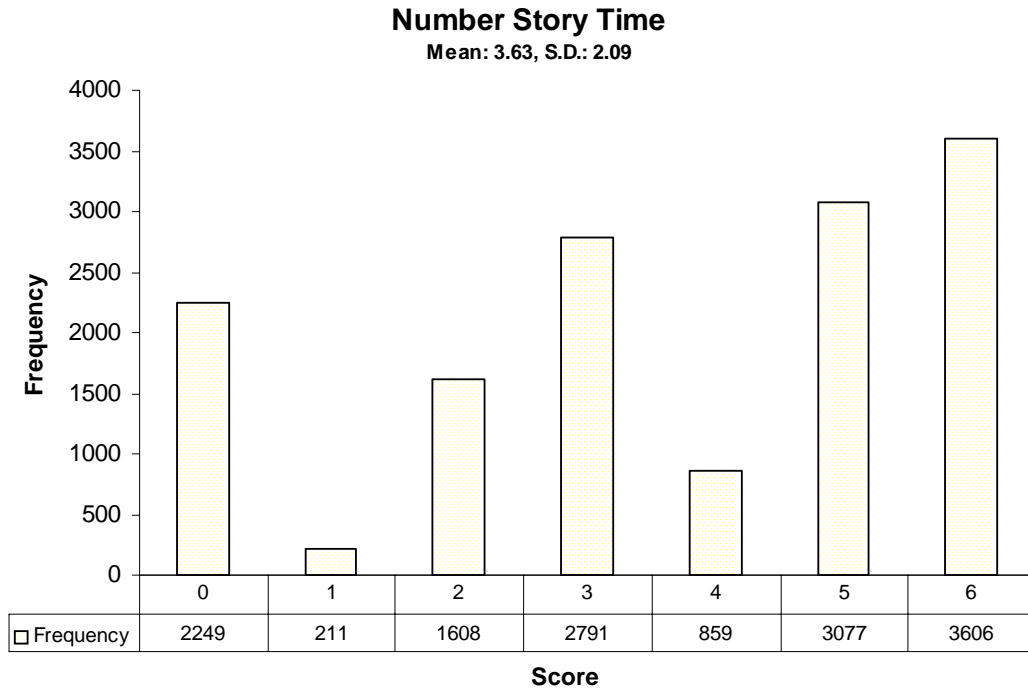
Distribution of sampling means - Grade 5 - Language Fluency



5th grade Task 1 Number Story Time

Student Task	Solve a multi-step money story problem. Write a story problem to illustrate a division problem with a remainder amount.
Core Idea 2 Number Operations	<p>Understand the meanings of operations and how they relate to each other, make reasonable estimates and compute fluently.</p> <ul style="list-style-type: none"> Reason about and solve problem situations that involve more than one operation in multi-step problems Develop fluency in dividing whole numbers Understand the meaning of remainders by modeling division problems

Frequency Distribution for each Task – Grade 5 - Number Story Time



Score:	0	1	2	3	4	5	6
% < =	15.6%	17.1%	28.2%	47.6%	53.6%	75.0%	100.0%
% > =	100.0%	84.4%	82.9%	71.8%	52.4%	46.4%	25.0%

The maximum score available for this task is 6 points.
The cut score for a level 3 response is 4 points.

Most students (about 85%) could show a correct process for finding the change. About half the students could solve correctly the number story problem about change and write their own number story that required dividing by 5. 25% of the students met all the demands of the task. About 15% of the students scored no points on this task.

Based on teacher observations, this is what fifth grade students seemed to know and be able to do:

- Determine a correct process for adding expenditures and finding change using decimals
- Write a sharing or partitive number story
- Understand the purpose of number operations and choose them appropriately when creating and solving problems

Areas of difficulty for fifth graders, fifth grade students struggled with:

- Monetary notation
- Writing questions to include dealing with the remainder
- Knowing or being comfortable with a variety of division situations and contexts

Instructional Implications:

Students need to learn to use proper monetary notation. Students also need to develop an understanding of different situations that require division of whole numbers with remainders. Their classroom experiences should include a variety of all the major types of division problems. Besides sharing or partitive, another important type of division is the measurement model, “How many groups fit into . . .?” A third type of division is products and factors (e.g. if a rectangle has an area of 60 sq. m. and one side is 20 m. How long is the other side?) Students need more experiences with writing word problems to match number sentences. Problem posing is an important skill to develop a deeper understanding of operations.

5th grade

Task 2

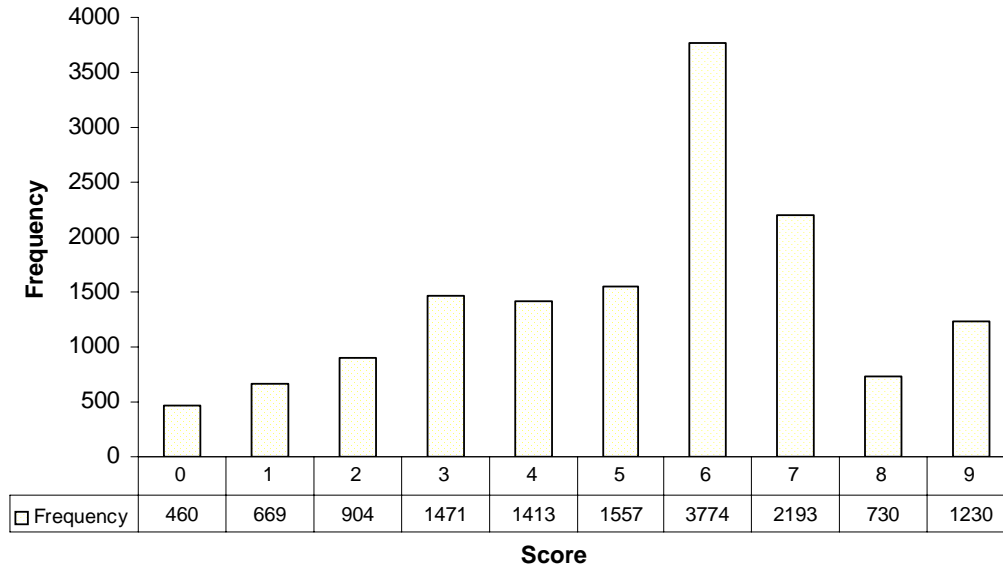
Raspberry Cake

Student Task	Halve the amounts of ingredients in a recipe. Solve a simple fraction problem in a practical context.
Core Idea 1 Number Properties	<p>Understand numbers, ways of representing numbers, relationships among numbers, and number systems.</p> <ul style="list-style-type: none"> • Use models, benchmarks, and equivalent forms to judge the size of fractions • Understand the place-value structure of the base-ten number system including being able to represent and compare rational numbers
Core Idea 2 Number Operations	<p>Understand the meanings of operations and how they relate to each other, make reasonable estimates and compute fluently.</p> <ul style="list-style-type: none"> • Develop and use strategies to solve problems involving number operations with fractions and decimals relevant to students’ experience

Grade 5 – Raspberry Cake

Raspberry Cake

Mean: 5.22, S.D.: 2.29



Score:	0	1	2	3	4	5	6	7	8	9
% < =	3.2%	7.8%	14.1%	24.3%	34.1%	45.0%	71.2%	86.4%	91.5%	100.0%
% > =	100.0%	96.8%	92.2%	85.9%	75.7%	65.9%	55.0%	28.8%	28.8%	8.5%

The maximum score available for this task is 9 points.

The cut score for a level 3 response is 5 points.

Most students (about 86 %) could cut the recipe in half. They may have made an error in finding the amount of cream or raspberries, which required working with fractional parts. More than half the students (about 66%) could cut the recipe in half with error in cream and/or raspberries and draw a model to represent fraction parts of a cake in a story problem. About 15% of the students could correctly find all the values for halving the recipe, draw a diagram and find the amount of raspberries needed to make a cake for 20 people. About 8% of the students met all the demands of the task. About 3% of the students scored no points on the task.

Instructional Implications:

Students need more practice working with mixed numbers and understanding the quantity represented by mixed numbers. Students at this grade level need to understand proportional reasoning and apply this in a practical context (e.g. increasing or decreasing quantities such as doubling, halving, and two-and-a-half times whole numbers, fractions, and mixed numbers). Students should have a variety of strategies to help them make sense of fractions, such as drawing pictures, decomposing them into simpler parts, and using algorithms. Students at this grade level should develop a facility with doing operations with common fractions and mixed numbers as mental math activities.

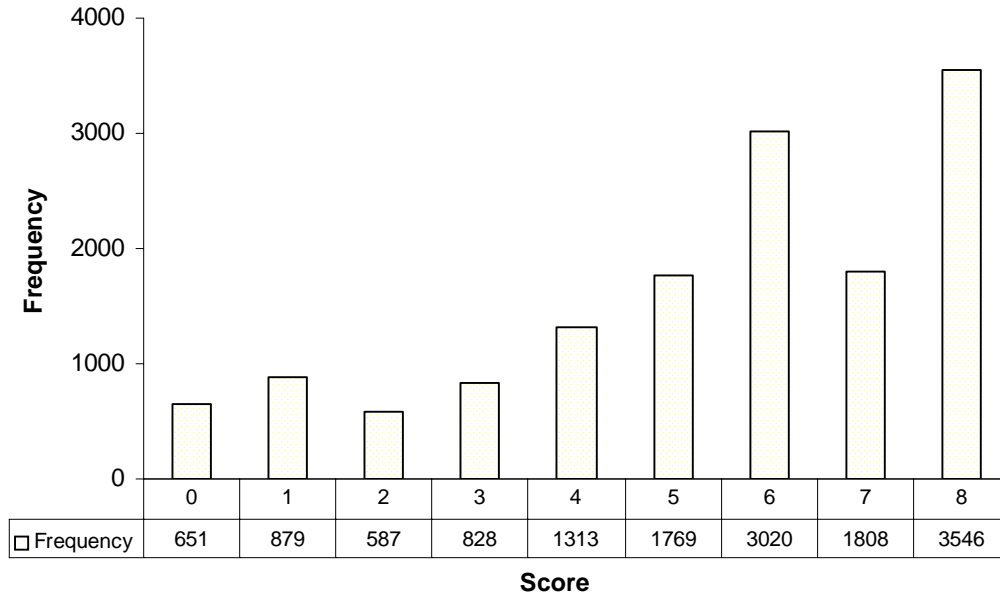
5th grade**Task 3****Buttons**

Student Task	Use a button arrangement pattern to describe, extend, and make generalization about its numeric pattern.
Core Idea 3 Patterns and Functions	Understand patterns and use mathematical models such as algebraic symbols and graphs to represent and understand quantitative relationships. <ul style="list-style-type: none">• Describe and extend numeric patterns (3rd grade)• Represent and analyze patterns and functions using words (4th grade)• Investigate how a change in one variable relates to a change in a second variable

Grade 5 – Buttons

Buttons

Mean: 5.40, S.D.: 2.36



Score:	0	1	2	3	4	5	6	7	8
% < =	4.5%	10.6%	14.7%	20.4%	29.6%	41.9%	62.8%	75.4%	100.0%
% > =	100.0%	95.5%	89.4%	85.3%	79.6%	70.4%	58.1%	37.2%	37.2%

The maximum score available for this task is 8 points.

The cut score for a level 3 response is 4 points.

Most students (about 89%) could draw pattern 4 and give the correct number of buttons for pattern 5 and pattern 6. Many students (about 80%) could draw and extend the pattern and explain in words how the patterned worked. More than half the students (about 60%) could draw the pattern, extend the pattern for 5 and 6 and explain how it grew, find and explain the number of white buttons for pattern 11, and could find the total buttons for pattern number 24. Almost 26% of the students met all the demands of the task. About 5% of the students scored no points on this task.

Based on teacher observations, this is what fifth grade students know and are able to do:

- Continue a pattern using pictures and numbers
- Explain how a pattern grows and use that algorithm to solve for larger numbers in the pattern

Areas of difficulty for fifth graders, fifth grade students struggled with:

- Distinguishing between part of a pattern and the whole pattern
- Explaining a pattern in words

Instructional Implications:

Fifth grade students need more experiences that require them to move beyond drawing the next figure in the pattern so that they analyze the pattern and represent the growth numerically. Fifth graders need to move beyond thinking about “what comes next?” to thinking about the problem as a whole: this involves generalizing what is happening with the growth, but need not necessarily involve variables or algebraic equations. Being able to see what remains the same and what changes in a pattern helps students develop algebraic thinking and the ability to make a generalization. Asking questions about how the pattern changes helps students to move beyond counting and drawing strategies to rules that will solve for any number in the pattern.

5th grade

Task 4

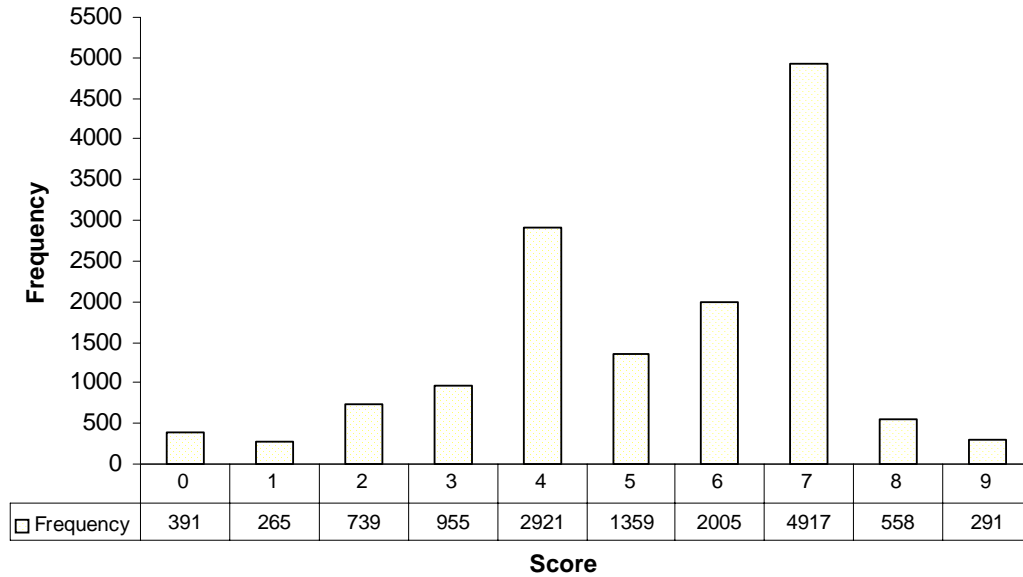
Winter Sports

Student Task	Use two tables of winter sports information to interpret data and to make calculations.
Core Idea 5 Data Analysis	Collect, organize, represent and interpret numerical and categorical data, and clearly communicate their findings. Display, analyze, compare, and interpret different data sets. <ul style="list-style-type: none"> • Interpret data to answer questions about a situation (4th grade) • Compare different sets of data.
Core Idea 3 Patterns, Functions, And Algebra	Understand quantitative relationships. <ul style="list-style-type: none"> • Identify and describe situations with rates and compare them.
Core Idea 2 Number Operations	Understand the meanings of operations and how they relate to each other, make reasonable estimates and compute fluently. <ul style="list-style-type: none"> • Reason about and solve problem situations that involve more than one operation in multi-step problems

Grade 5 – Winter Sports

Winter Sports

Mean: 5.32, S.D.: 1.99



Score:	0	1	2	3	4	5	6	7	8	9
% <=	2.7%	4.6%	9.7%	16.3%	36.6%	46.0%	60.0%	94.1%	98.0%	100.0%
% >=	100.0%	97.3%	95.4%	90.3%	83.7%	63.4%	54.0%	40.0%	5.9%	2.0%

The maximum score available for this task is 9 points.

The cut score for a level 3 response is 5 points.

Most students (about 81%) were able to use information on the table to find missing values in the table and answer simple questions about the table. A little more than half the students (about 54%) could fill in the table, answer questions about the table, use a weighted point system to find the points for Finland and explain how they solved for it. A little less than half of the students could also find the point total for Norway. Few students (less than 3%) could make a mathematical comparison to show how they knew Norway won. Less than 3% of the students scored no points on this task.

Based on teacher observations, this is what fifth grade students seemed to know and be able to do:

- Answer simple questions about a table.
- Fill in missing values in a table.
- Use a weighted point value to find a score.

Areas of difficulty for fifth graders, fifth grade students struggled with:

- Using a weighted point value to find the points for Norway.
- Making a comparison by showing values for countries other than the leader, not referencing other countries in the comparison set

Implications for Instruction:

Students need to be able to interpret and use data from different tables to make calculations. Students should have experiences working with weighted values and be exposed to problem-solving situations that highlight the need for them. At this grade level students need to become familiar with what makes a good comparison. To convince someone, you need to show more than just the best solution, but also why other choices are not as favorable

5th grade

Task 5

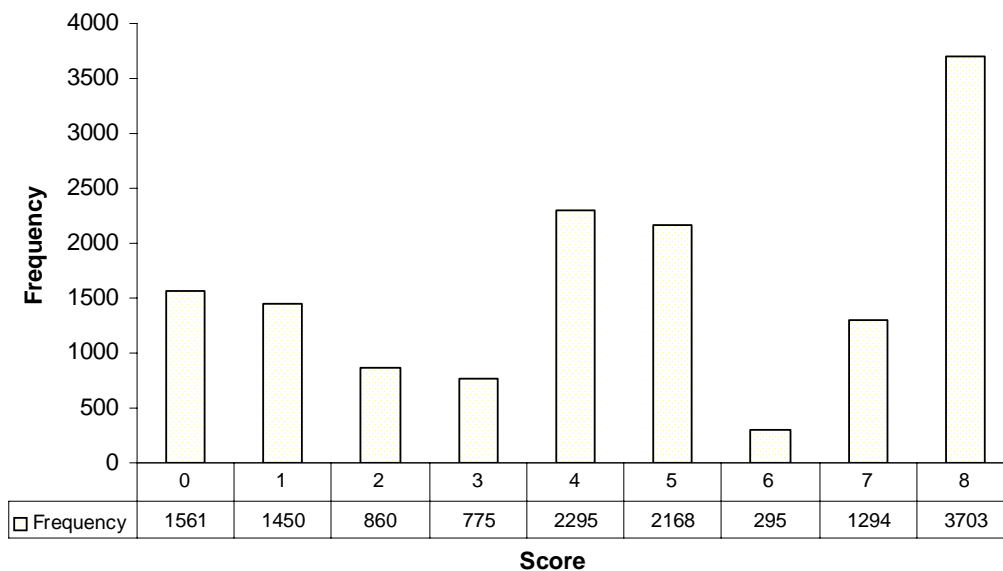
Juan’s Shapes

Student Task	Use a grid to find the perimeter and area of shapes. Draw shapes of same area/perimeter on a grid.
Core Idea 4 Geometry and Measurement	<p>Students will analyze characteristics and properties of two- and three-dimensional geometric shapes, understand attributes, and apply appropriate techniques, tools, and formulas to determine measurements.</p> <ul style="list-style-type: none"> • Understand such attributes as length, area, weight, volume, and angle size and select the appropriate type of unit for measuring each attribute • Develop strategies for estimating or calculating the perimeters and areas of irregular shapes • Explore and determine what happens to perimeter and area of a two-dimensional figure when its shapes changed in some way.

Grade 5 – Juan’s Shapes

Juan's Shapes

Mean: 4.58, S.D.: 2.78



Score:	0	1	2	3	4	5	6	7	8
% < =	10.8%	20.9%	26.9%	32.3%	48.2%	63.3%	65.3%	74.3%	100.0%
% > =	100.0%	89.2%	79.1%	73.1%	67.7%	51.8%	36.7%	34.7%	25.7%

The maximum score available for this task is 8 points.

The cut score for a level 3 response is 4 points.

Most students (about 90%) could find the area of shapes on a grid. Many students (about 70%) could find area of shapes on a grid and draw a different shape with the same area. A little less than half of the students (about 40%) could find area and perimeter for shapes on a grid and draw a shape either maintaining the same area or maintaining the same perimeter. About 26% of the students met all the demands of the task. About 11% of the students scored no points on this task.

Based on teacher observations, this is what fifth grade students seemed to know and be able to do:

- Count to find the area and perimeter of simple shapes
- Design a shape to satisfy a constraint about area

Areas of difficulty for fifth graders, fifth grade students struggled with:

- Finding perimeters around concave angles
- Designing a shape with a constraint about perimeter

Instructional Implications:

Students need to understand the difference between area and perimeter. They need to be able to calculate each for the shapes drawn on the grid. Students also need to be able to construct simple shapes with certain area and perimeter restrictions. Most students with this score had difficulty with finding the perimeter of shapes and designing their own shapes. Practice measuring real objects in different units would be beneficial

Blue Print for Sixth Grade MAC

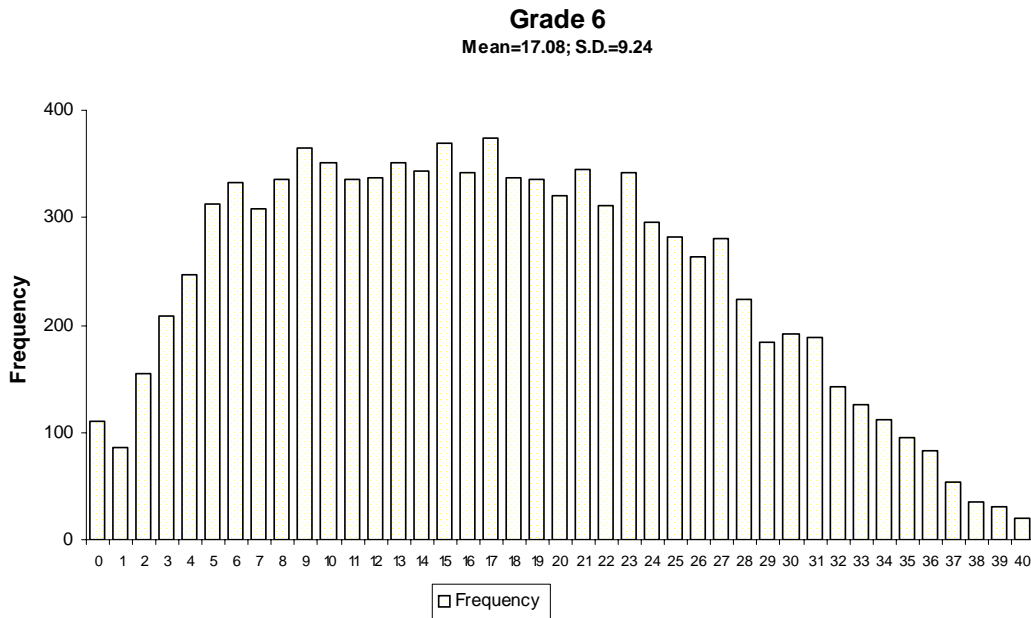
Alignment of Tasks to NCTM Content and Process Standards

	Baseball Players	Gym	Square Elk	Spinners	Rabbit Costumes
NUM	S	S		S	P
ALG		P			S
GEO			S		
MEAS			P		S
DATA	P	S		P	
PS	*	*	*	*	*
REAS	*	*	*	*	*
COMM	*	*		*	*
CONN		*	*		
REP					

NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

- P denotes Primary NCTM Content Standard
- S denotes Secondary NCTM Content Standard
- denotes NCTM Process Standard

Overall Frequency Distribution by Total Score



Level Frequency Distribution Chart and Frequency Distribution

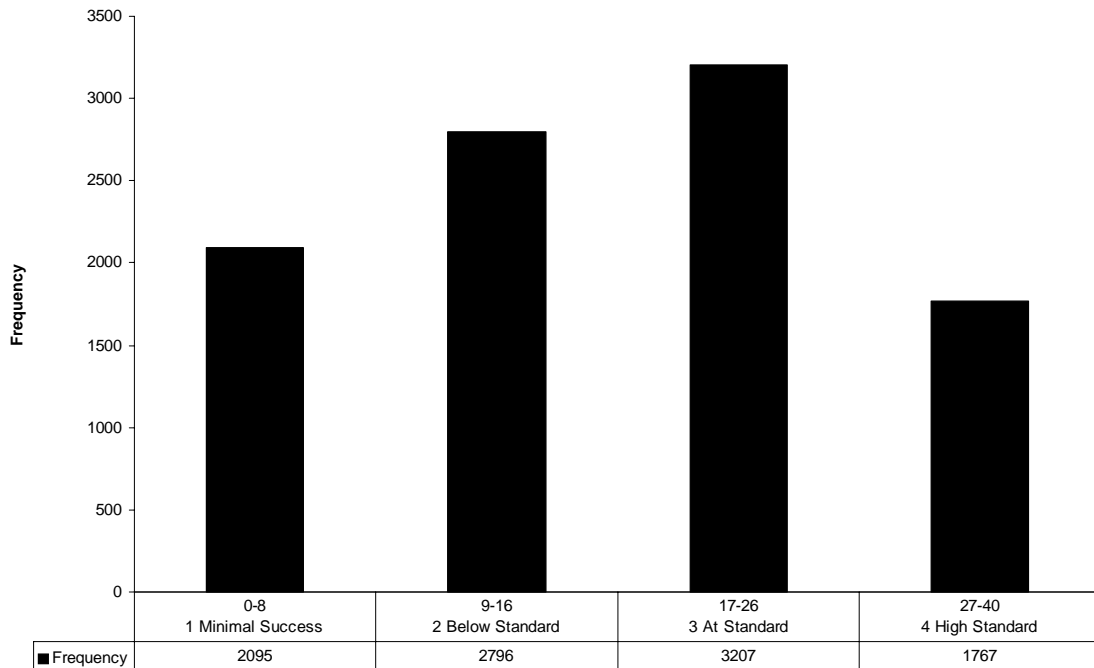
2003 - Numbers of students tested: Grade 6: 9865

Grade 6 2000 - 2001

Level	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	26%	100%	16%	100%
2	47%	74%	43%	84%
3	17%	27%	28%	41%
4	10%	10%	13%	13%

Grade 6 2002 - 2003

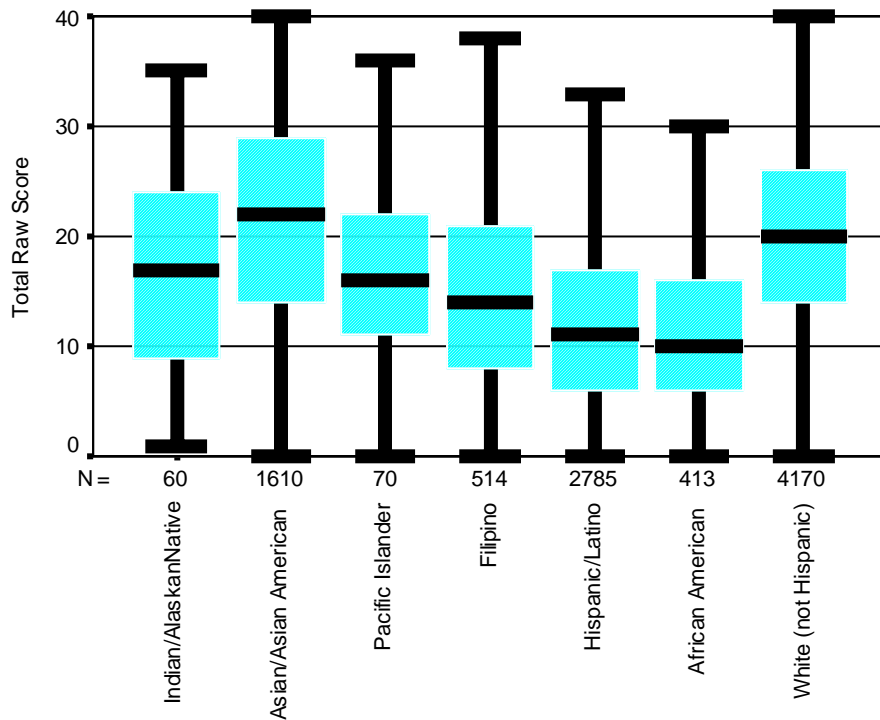
Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	21%	100%	21%	100%
2	25%	79%	28%	79%
3	36%	55%	33%	50%
4	19%	19%	18%	18%



Grade 6 Distribution by Performance Levels

Grade 6 - Ethnicity

The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values* within a category lie between the highest and lowest horizontal bars.



Distribution of sampling means - Grade 6 - Ethnicity

In this section, test scores are compared across different ethnic groups¹⁶. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if the scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than group B. When the two scores overlap, then there is not significant difference between the groups.

The scores of Indian/Alaskan Native students are significantly lower than those of Asian/Asian American students, and are significantly higher than those of Hispanic/Latino, and African American students. There are no other significant differences between the scores of Indian/Alaskan Native students and any other ethnic group.

The scores of Asian/Asian American students are significantly higher than the scores of students of all other ethnic groups.

The scores of Pacific Islander students are significantly lower than those of the Asian/Asian American and White students, and significantly higher than those of the Hispanic/Latino and African American students. There are no other significant differences between the scores of Pacific Islander students and any other ethnic group.

The scores of Filipino students are significantly lower than those of the Asian/Asian American and White students, and significantly higher than those of the Hispanic/Latino and African American students. There are no significant differences between the scores of Filipino students and Indian/Alaskan Native students or Pacific Islanders.

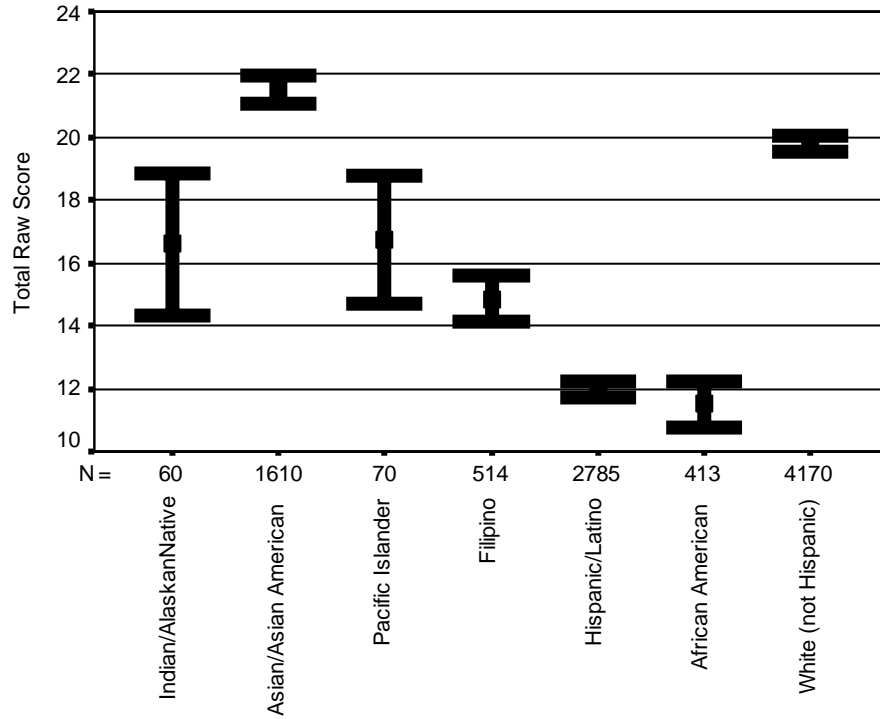
The scores of Hispanic/Latino students are significantly lower than those of all other ethnic groups except African American. The difference between the Hispanic/Latino scores and the African American student scores is not statistically significant.

The scores of African American students are significantly lower than those of all other ethnic groups except Hispanic/Latino. The difference between the African American scores and the Hispanic/Latino student scores is not statistically significant.

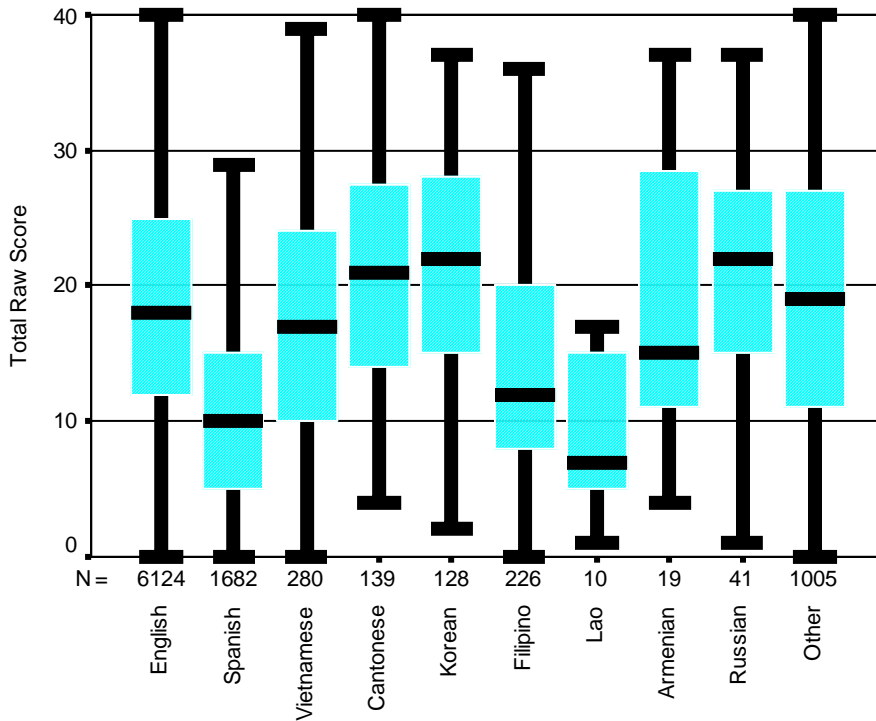
The scores of White students are significantly higher than those of Pacific Islander, Filipino, Hispanic/Latino, and African American students. The scores of White students were significantly lower than those of the Asian/Asian American students, and not significantly different from the scores of Indian/Alaskan Native students.

¹⁶ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 6 - Ethnicity



Grade 6 - Home Language



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 6 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home¹⁷. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with a home language of Cantonese, Korean, or “Other”. The scores of students with English as a home language are significantly higher than those of students with a home language of Spanish or Filipino. There is no significant difference between scores of students with a home language of English, and those with a home language of Vietnamese, Khmer, Lao, Armenian, or Russian.

The scores of students with Spanish as a home language are significantly lower than those of students with any other home language except Lao. The difference between students with Spanish as a home language and those with Lao as a home language is not significant.

The scores of students with Vietnamese as a home language are significantly higher than those of students with Spanish or Filipino as a home language, and lower than those of students with Cantonese, Korean, or “Other” as a home language. There are no other significant differences between the scores of students with Vietnamese as a home language and any other group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or Lao as a home language. There are no other significant differences between the scores of students with Cantonese as a home language and any other group.

The scores of students with Korean as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or Lao as a home language. There are no other significant differences between the scores of students with Korean as a home language and any other group.

The scores of students with Filipino as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a

¹⁷ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

home language, and significantly higher than those of students with Spanish as a home language. There are no other significant differences between the scores of students with Filipino as a home language and any other group.

The scores of students with Khmer as a home language are significantly lower than the scores of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a home language.

The scores of students with Lao as a home language are significantly lower than the scores of students with Cantonese, Korean, Russian, or “Other” as their home language, but there are no other significant differences with any other home language category.

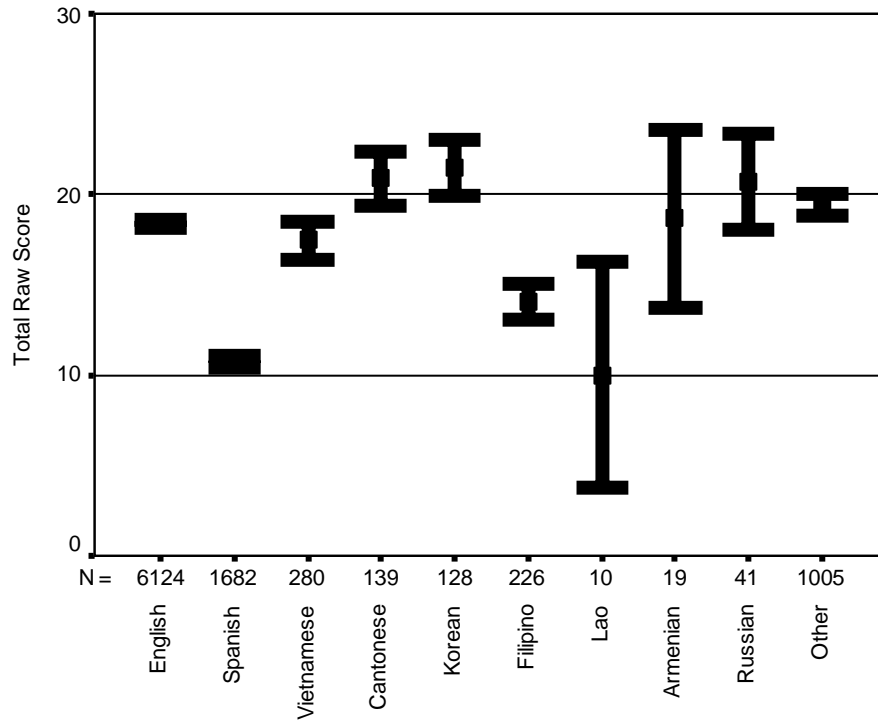
The scores of students with Armenian as a home language are significantly higher than the scores of students with Spanish as their home language, but there are no other significant differences with any other language group.

The scores of students with Russian as a home language are significantly higher than those of students with Spanish, Filipino, or Lao as a home language. There are no other significant differences between the scores of students with Russian as a home language and any other language group.

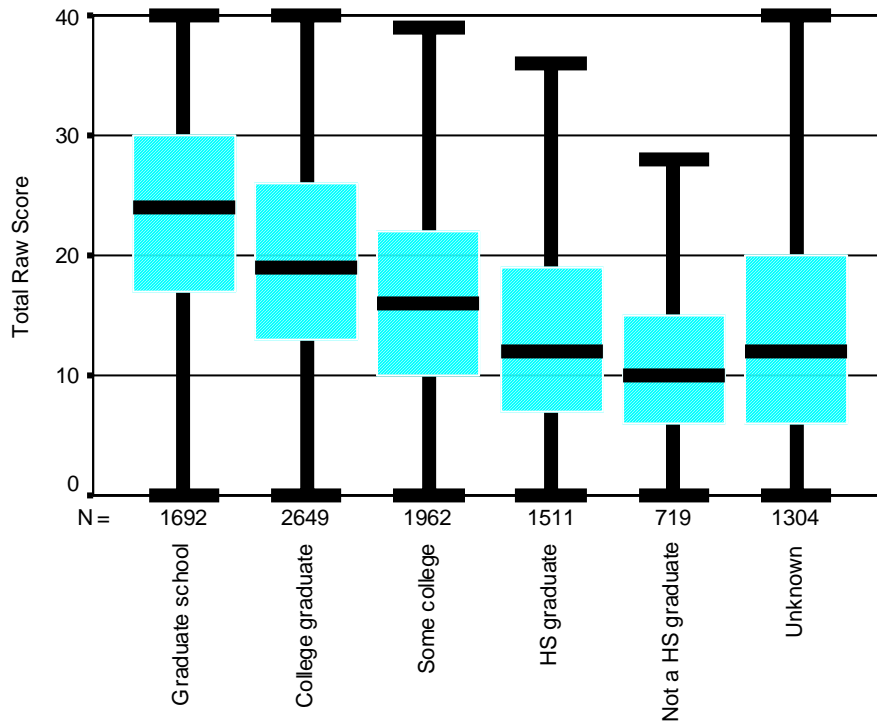
The scores of students with “Other” as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or Lao as a home language. There are no other significant differences between the scores of students with “Other” as a home language and any other language group.

Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 6 - Home Language



Grade 6 - Parent Education



Distribution of sampling means - Grade 6 -Parent Education

In this section, test scores are compared across groups of different levels of parent education¹⁸. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate school education are significantly higher than those of students in all other Parent Education categories.

The scores of students whose parents have a college education are significantly lower than those whose parents have a graduate education, and significantly higher than all other groups.

The scores of students whose parents have some college education are significantly higher than those of students whose parents are not High School graduates, those whose parents are High School graduates, and those whose parents' education is "unknown", and significantly lower than those whose parents have a college education or a graduate education.

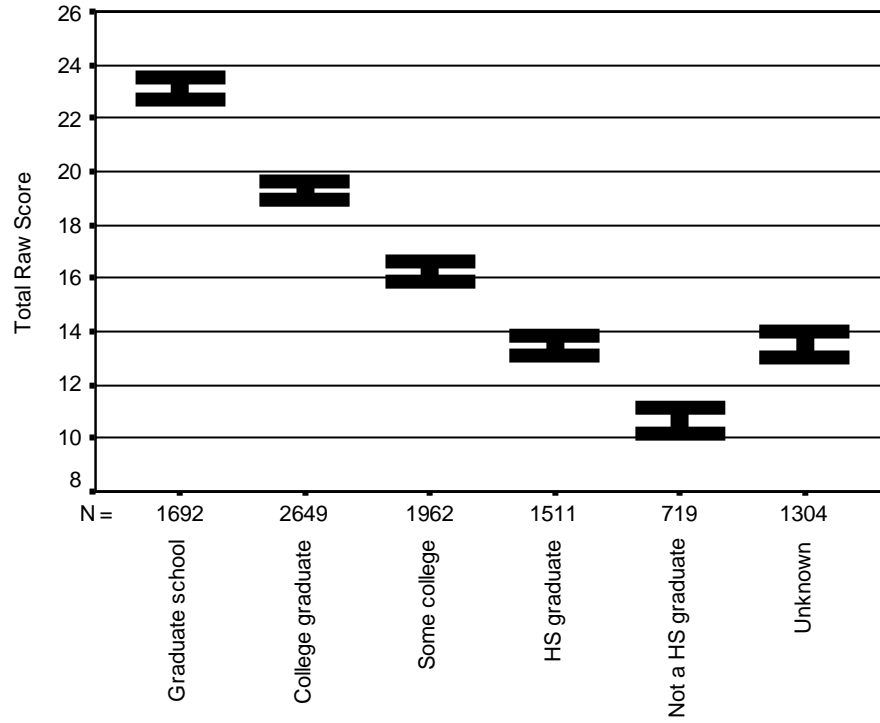
The scores of students whose parents are High School graduates are significantly higher than those of students whose parents are not High School graduates, and significantly lower than those whose parents have some college education, a college education or a graduate education. They are not significantly different from scores of students whose parent education level is unknown.

The scores of students whose parents are not High School graduates are significantly lower than the scores of students in all other Parent Education categories.

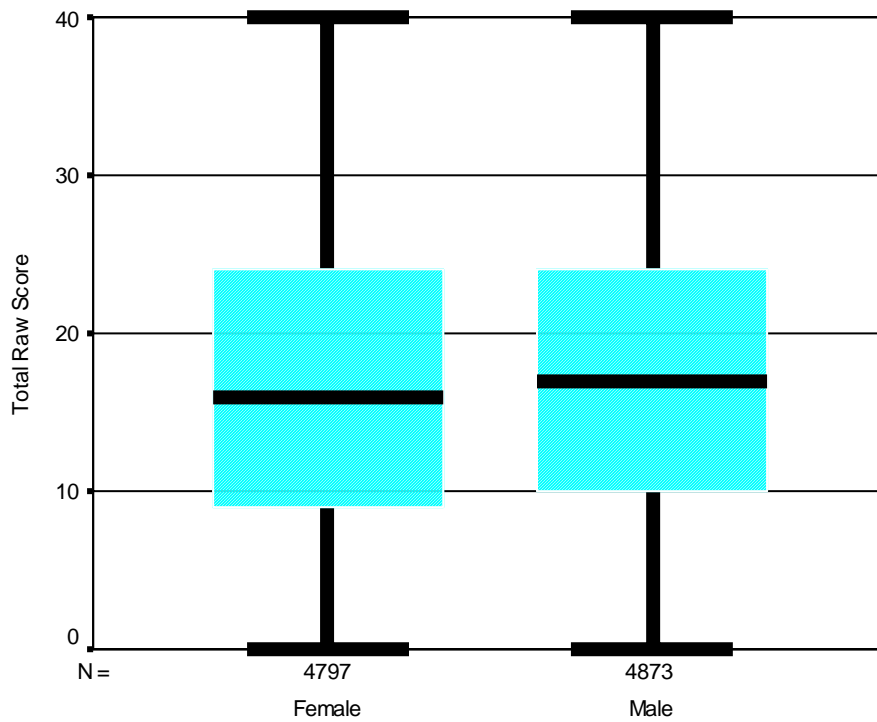
The scores of students whose parents' education is unknown are significantly higher than those of students whose parents are not High School graduates, and significantly lower than those whose parents have some college education, a college education or a graduate education. They are not significantly different from scores of students whose parents are High School graduates.

¹⁸ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 6 - Parent Education



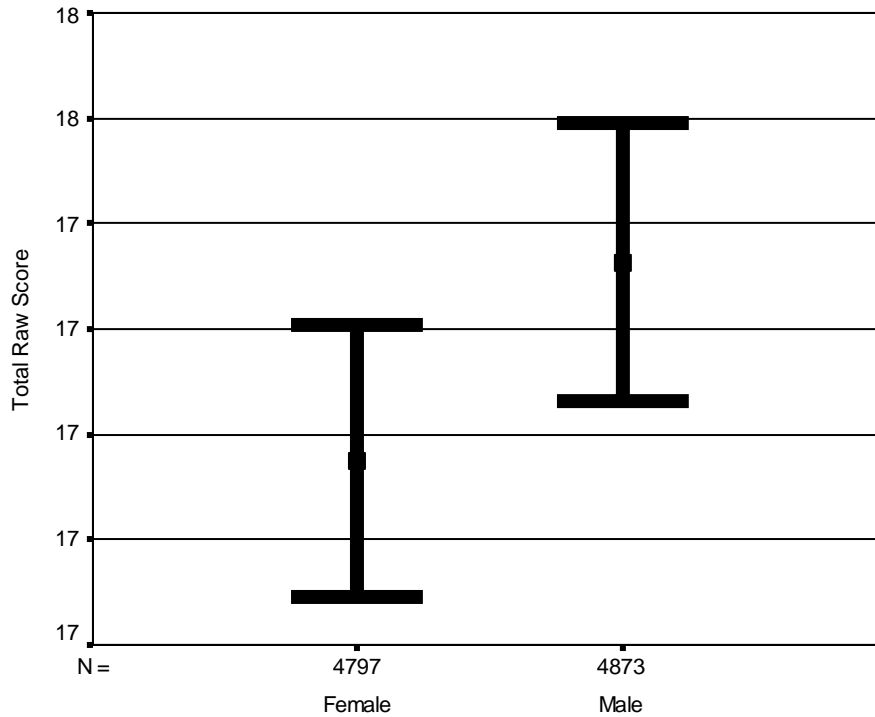
Grade 6 - Gender



Distribution of sampling means - Grade 6 - Gender

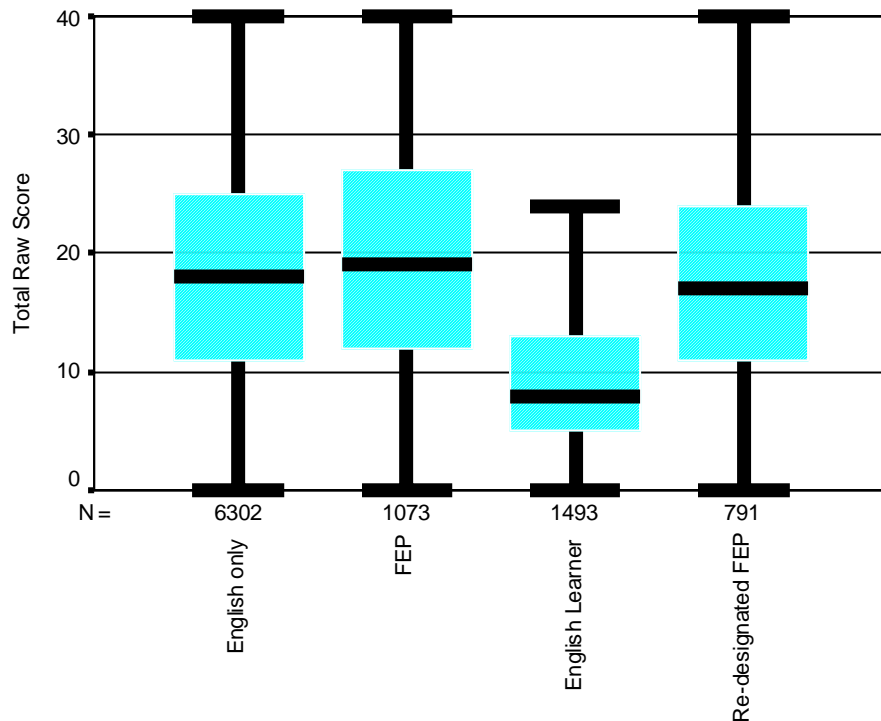
In this section, test scores are compared across gender¹⁹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

There is no statistically significant difference in the scores of females and males.



¹⁹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 6 - Language Fluency



Distribution of sampling means - Grade 6 - Language Fluency

In this section, test scores are compared across groups of different language fluency²⁰. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly lower than those of students described as Full English Proficiency (FEP), and significantly higher than those students described as English Learners. The scores of students with English Only are not significantly different from those of students in the re-designated FEP group.

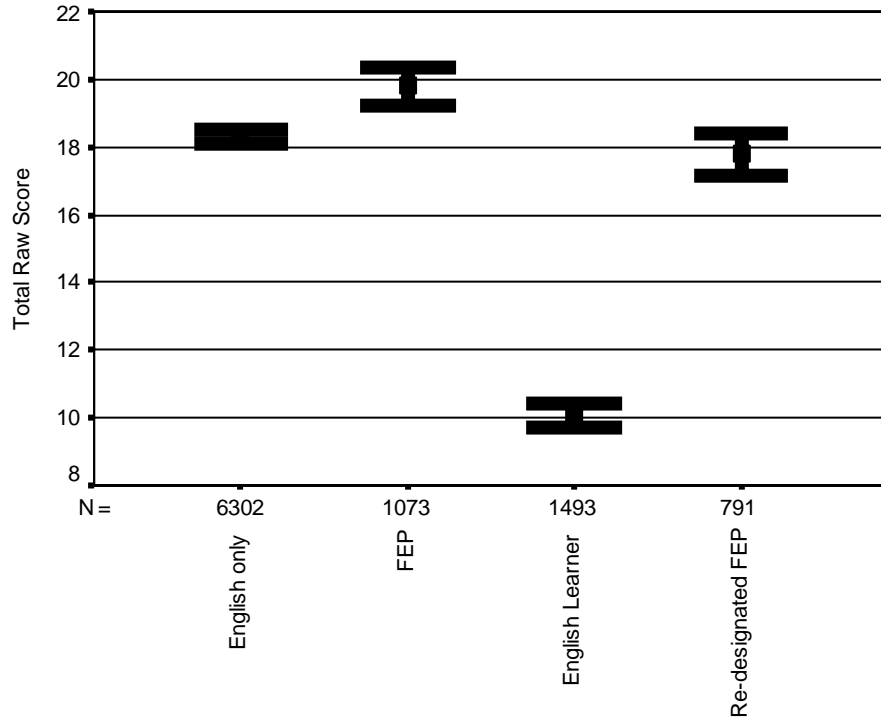
The scores of students in the FEP category are significantly higher than those of students in any other category.

The scores of students in the English Learner category are significantly lower than those of the students in all other Language Fluency categories.

The scores of students in the Re-designated FEP category are significantly higher than those of students in the English Learner group, lower than those in the FEP group, and not significantly different from student scores from the English Only category of Language Fluency.

²⁰ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 6 - Language Fluency



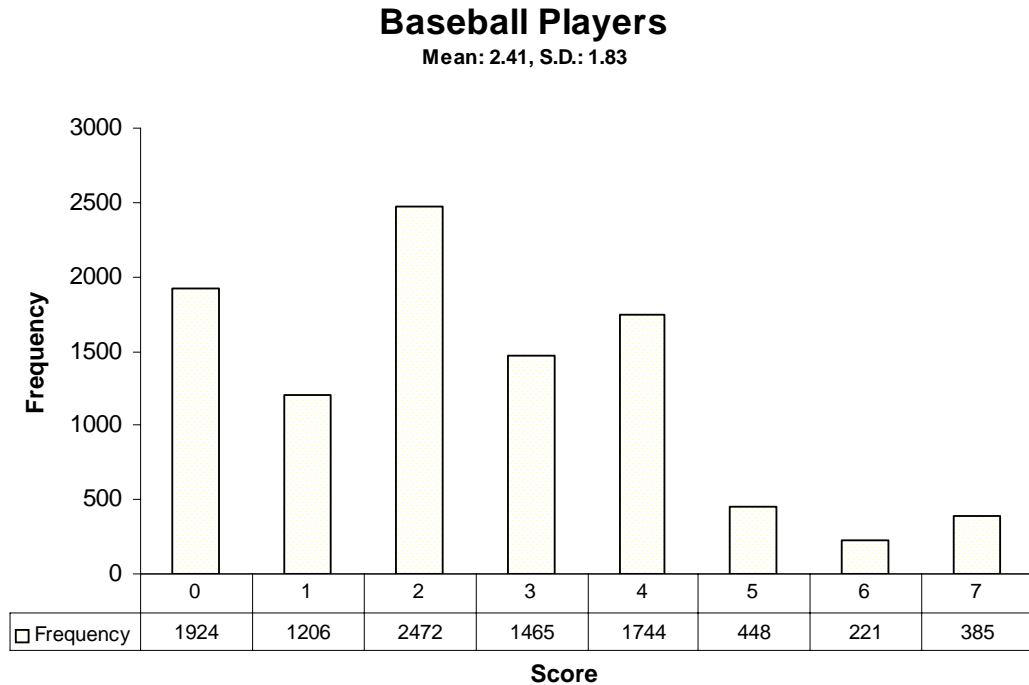
6th grade

Task 1

Baseball Players

Student Task	Determine measures of center and spread for two baseball teams.
Core Idea 5 Statistics	<p>Select and use appropriate statistical methods to display, analyze, compare and interpret different data sets.</p> <ul style="list-style-type: none"> Use measures of center and spread (mean, median, and range) and understand what each does and does not indicate about the data set. (5th grade)

Frequency Distribution for each Task – Grade 6
Grade 6 – Baseball Players



Score:	0	1	2	3	4	5	6	7
% < =	19.5%	31.7%	56.8%	71.6%	89.3%	93.9%	96.1%	100.0%
% > =	100.0%	80.5%	68.3%	43.2%	28.4%	10.7%	6.1%	3.9%

The maximum score available for this task is 7 points.
The cut score for a level 3 response is 3 points.

Most students (about 80%) could find the range or median of a set of numbers. A little less than half (about 40%) could find the total weight of players given the average weight and find the range or median. 6% of the students could meet all the demands of the task. Almost 20% of the students scored no points on this task. 90% of those students attempted the problem.

Based on teacher observations, this is what sixth grade students seemed to know and be able to do:

- Find the total weight of players if they knew the average weight

- Find the median

Areas of difficulty for sixth graders, sixth grade students struggled with:

- Finding the range
- Understanding the relationship between individual weights and average weights

- Thinking about how additional data affects the average, how new weights must be higher than the average to raise the total average

Instructional Implications:

Students need more understanding of the meaning and application of mean (average) in a real situation. While many students can calculate the mean, they don't understand the relationship of that answer to the problem. They can't work backwards from the average to the total. This makes it difficult for them to reason about how changes in the situation will affect the average. Students need more experiences understanding what the mean reflects about the data and how changes in data will affect the average. The purpose of statistics is to give a picture about the data. Students need to be able to use other measures of center, like median and range, to help them make sense of the situation.

6th grade

Task 2

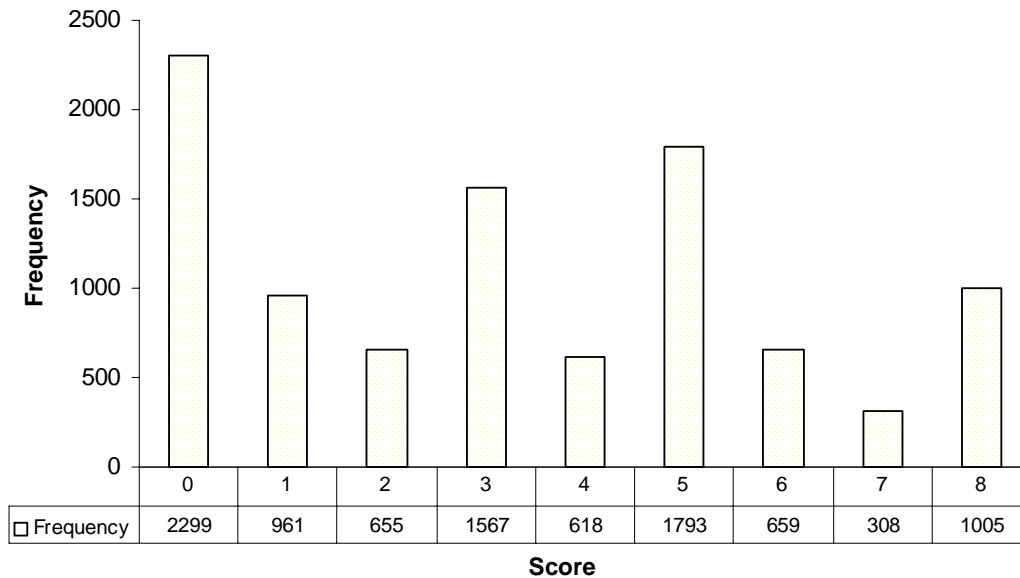
Gym

Student Task	Analyze gym membership costs to solve a practical money problem.
Core Idea 3 Algebra and Functions	Understand relations and functions, analyze mathematical situations, and use models to solve problems involving quantity and change. <ul style="list-style-type: none"> • Model and solve contextualized problems using various equations
Core Idea 1 Number and Operation	Understand number systems, the meanings of operations, and ways of representing numbers, relationships, and number systems <ul style="list-style-type: none"> • Understand and use proportional reasoning to represent quantitative relationships

Grade 6 – Gym

Gym

Mean: 3.30, S.D.: 2.63



Score:	0	1	2	3	4	5	6	7	8
% < =	23.3%	33.0%	39.7%	55.6%	61.8%	80.0%	86.7%	89.8%	100.0%
% > =	100.0%	76.7%	67.0%	60.3%	44.4%	38.2%	20.0%	13.3%	10.2%

The maximum score available for this task is 8 points.
The cut score for a level 3 response is 3 points.

Most students (about 80%) could find the cost of the Pay as you go plan and the All-in-one plan. A little more than half of the students (about 67%) could correctly calculate the costs of all three options and compare to find the better deal. Almost 44% could compare the 3 options and find the number of visits that would make all-in-one and regular deal the same. About 10% of the students could meet all the demands of the task, including comparing the yearly costs and savings of two different gym plans. Almost 20% of the students scored no points on this task. 90% of those students attempted the task.

Based on teacher observations, this is what sixth grade students seemed to know and be able to do:

- Calculate the costs of different options for gym memberships for one month
- Compare costs of the different options
- Find the number of visits needed in one plan to equal the cost with another plan

Areas of difficulty for sixth graders, sixth grade students struggled with:

- Identifying and using multiple constraints within the different gym options
- Analyzing the affects of operations on units (e.g. adding visits to dollars)
- Converting all constraints to the same measure to allow a comparison (e.g. forgetting to converting cost of visits per month to cost of visits per year)

Instructional Implications:

Students need more experiences making comparisons. They need to calculate the full costs of both options before making a decision. It is not enough to just calculate the cost for the favored choice. Students might consider different choices about family entertainment options or cell phone plans to determine the best price. Students lack experience analyzing the effects of operations on units or systems for keeping track of units as they work through steps in a problem. Dimensional analysis is a topic that many teachers need to deal with more explicitly. Students also have trouble identifying the various constraints of within each option and seeing how they interconnect. Students need to work with more problems in context instead of working isolated number calculation skill sets

6th grade

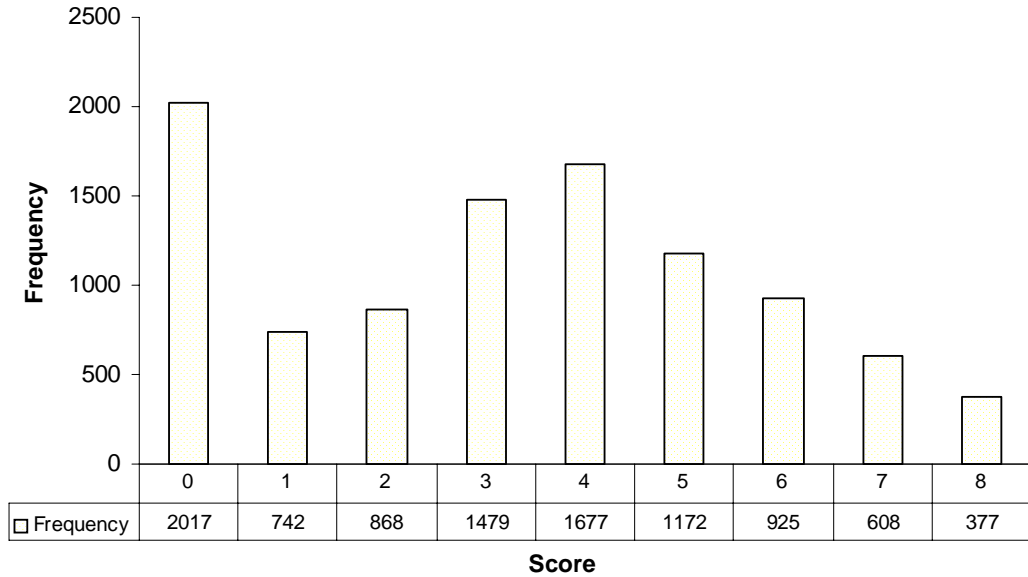
Task 3

Square Elk

Student Task	Find the area and perimeter of letter shapes on a square grid.
Core Idea 4 Geometry and Measurement	Apply the appropriate techniques, tools, and formulas to determine measurements. <ul style="list-style-type: none">• Select and apply techniques and tools to accurately find length and area to appropriate levels of precision• Investigate, describe, and reason about he results of subdividing, combining, and transforming shapes

Grade 6 – Square Elk

Square Elk
 Mean: 3.27, S.D.: 2.36



Score:	0	1	2	3	4	5	6	7	8
% < =	20.4%	28.0%	36.8%	51.8%	68.8%	80.6%	90.0%	96.2%	100.0%
% > =	100.0%	79.6%	72.0%	63.2%	48.2%	31.2%	19.4%	10.0%	3.8%

The maximum score available for this task is 8 points.
 The cut score for a level 3 response is 4 points.

Most students (about 80%) could find one or more of the areas in elk. A little less than half the students (about 49%) could find all the areas in part 1 and 3. Only about 4% of the students could meet all the demands of the task. About 20% of the students scored no points on the task. Most of those attempted the problem.

Based on teacher observations, this is what sixth graders seemed to know and be able to do:

- Count squares to find area
- Understand the relationship between right triangles and rectangles to find area
- Understand that the distance around a shape was equal to the perimeter

Areas of difficulty for sixth graders, sixth grade students struggled with:

- Using a key to find perimeter of shapes representing the hypotenuse of a right triangle
- Recognizing rotations of a shape
- Systematically identifying and counting all the parts of a perimeter

Implications for Instruction:

Students who had difficulty with this problem could not break down the shape into simpler parts. When calculating the area, they could not find the fractional parts. When finding perimeter, they did not recognize the connection between the scale model and the diagonals on the diagram. They need more practice with spatial visualization and rotations. Students at this grade level need to be able to work with more complex shapes. When working with detailed figures, they need to develop strategies for keeping track of their calculations, what’s been done, and what still needs to be calculated.

6th grade

Task 4

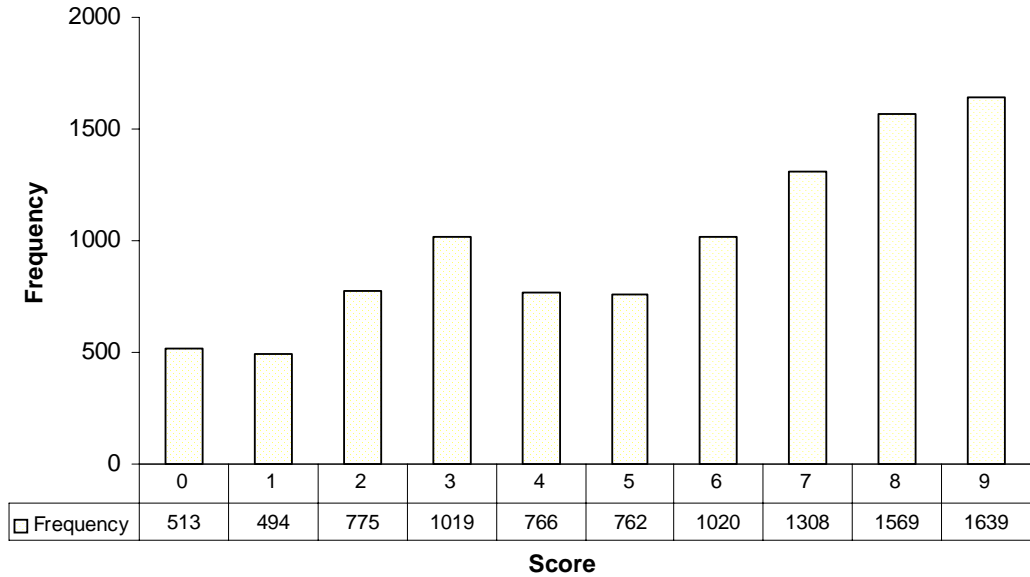
Spinners

Student Task	Work with probabilities for two different spinners. Identify the sample space and the likelihood for certain events.
Core Idea 2 Probability	<p>Demonstrate understanding and use of probability in problem situations.</p> <ul style="list-style-type: none"> • Determine theoretical and experimental probabilities and use these to make predictions about events. • Represent the sample space for a given event in an organized way (e.g., table, diagram, organized list, and tree diagram)

Grade 6 – Spinners

Spinners

Mean: 5.53, S.D.: 2.78



Score:	0	1	2	3	4	5	6	7	8	9
% < =	5.2%	10.2%	18.1%	28.4%	36.2%	43.9%	54.2%	67.5%	83.4%	100.0%
% > =	100.0%	94.8%	89.8%	81.9%	71.6%	63.8%	56.1%	45.8%	56.1%	16.6%

The maximum score available for this task is 9 points.
 The cut score for a level 3 response is 4 points.

Most students (about 82%) could determine which spinner was more likely to get a certain number. Many students (about 72%) could also recognize equally likely events. About half the students could recognize more likely and equally likely events and find all the scores when the numbers on the two spinners were multiplied. Almost 17% of the students could meet all the demands of the task including developing a convincing mathematical argument for why scores would be less than 10. About 5% of the students scored no points on this task. All the students in the sample set attempted this task or did not finish any parts on this task or the rest of test.

Based on teacher observations, this is what sixth grade students seemed to know and be able to do:

- Compare areas on a spinner to determine which is more likely
- Multiply numbers from 2 spinners to find possible outcomes

Areas of difficulty for sixth grade students, sixth graders struggled with:

- Combining areas of noncontiguous sections on a spinner
- Determining equally likely events
- Constructing a convincing, specific mathematical argument for why some scores are not possible
- Organizing information or lists in a manner to determine when all possibilities have been given

Implications for Instruction:

Students need more experiences with probability situations. They need experience analyzing areas with different sizes to determine if events are equally likely. They need to be able to visualize how sections can be divided or combined on a spinner and still have the same probability. Students often think the position on the spinner affects the probability, e.g. if a red space appears on both sides of a spinner, then it is more likely than if two red spaces of the same size appear side-by-side. Students also have difficulty combining information from two events like two spinners, two coins, a coin and a spinner. Students need many more opportunities to experiment with probability situations to better their understanding between models and the probability of an event happening. Learning to organize information from the situation by making charts or lists can also help them make sense of the possible outcomes.

6th grade

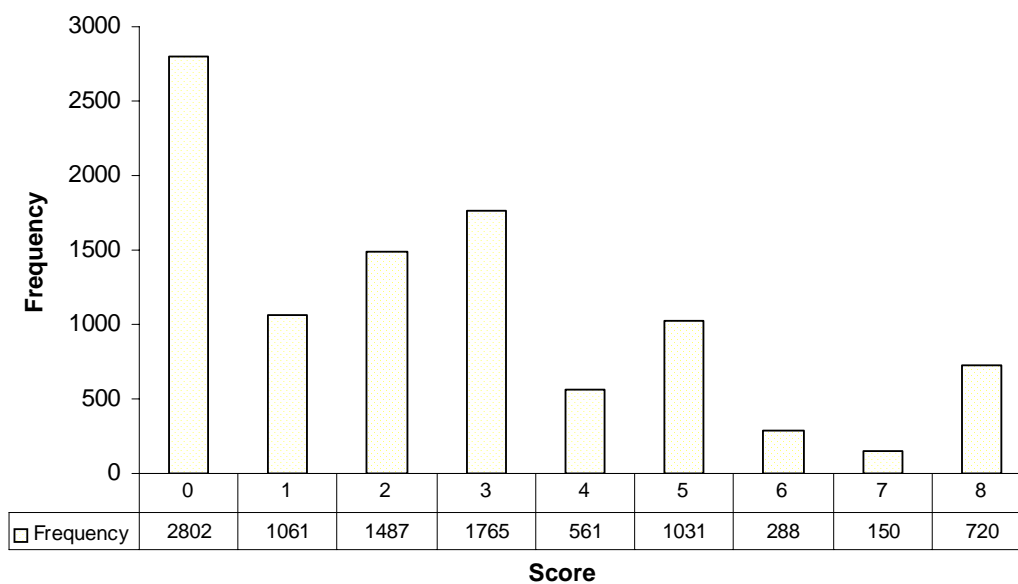
Task 5

Rabbit Costumes

Student Task	Use proportional reasoning involving fractions to find material needs for making costumes.
Core Idea 1 Number and Operations	<p>Understand number systems, the meanings of operations, and ways of representing numbers, relationships, and number systems</p> <ul style="list-style-type: none"> • Understand and use proportional reasoning to represent quantitative relationships • Select appropriate methods and tools for computing with fractions, and decimals from among mental computation, estimation, calculators, and paper-and –pencil, depending on the situation, and apply selected methods

Rabbit Costumes

Mean: 2.56, S.D.: 2.40



Score:	0	1	2	3	4	5	6	7	8
% < =	28.4%	39.2%	54.2%	72.1%	77.8%	88.3%	91.2%	92.7%	100.0%
% > =	100.0%	71.6%	60.8%	45.8%	27.9%	22.2%	11.7%	8.8%	8.8%

The maximum score available on this task is 8 points.

The cut score for a level 3 response is 3 points.

Many students (about 71%) could find the amount of blue fabric needed to make 8 costumes. Almost half the students (45%) could find the amount of fabric needed for each of the 3 colors. A little more than 10% of the students checked all 3 constraints before making a comparison. Less than 10% of the students could meet all the demands of the task. Almost 30% of the students scored no points on this task. About 1/4 of those students did not attempt the task.

Based on teacher observations, this is what sixth grade students seemed to know and be able to do:

- Find the amount of fabric needed for blue fabric (multiply by whole numbers)
- Find the amount of fabric needed for pink fabric (multiple by a fraction)

Areas of difficulty for sixth graders, sixth grade students struggled with:

- Multiplying and dividing mixed numbers
- Checking all constraints before making a comparison
- Connecting results of calculations to the context of the problem (e.g. being able to calculate how many costumes could be made from each color of fabric, but then using other information to decide which fabric would run out first)

Implications for Instruction:

Students need more experiences multiplying fractions and mixed numbers. Students also need to show all calculations before making a comparison. When given information on the amount needed for one item they need to determine the amount needed for any number of items. Students also need to work backwards. When they are given the total amount of material available, they should be able to calculate the amount of costumes that can be made. Students had difficulty grasping the difference between having a small amount of something, compared to how many costumes that would make.

Blue Print for Seventh Grade MAC

Alignment of Tasks to NCTM Content and Process Standards

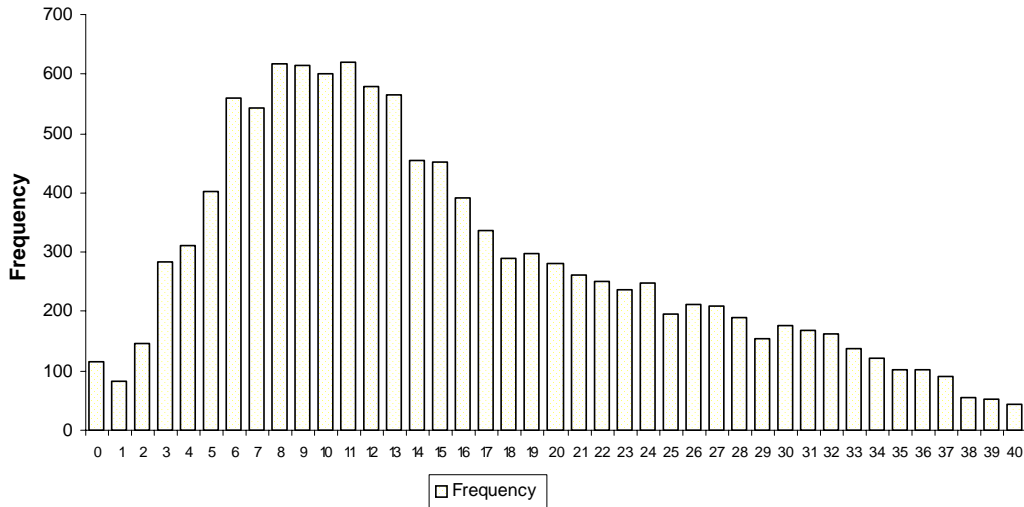
	Mixing Paints	Hexagons	Pattern	Fair Game?	Yogurt
NUM	P			S	S
ALG	S	P			S
GEO		S	P		
MEAS		S	S		P
DATA				P	
PS	*	*		*	*
REAS	*	*	*	*	*
COMM	*	*	*	*	*
CONN			*		*
REP		*	*		

NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

P denotes Primary NCTM Content Standard
 S denotes Secondary NCTM Content Standard
 • denotes NCTM Process Standard

Overall Frequency Distribution by Total Score

Grade 7
Mean=15.37; S.D.=9.22



Level Frequency Distribution Chart and Frequency Distribution

2003 - Numbers of students tested: Grade 7: 11715

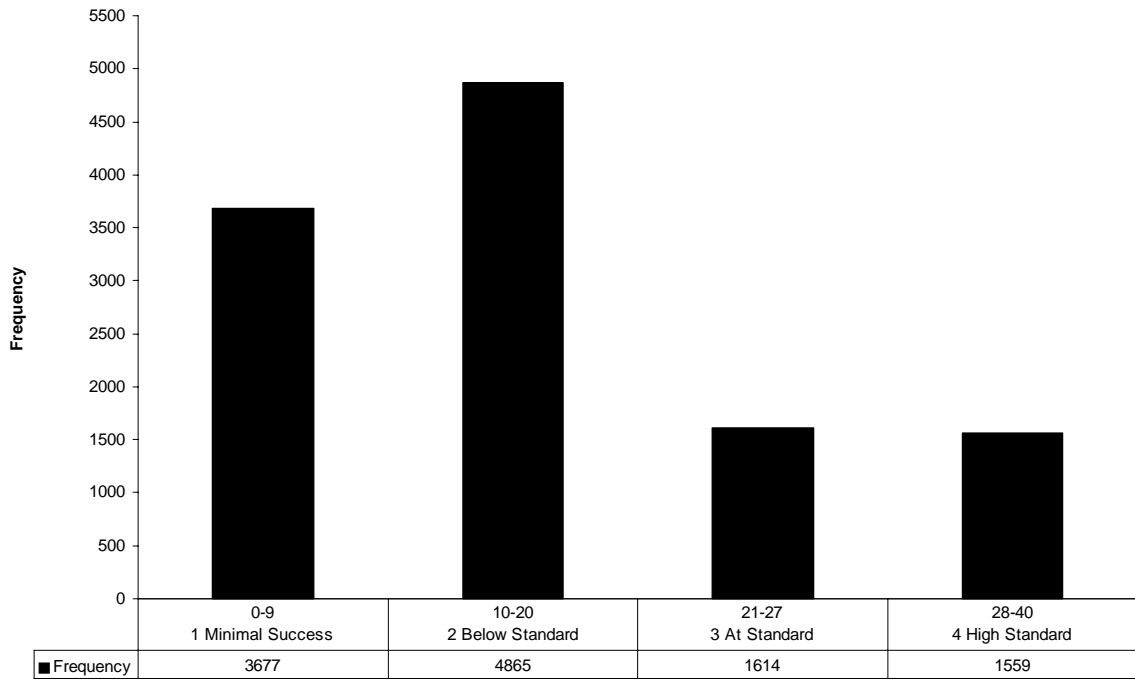
Grade 7 1999 - 2001

Level	% at ('99)	% at least ('99)	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	22%	100%	16%	100%	23%	100%
2	41%	78%	39%	74%	37%	77%
3	24%	37%	29%	45%	26%	40%
4	13%	13%	16%	16%	14%	14%

Grade 7 2002 - 2003

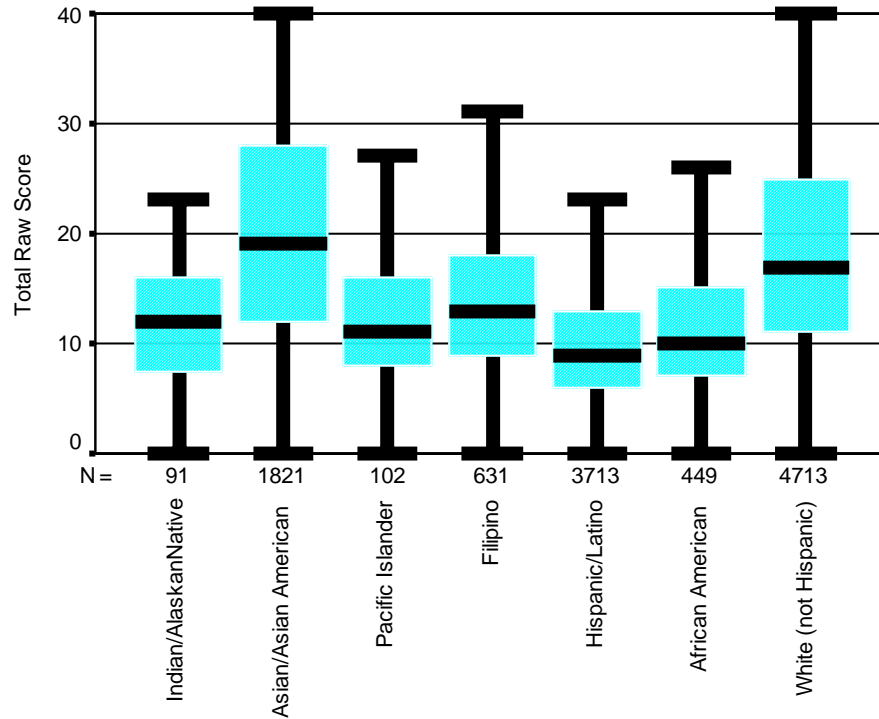
Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	31%	100%	31%	100%
2	37%	69%	42%	69%
3	24%	32%	14%	27%
4	8%	8%	13%	13%

Grade 7 – Distribution by Performance Levels



Grade 7 – Ethnicity

The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values* within a category lie between the highest and lowest horizontal bars.



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Distribution of sampling means - Grade 7 - Ethnicity

In this section, test scores are compared across different ethnic groups²¹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if the scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than group B. When the two scores overlap, then there is not significant difference between the groups.

The scores of Indian/Alaskan Native students are significantly lower than those of Asian/Asian American and White students. There are no other significant differences between the scores of the Indian/Alaskan Native students and any other ethnic group.

²¹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

The scores of Asian/Asian American students are significantly higher than those of students of all other ethnic groups.

The scores of Pacific Islander students are significantly lower than those of Asian/Asian American and White students. There are no other significant differences between the scores of Pacific Islander students and any other ethnic group.

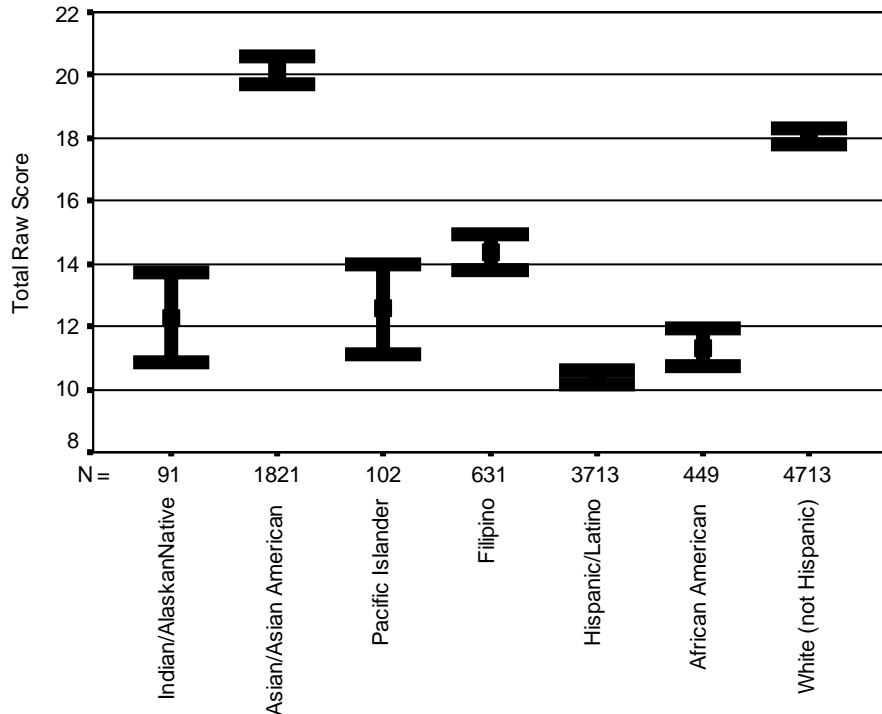
The scores of Filipino students are significantly lower than those of Asian/Asian American and White students, and significantly higher than those of Hispanic/Latino and African American students. There are no other significant differences between the scores of Filipino students and those of any other ethnic group.

The scores of Hispanic/Latino students are significantly lower than those of Asian/Asian American, Filipino, and White students. There are no other significant differences between the scores of Hispanic/Latino students and those of any other ethnic group.

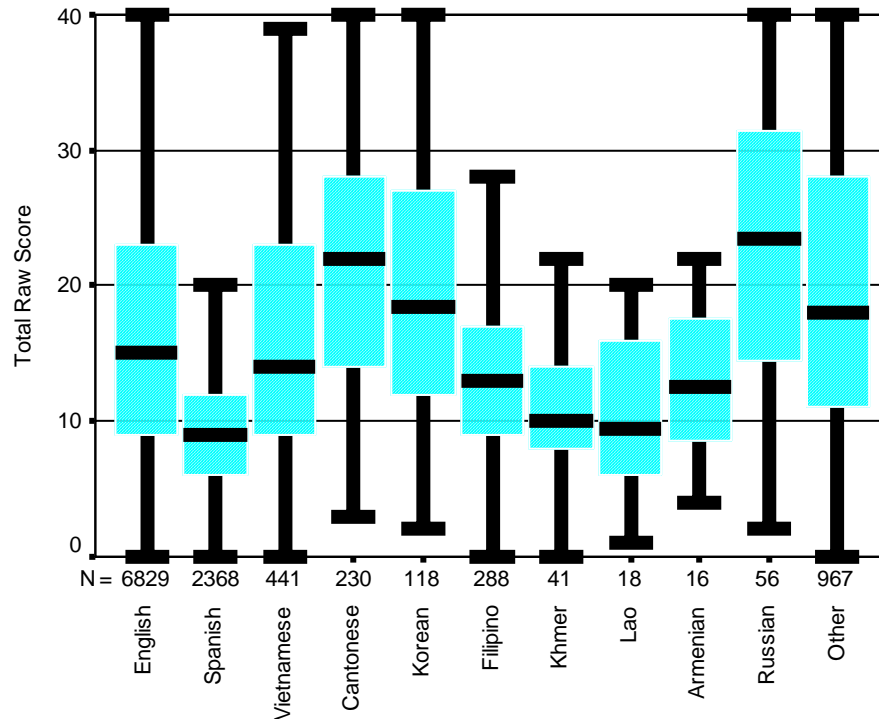
The scores of African American students are significantly lower than those of Asian/Asian American, Filipino, and White students. There are no other significant differences between the scores of African American students and any other ethnic group.

The scores of White students are significantly higher than those of Indian/Alaskan Native, Pacific Islander, Filipino, Hispanic/Latino, and African American students. The scores of White students were significantly lower than those of the Asian/Asian American students.

Distribution of sampling means - Grade 7 - Ethnicity



Grade 7 -Home Language



Distribution of sampling means - Grade 7 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home²². One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with a home language of Cantonese, Korean, Russian, or “Other”. The scores of students with English as a home language are significantly higher than those of students with a home language of Spanish, Filipino, or Khmer. There is no significant difference between scores of students with a home language of English and those with a home language of Vietnamese, Lao or Armenian.

The scores of students with Spanish as a home language are significantly lower than those of students with any other home language except Khmer, Lao and Armenian. The

²² Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

differences between the scores of students with Spanish as a home language and the scores of students in these latter three groups are not significant.

The scores of students with Vietnamese as a home language are significantly higher than those of students with Spanish or Filipino as a home language, and lower than those of students with Cantonese, Korean, Russian, or “Other” as a home language. There are no other significant differences between students with Vietnamese as a home language and any other group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, Khmer, Lao, or “Other” as a home language. There are no significant differences between the scores of students with Cantonese as a home language and any other group.

The scores of students with Korean as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, Khmer, or Lao as a home language. There are no other significant differences between the scores of students with Korean as a home language and any other group.

The scores of students with Filipino as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a home language, and significantly higher than those of students with Spanish as a home language. There are no other significant differences between the scores of students with Filipino as a home language and any other group.

The scores of students with Khmer as a home language are significantly lower than those of students with English, Cantonese, Korean, Russian, or “Other” as a home language. There are no other significant differences between the scores of students with Khmer as a home language and any other group.

The scores of students with Lao as a home language are significantly lower than those of students with Cantonese, Korean, Russian, or “Other” as a home language. There are no other significant differences with any other home language category.

The scores of students with Armenian as a home language are significantly lower than the scores of students with Russian as their home language. There are no other significant differences with any other language group.

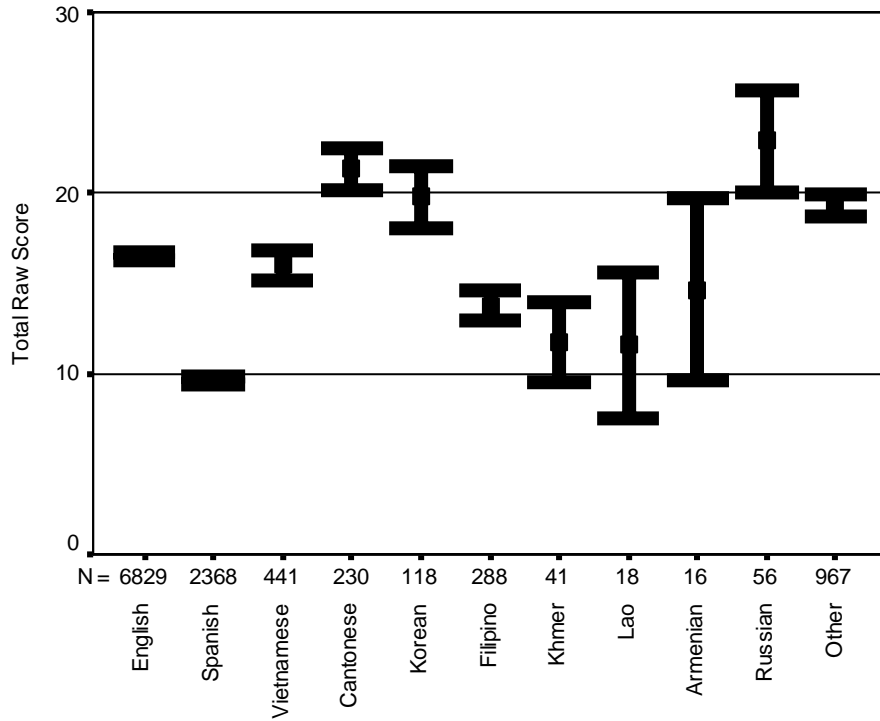
The scores of students with Russian as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, Khmer, Lao, or Armenian as a home language. There are no other significant differences between the scores of students with Russian as a home language and any other language group.

The scores of students with “Other” as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, Khmer, or Lao as a home language, and significantly lower than those of students with Cantonese as a home

language. There are no other significant differences between the students with “Other” as a home language and any other language group.

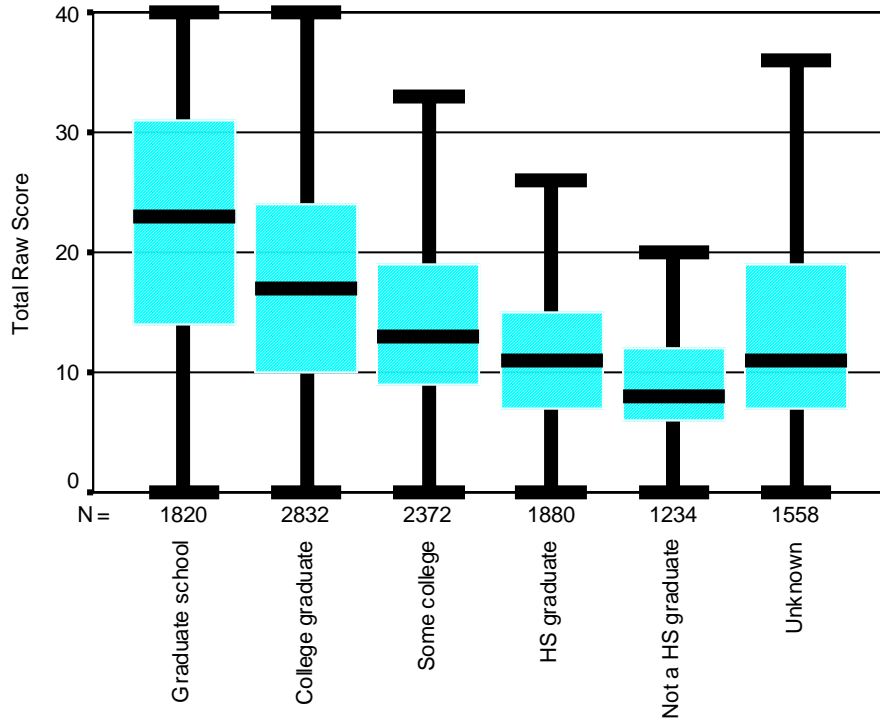
Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 7 - Home Language



Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmu, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Grade 7 - Parent Education



Distribution of sampling means - Grade 7 - Parent Education

In this section, test scores are compared across groups of different levels of parent education²³. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate school education are significantly higher than those of students in all other Parent Education categories.

The scores of students whose parents have a college education are significantly lower than those whose parents have a graduate education, and significantly higher than all other groups.

The scores of students whose parents have some college education are significantly higher than those of students whose parents are not High School graduates and those whose parents are High School graduates, and significantly lower than those whose parents have a college education or a graduate education. There is not a significant

²³ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

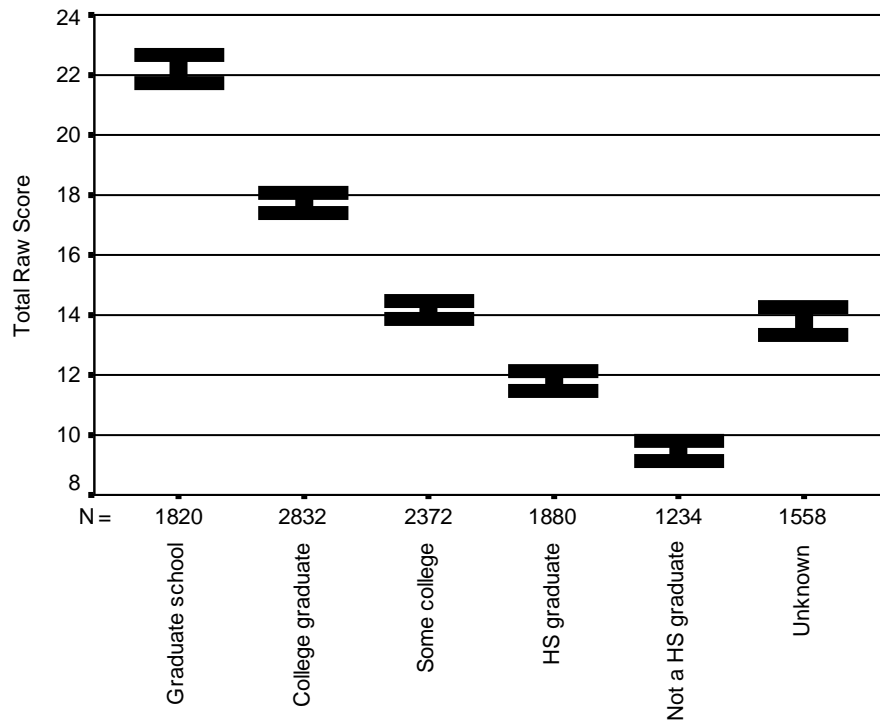
difference between scores of students whose parent education level is unknown and those whose parents have some college education.

The scores of students whose parents are High School graduates are significantly higher than those of students whose parents are not High School graduates, and significantly lower than the scores of students in all other Parent Education categories.

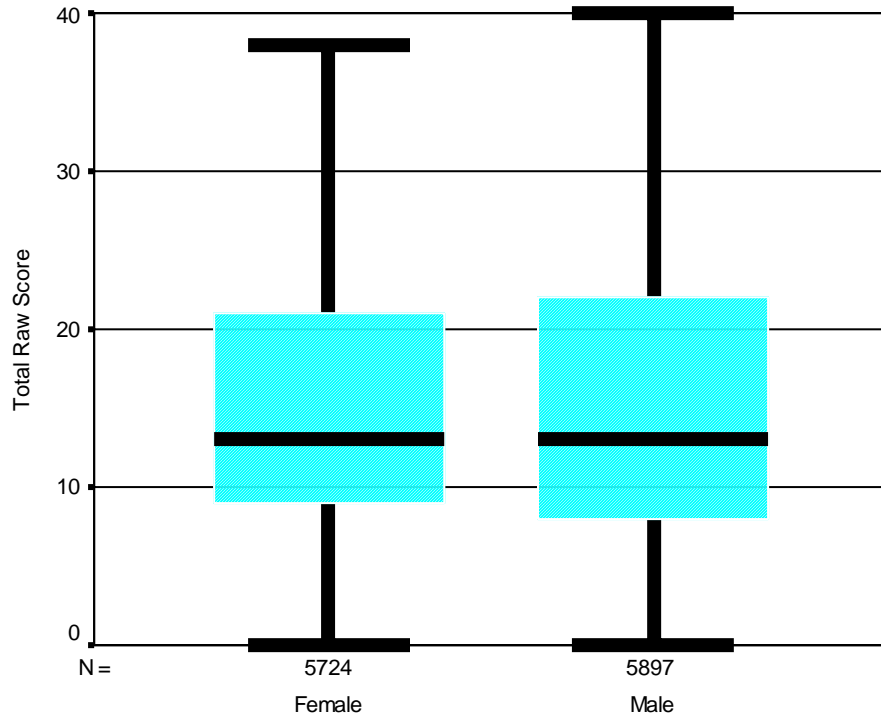
The scores of students whose parents are not High School graduates are significantly lower than the scores of students in all other Parent Education categories.

The scores of students whose parents' education is unknown are significantly higher than those of students whose parents are not High School graduates and those whose parents are High School graduates. The scores of students whose parents' education is unknown are significantly lower than those of students whose parents have a college degree or have a graduate education. There is not a significant difference between scores of students whose parent education level is unknown and those whose parents have some college education.

Distribution of sampling means - Grade 7 - Parent Education



Grade 7 - Gender



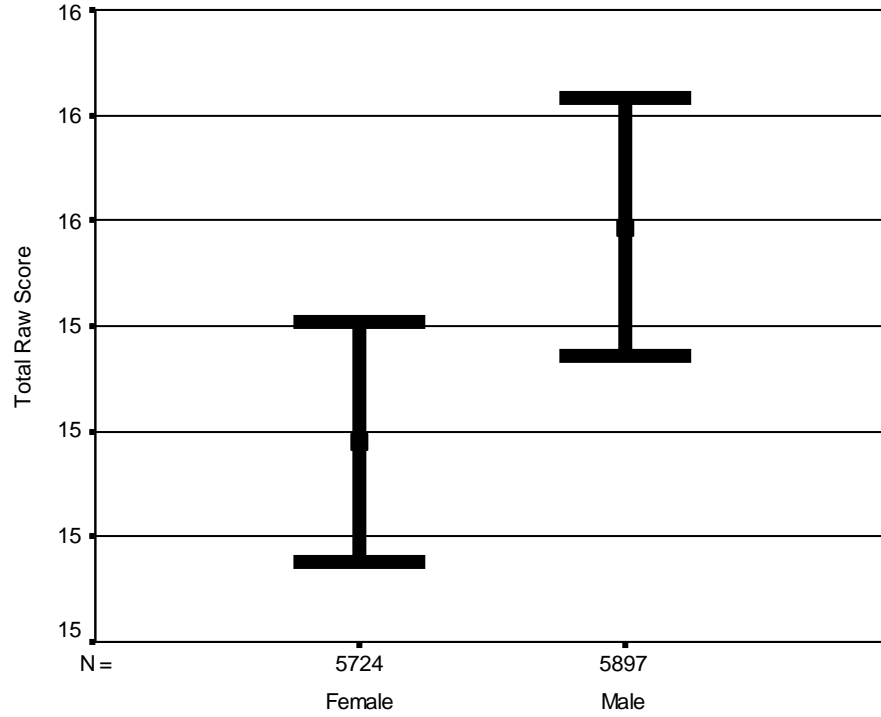
Distribution of sampling means - Grade 7 - Gender

In this section, test scores are compared across gender²⁴. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

There is no statistically significant difference in the scores of females and males.

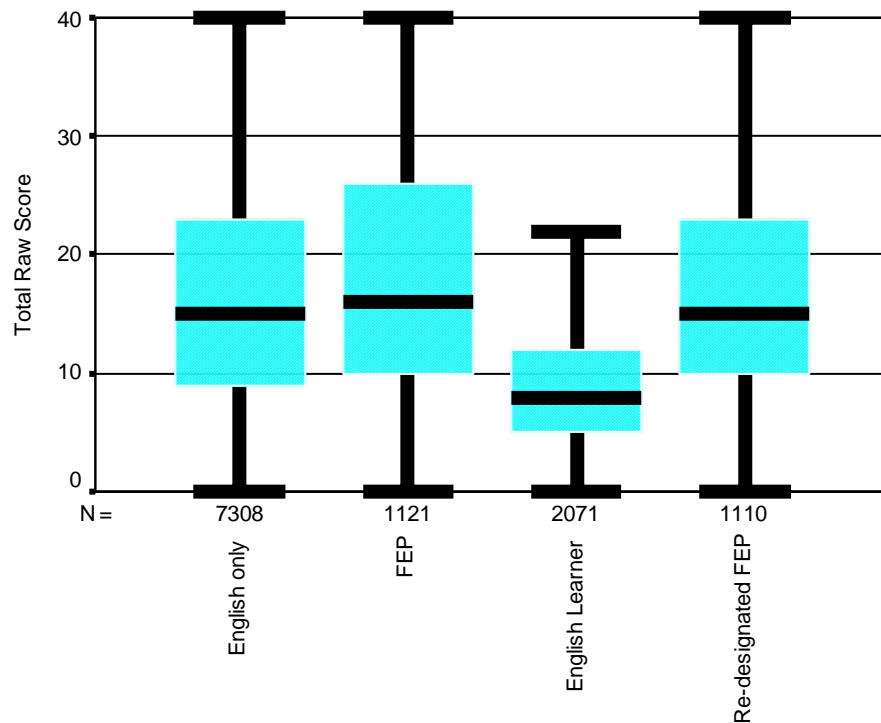
²⁴ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 7 - Gender



¹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 7 - Language Fluency



Distribution of sampling means - Grade 7 - Language Fluency

In this section, test scores are compared across groups of different language fluency²⁵. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly lower than those of students described as Full English Proficiency (FEP) and significantly higher than those students described as English Learners. There is no significant difference between scores of students described as English Only, and those described as re-designated FEP.

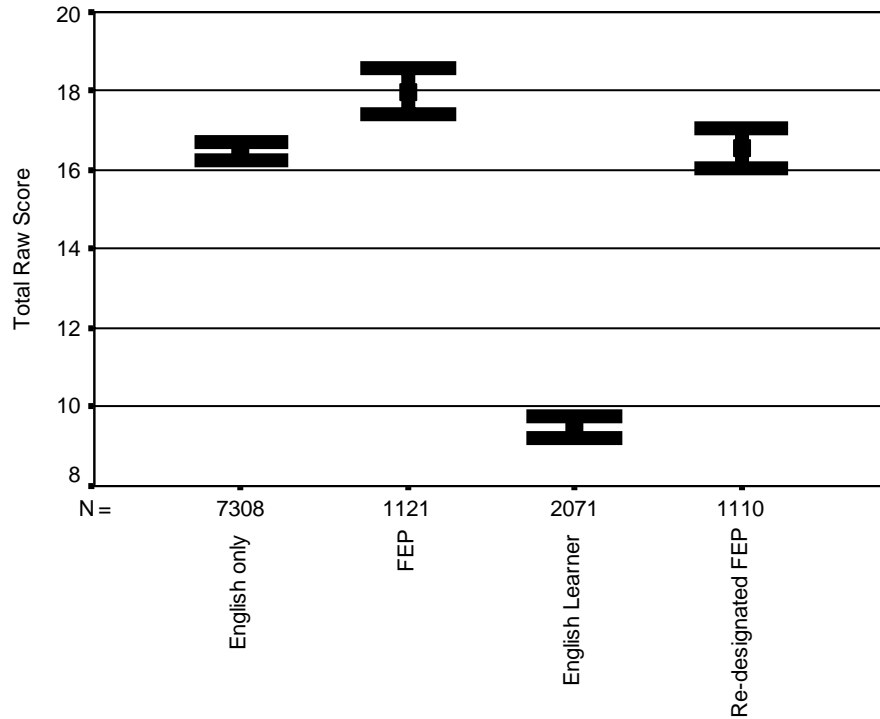
The scores of students in the FEP category are significantly higher than those of students in any other language fluency category.

The scores of students in the English Learner category are significantly lower than those of the students in all other category of language fluency.

The scores of students in the Re-designated FEP category are significantly higher than those of students in the English Learner category, and lower than those of students in the FEP category. There is no significant difference between scores of students described as re-designated FEP, and those described as English Only.

²⁵ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 7 - Language Fluency



7th grade

Task 1

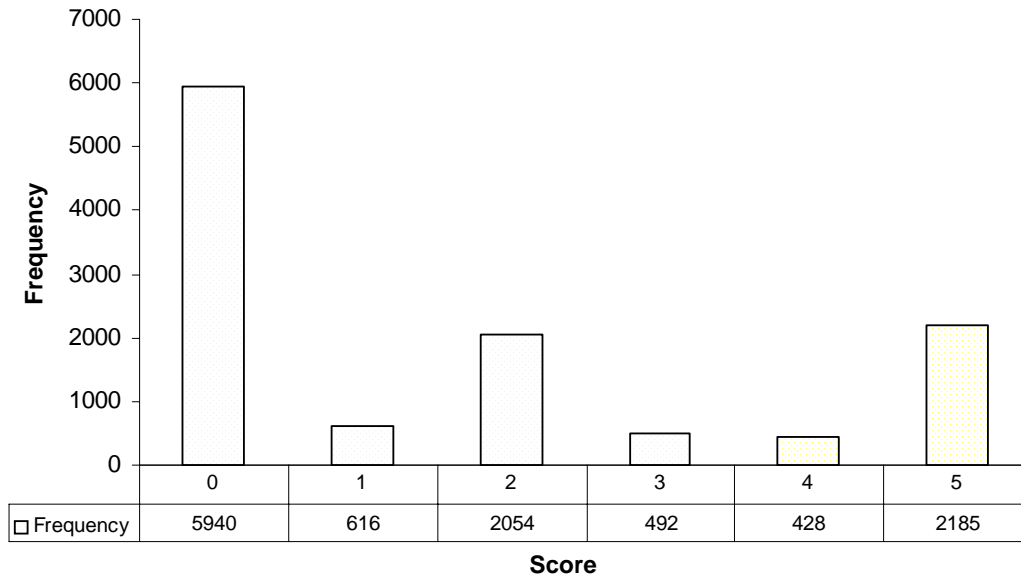
Mixing Paint

Student Task	Use ratios and percents to solve a practical problem involving the mixing of paint.
Core Idea 1 Number and Operations	<p>Understand number systems and ways of representing number, relationships, and number systems.</p> <ul style="list-style-type: none"> • Develop, analyze and explain methods for solving problems involving proportional reasoning, such as scaling and finding equivalent ratios • Understand the meaning and effects of operations with rational numbers. • Develop and use strategies to estimate the results of rational number computations, and judge the reasonableness of results. • Work flexibly with fractions, decimals, and percents to solve problems.

Frequency Distribution for each Task – Grade 7 – Mixing Paints

Mixing Paints

Mean: 1.61, S.D.: 1.95



Score:	0	1	2	3	4	5
% < =	50.7%	56.0%	73.5%	77.7%	81.3%	100.0%
% > =	100.0%	49.3%	44.0%	26.5%	22.3%	18.7%

The maximum score available on this task is 5 points.

The cut score for a level 3 response is 2 points.

A little less than half the students could find the amount of red or blue paint needed. Less than 20% could find the amount of red and blue paint and calculate the percentage of total paint that represented. A little more than 50% of the students scored zero points on this task. About 7% of all the students did not attempt this task, even though it was the first problem in the test. 43% of the students attempted the task with no success.

Based on teacher observations, this is what seventh grade students knew and were able to do:

- Find the fractional amount of red or blue paint and use that to find the number of quarts

Areas of difficulty for seventh graders, seventh grade students struggled with:

- Understanding the whole in a multi-step problem
- Calculating percentages
- Tracking part/whole relationships

- Using fractions instead of taking a fractional part of the whole

Instructional Implications:

Students need strategies to help them make sense of problem situations. Being able to draw a picture or make a model helps them see part/whole relationships. Until the student can see how the parts fit together, the student cannot calculate with fractions or find percentages for the different components. Students need more experiences working with problems in which “the whole” changes. Students need more experience using information from problems to get the answer instead of repeating numerical information that has been given in the problem. Too many students thought the numbers in the text of the problem were the solutions. Students at this grade level should be able to convert mentally between common fractions and percents. At this grade level, students should “just know” the relationship between common fractions and the equivalent percents.

7th grade

Task 2

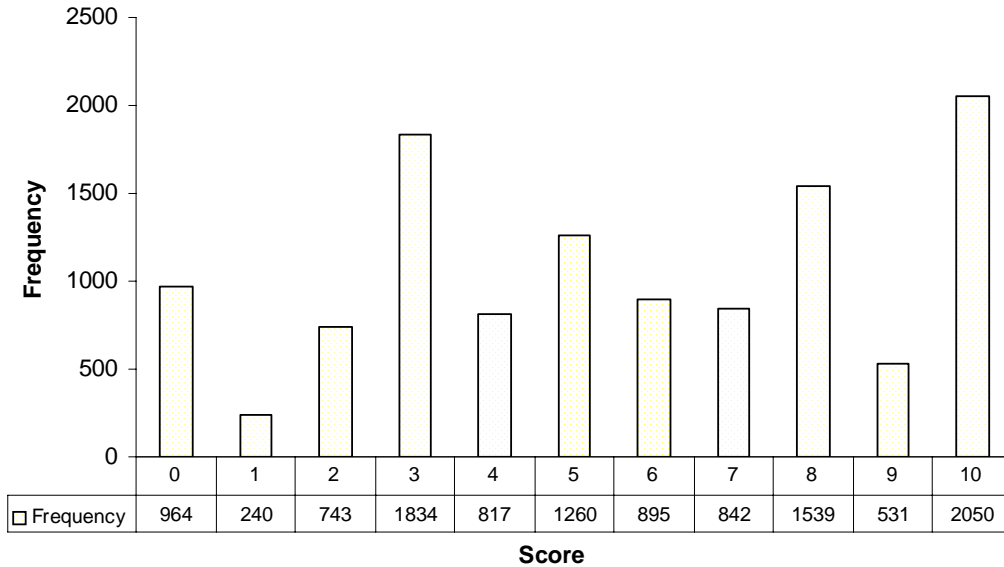
Hexagons

Student Task	Find and extend a number pattern in the context of hexagonal designs. Give a rule or formula for determining the perimeter of each growing design.
Core Idea 2 Algebra and Functions	<p>Understand relations and functions, analyze mathematical situations and use models to solve problems involving quantity and change.</p> <ul style="list-style-type: none"> • Represent, analyze, and generalize a variety of functions including linear relationships • Express mathematical relationships using expressions and equations • Use symbolic algebra to represent situations to solve problems

Grade 7 – Hexagons

Hexagons

Mean: 5.60, S.D.: 3.15



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	8.2%	10.3%	16.6%	32.3%	39.2%	50.0%	57.6%	64.8%	78.0%	82.5%	100.0%
% > =	100.0%	91.8%	89.7%	83.4%	67.7%	60.8%	50.0%	42.4%	35.2%	42.4%	17.5%

The maximum score available on this task is 10 points.
The cut score for a level 3 response is 6 points.

Most students (about 92%) could look at a geometric pattern and continue it on a table. About half the students could extend the pattern in a table and slightly beyond the numbers in the table and find the perimeter for a row of 66 tiles.(score of 6 or above). Almost 18% of the students could meet all the demands of the task including writing a rule for extending the pattern in words or algebraic notation. 8% of the students scored no points on this task.

Based on teacher observations, this is what seventh graders seemed to know and be able to do:

- Find a geometric pattern and extend it using a table.
- Extend a geometric pattern beyond the table using drawing or repeated addition.

Areas of difficulties for seventh graders, seventh grade students struggled with:

- Bridging from drawing, extending tables, and repeated addition to using multiplicative relationships
- Seeing the connections between independent and dependent variables to write rules or formulas, instead of writing recursive relationships
- Understanding the role of constants in functions to know that unless the relationship is a proportion different numbers in the table cannot be added to find further numbers in the pattern without considering the change in the constant
- Using inverse relationships or algebra to work from the perimeter (dependent variable) to the number of tiles (independent variable)

Implications for Instruction:

When students look at pattern problems, it is helpful to visualize what is changing and what is staying the same. As they progress through the grades this information could be used to help them write a rule or formula. At this grade level, they should no longer rely on drawing pictures or doing repeated addition to find the solutions to complex problems. They should develop the ability to see and use multiplicative relationships. Students also need experience working backward with patterns to knowing the dependent variable and using it to find the independent variable. Students at this grade level should be proficient at solving a variety of questions about growing patterns.

7th grade

Task 3

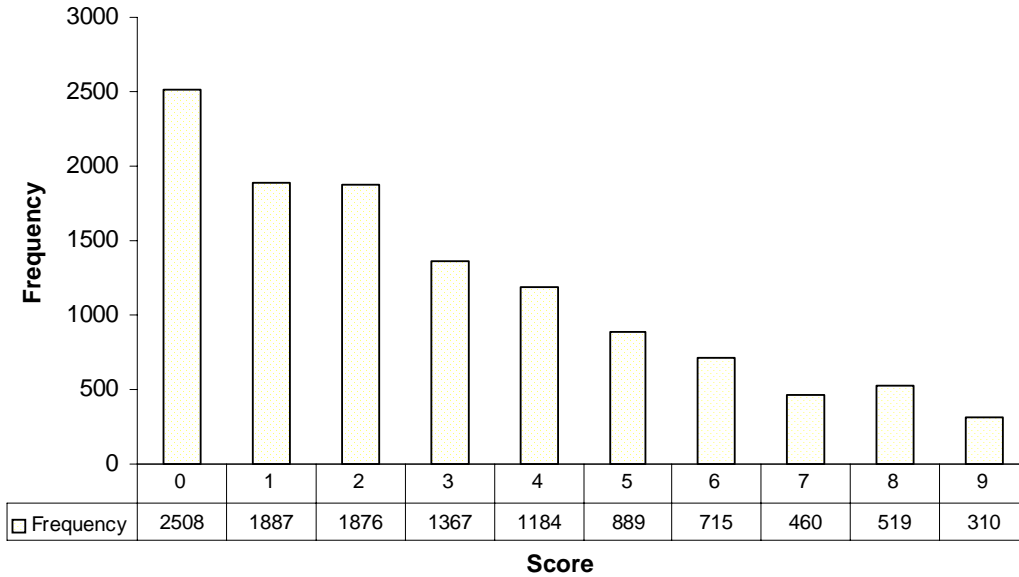
Pattern

Student Task	Determine symmetrical properties of shape. Calculate lengths and angles in symmetrical figures.
Core Idea 3 Geometry and Measurement	<p>Apply congruence and similarity to analyze mathematical situations. Apply appropriate techniques, tools, and formulas to determine measurements.</p> <ul style="list-style-type: none"> • Understand relationships among the angles, side lengths, perimeters, and the areas of similar objects • Develop and critique inductive and deductive arguments concerning geometric ideas and relationships. • Understand line of symmetry. (6th grade) • Investigate and reason about the results of subdividing, combining, and transforming shapes. (6th grade)

Grade 7 – Pattern

Pattern

Mean, 2.85, S.D.: 2.52



Score:	0	1	2	3	4	5	6	7	8	9
% < =	21.4%	37.5%	53.5%	65.2%	75.3%	82.9%	89.0%	92.9%	97.4%	100.0%
% > =	100.0%	78.6%	62.5%	46.5%	34.8%	24.7%	17.1%	11.0%	7.1%	2.6%

The maximum score available for this task is 9 points.

The cut score for a level 3 response is 4 points.

Many students (about 80%) could find the lines of symmetry for the rhombus or for the total design. More than half (about 63%) the students could find the lines of symmetry for the rhombus and either the degrees in an equilateral triangle or the lines of symmetry in an equilateral triangle. About one fourth of the students showed understanding of symmetry, perimeter, and some knowledge about angle, usually finding the degrees in the acute angle of the rhombus. Less than 3 % of the students could meet all the demands of the task. 21% of the students scored no points on the task. 82% of the students with a score of zero attempted the task.

Based on teacher observations, this is what seventh grade students seemed to know and be able to do:

- Find lines of symmetry in a rhombus
- Find at least one line of symmetry in an equilateral triangle
- Find the perimeter of the total design and calculate accurately with decimals

Areas of difficulty for seventh graders, seventh grade students struggled with:

- Knowing basic facts about angles, like degrees in an equilateral triangle or in a circle
- Interpreting a diagram and using know facts to find the degrees of other angles
- Finding lines of symmetry for a complex shape

Instructional Implications:

Students need more opportunities to analyze compound geometric figures. Students need a variety of experiences such as folding papers, building designs with blocks, using mirrors to illustrate line of symmetry. Students had difficulty identifying which part of a complex design related to the perimeter of the shape. Some students tried to find the perimeter of all the individual pieces, instead of thinking about the shape as a whole unit. Students also struggled with angle measurement. They could not apply information about the number of degrees in a triangle or number of degrees in a circle to a problem-solving situation. Students at this grade level should be comfortable working with much more complex ideas about angles and angle properties, like understanding vertical angles and alternating angles. Even student guesses showed limited knowledge of basic angle measurement, such as previous experience with protractors. Even if answers were considered estimates, they were often too far off to be reasonable.

7th grade

Task 4

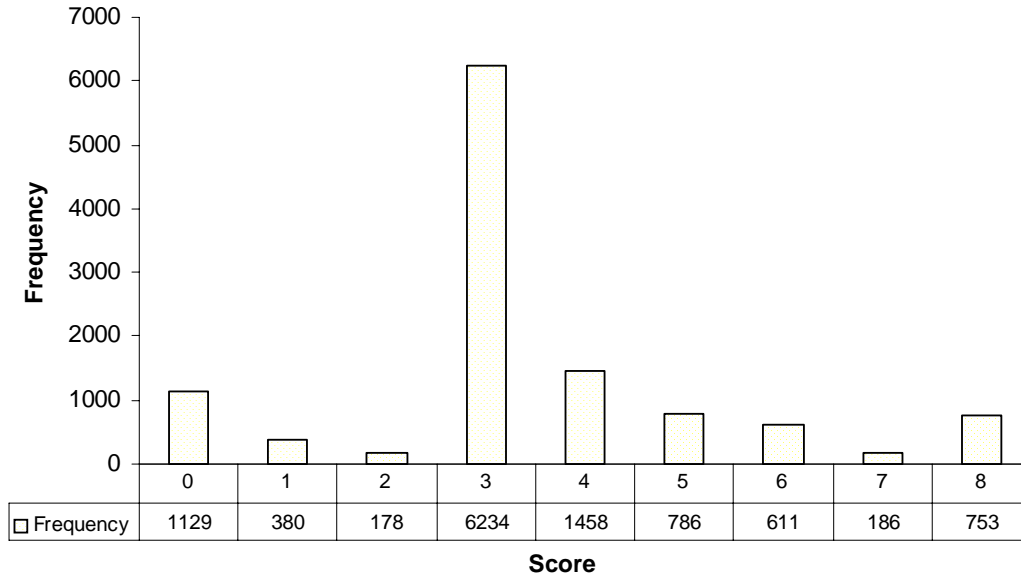
Fair Game?

Student Task	Use probability to judge the fairness of a game involving a dice and coin toss.
Core Idea 2 Probability	<p>Apply and deepen the understanding of theoretical and empirical probability.</p> <ul style="list-style-type: none"> • Represent the sample space for simple and compound events in an organized way (table, diagram, organized list, and tree diagram) and express the theoretical probability of each outcome. • Use data to estimate the probability of future events. • Know the relationship between the probability of an event and its complement (i.e., if the probability of an event is P, then the probability of the event not occurring is $1-P$). • Determine theoretical probabilities and use these to make predictions about events.

Grade 7 – Fair Game?

Fair Game?

Mean, 3.43, S.D.: 1.88



Score:	0	1	2	3	4	5	6	7	8
% < =	9.6%	12.9%	14.4%	67.6%	80.1%	86.8%	92.0%	93.6%	100.0%
% > =	100.0%	90.4%	87.1%	85.6%	32.4%	19.9%	19.9%	8.0%	6.4%

The maximum score available for this task is 8 points.

The cut score for a level 3 response is 4 points.

Many students (about 85%) could apply the two different rules to fill out the table of outcomes. Less than 20% could find the number of primes and correctly calculate the probability for getting a prime. Less than 10% of the students could meet all the demands of the task. 10% of the students scored no points on this task. 5% of the students appeared to have run out of time, because there is no work on task 4 or task 5.

Based on teacher observations, this is what seventh grade students seemed to know and be able to do:

- Follow rules to fill out a table

Areas of difficulty for seventh graders, seventh grade students struggled with:

- Identifying 2 as a prime number
- Calculating probabilities from a given set of outcomes or recognizing the difference between the random number generators and the outcomes (*many students thought things were 50-50 because there are two sides on a coin or three of the 6 numbers on the cube are prime*)
- Using probability to find a expected outcomes in a game situation. Giving a good mathematical argument for why a game is fair combining probabilities and rules of the game

Instructional Implications:

Students need more experiences with games and probability. Students do not know how to combine two events, like tossing a coin and a number cube, to find the probability. They confuse the random number generators with the outcomes. While students can complete tables showing what is possible, they have difficulty determining how to answer questions based on that table. When it is possible to move a different number of spaces for each player, they don't know how to use that information to determine who has a better chance of winning. Students do not understand that probability or expected outcomes are determined by a large number of events. To compare whether a game is fair, the outcomes of several games needs to be considered. Students also needed to know mathematical vocabulary like prime number, fair, and trial to work this problem successfully.

7th grade

Task 5

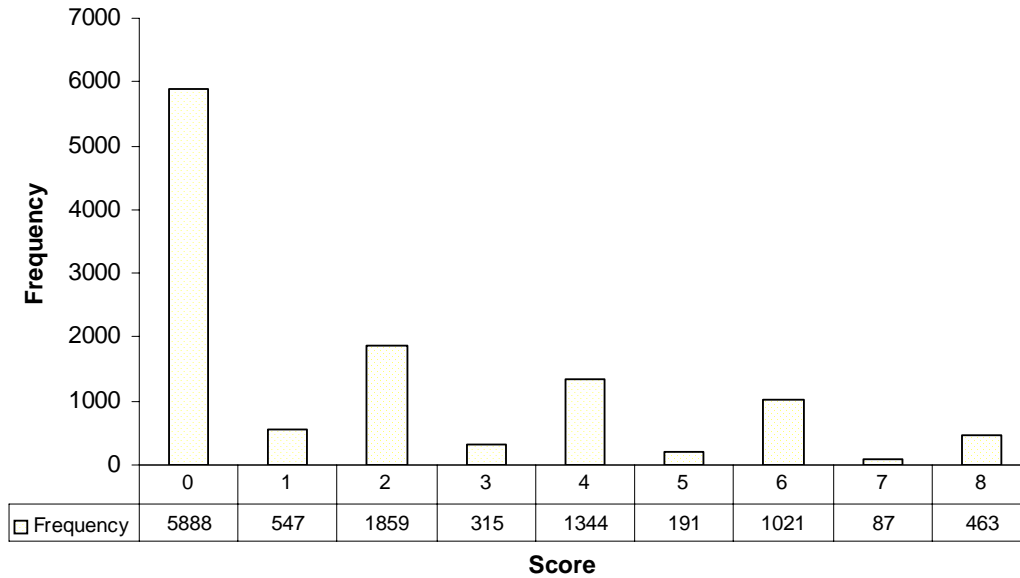
Yogurt

Student Task	Use fractions and percents to calculate the answers to questions involving the production and profit of yogurt packaging.
Core Idea 1 Number and Operation	<p>Understand number systems, the meanings of operations, and ways of representing numbers, relationships, and number systems.</p> <ul style="list-style-type: none"> • Develop, analyze, and explain methods for solving problems involving proportional reasoning • Understand the meaning and effects of operations with rational numbers • Use the associative and commutative properties of multiplication and the distributive property of multiplication to simplify computations • Work flexibly with fractions, decimals, and percents to solve problems • Solve problems involving proportional reasoning and scaling.

Grade 7 – Yogurt

Yogurt

Mean: 1.88, S.D.: 2.38



Score:	0	1	2	3	4	5	6	7	8
% < =	50.3%	54.9%	70.8%	73.5%	85.0%	86.6%	95.3%	96.0%	100.0%
% > =	100.0%	49.7%	45.1%	29.2%	26.5%	15.0%	15.0%	4.7%	4.0%

The maximum score available for this task is 7 points.
The cut score for a level 3 response is 4 points.

About half the students could find 20% of a number, although there might have been errors in monetary notation. Some students used a percent sign for the answer instead of cents. Others combined decimal points and cent signs. About 30% of the students could correctly find 20% of a number using correct monetary notation and find the correct relationship between gallons per hour and gallons for 5 days or could correctly convert from 1/2-cup containers to gallons per hour to gallons per week. Less than 5% of the students could meet all the demands of the task. Almost 50% of the students scored no points on this task. While time may have been an issue for some students with this score, almost 80% of the students with a score of zero attempted the task.

Based on teacher observations, this is what seventh grade students seemed to be able to do:

- Convert from gallons per hour to gallons per week
- Find 20% profit and use correct monetary notation

Areas of difficulty for seventh graders, seventh grade students struggled with:

- Converting among different units of measurement
- Choosing the appropriate operation for converting between units of measurement
- Calculating the amount of increase before attempting to calculate the percent of increase
- Using the final amount to calculate the percent of increase instead of using the starting amount as the base
- Knowing how to calculate a percentage

Instructional Implications:

Students need more practice converting between different measurements. Many students are unfamiliar with cups, pints, quarts, and gallons. Students also have trouble working with detailed calculations. They need a strategy like labeling, to help them keep track of what their answer means, what quantity they have calculated, and what is still left to do. They also need more work around number operations. Many students multiplied instead of dividing to make their conversions. Students also need to develop a conceptual understanding of percents in a variety of contexts. They should have a variety of strategies for thinking and reasoning about percents. Further students do not show an understanding of looking at increase and comparing it to the base amount. They need to develop an understanding of percent as a tool for comparing quantities and understanding change, not as just parts of 100 or an algorithm. They need to do more reasoning on the use and purpose for percents.

Blue Print for Eighth Grade MAC

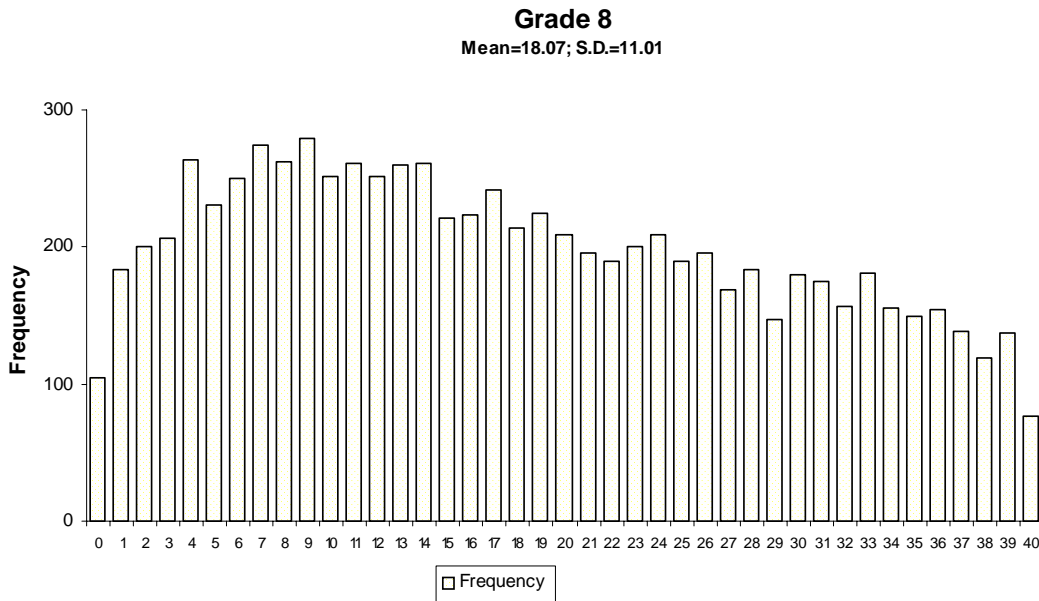
Alignment of Tasks to NCTM Content and Process Standards

	Pete's Number	Squares & Rectangles	Sports Injuries	Dots & Squares	Number Pairs
NUM	S		P		
ALG	P			P	S
GEO		P		S	P
MEAS		S	S	S	
DATA			P		S
PS	*		*		
REAS		*		*	*
COMM		*			
CONN	*	*	*		
REP	*			*	*

NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

P denotes Primary NCTM Content Standard
 S denotes Secondary NCTM Content Standard
 • denotes NCTM Process Standard

Overall Frequency Distribution by Total Score



Level Frequency Distribution Chart and Frequency Distribution

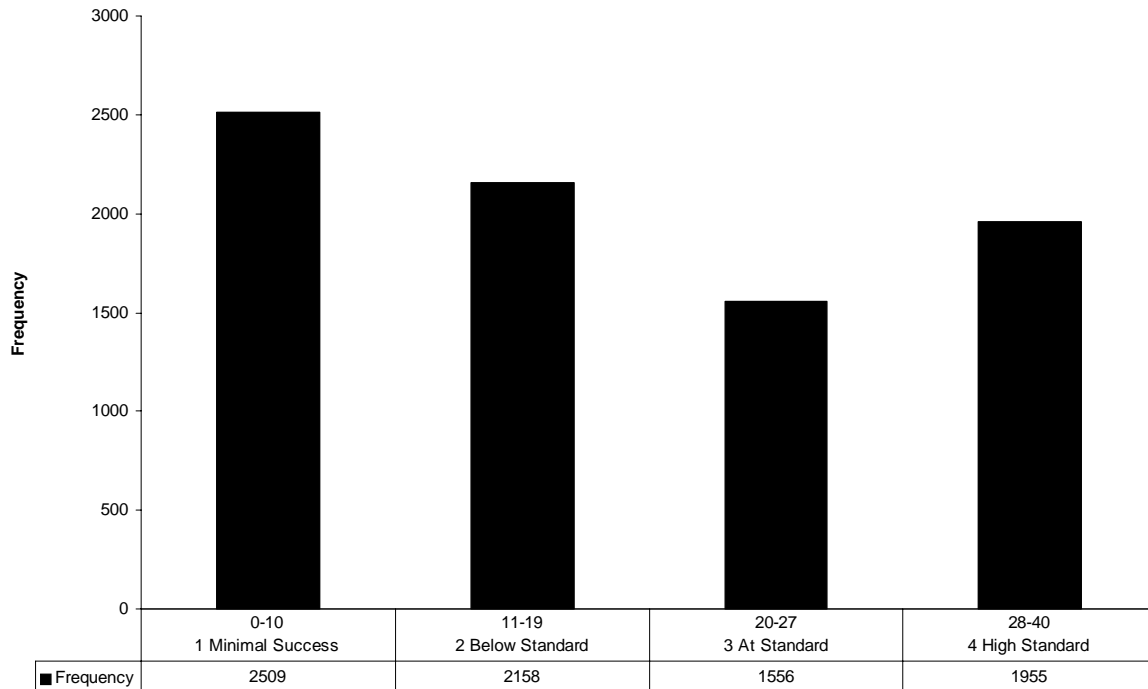
2003 - Numbers of students Grade 8: 8178
tested:

Grade 8 2000 - 2001

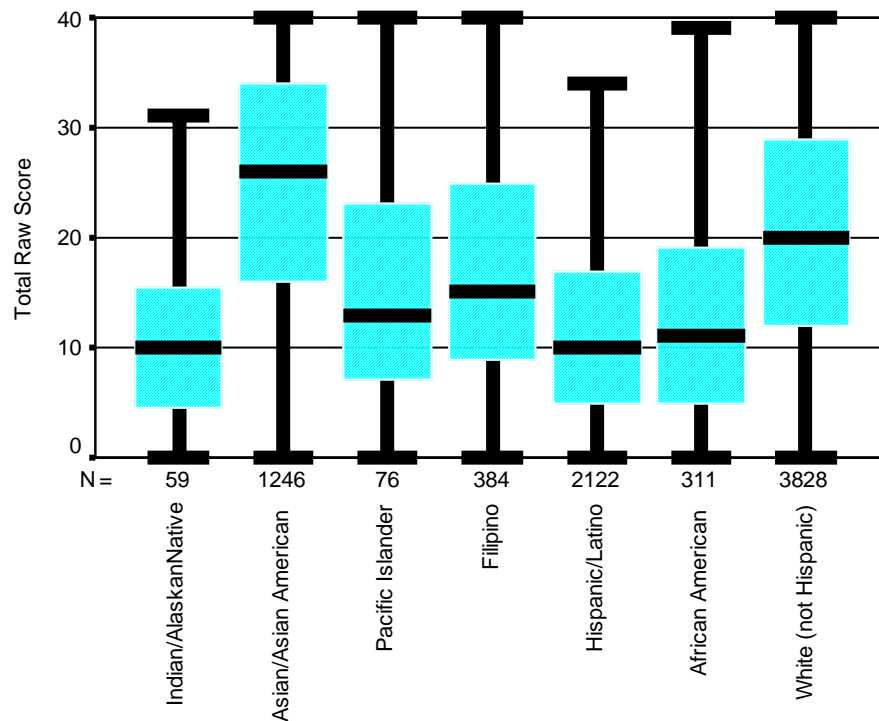
Level	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	19%	100%	19%	100%
2	47%	81%	42%	81%
3	26%	34%	31%	39%
4	8%	8%	8%	8%

Grade 8 2002 - 2003

Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	27%	100%	31%	100%
2	37%	73%	26%	69%
3	23%	36%	19%	43%
4	13%	13%	24%	24%



Grade 8 – Ethnicity - The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values* within a category lie between the highest and lowest horizontal bars.



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Distribution of sampling means - Grade 8 - Ethnicity

In this section, test scores are compared across different ethnic groups²⁶. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B, then group A is significantly higher than group B. Conversely, if the scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than group B. When the two scores overlap, then there is not significant difference between the groups.

The scores of Indian/Alaskan Native students are significantly lower than those of Asian/Asian American, Filipino or White students. There are no other significant differences between the scores of the Indian/Alaskan Native students and any other ethnic group.

The scores of Asian/Asian American students are significantly higher than those of students of all other ethnic groups.

The scores of Pacific Islander students are significantly lower than those of Asian/Asian American and White students, and higher than those of Hispanic/Latino students. There are no other significant differences between the scores of Pacific Islander students and those of any other ethnic group.

The scores of Filipino students are significantly lower than those of Asian/Asian American and White students, and significantly higher than those of Indian/Alaskan Native, Hispanic/Latino and African American students. There are no other significant differences between the scores of Filipino students and those of any other ethnic group.

The scores of Hispanic/Latino students are significantly lower than those of Asian/Asian American, Pacific Islander, Filipino, and White students. There are no other significant differences between the scores of Hispanic/Latino students and those of any other ethnic group.

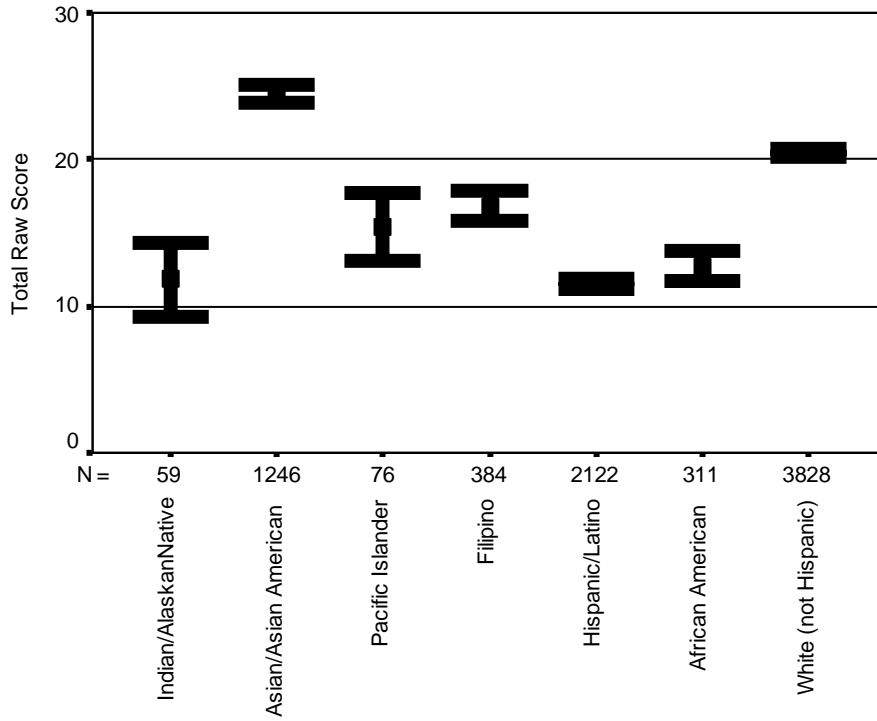
The scores of African American students are significantly lower than those of Asian/Asian American, Filipino, and White students. There are no other significant differences between the scores of African American students and those of any other ethnic group.

The scores of White students are significantly higher than those of Indian/Alaskan Native, Pacific Islander, Filipino, Hispanic/Latino, and African American students. The

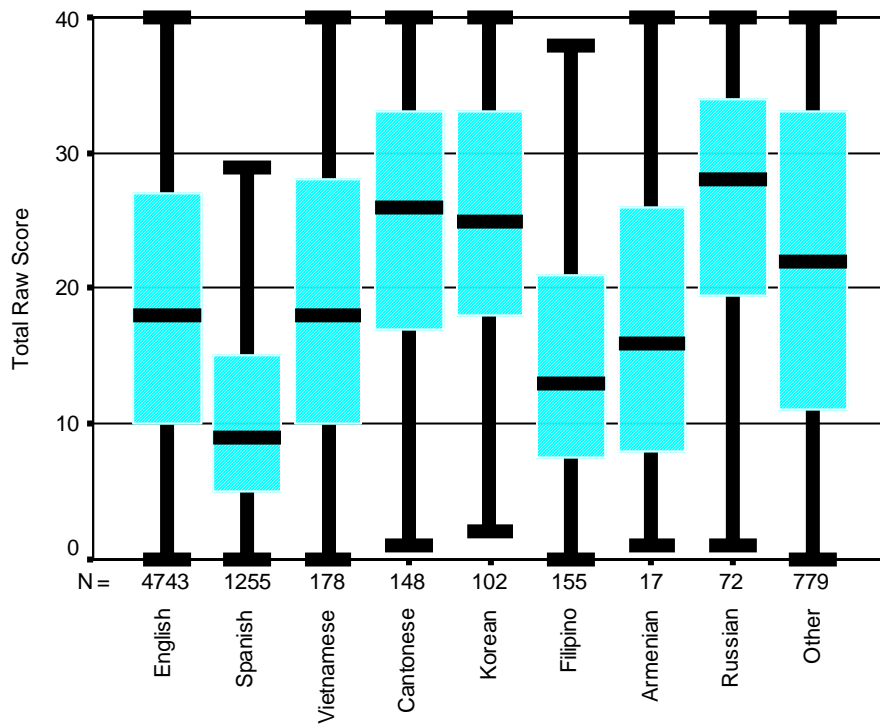
²⁶ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

scores of White students were significantly lower than those of the Asian/Asian American students.

Distribution of sampling means - Grade 8 - Ethnicity



Grade 8 - Home Language



Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Lahu, Lao, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 8 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home²⁷. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with a home language of Cantonese, Korean, Russian, or “Other”. The scores of students with English as a home language are significantly higher than those of students with a home language of Spanish or Filipino. There is no significant difference between scores of students with a home language of English, and those with Vietnamese or Armenian as a home language.

The scores of students with Spanish as a home language are significantly lower than those of students with any other home language.

The scores of students with Vietnamese as a home language are significantly higher than those of students with Spanish or Filipino as a home language, and lower than those of students with Cantonese, Korean, or Russian as a home language. There are no other significant differences between the scores of students with Vietnamese as a home language and those of any other language group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or “Other” as a home language. There are no other significant differences between the scores of students with Cantonese as a home language and those with Korean, Armenian or Russian as a home language.

The scores of students with Korean as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or “Other” as a home language. There are no other significant differences between the scores of students with Korean as a home language and any other group.

The scores of students with Filipino as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, Korean, Russian, or “Other” as a home language, and significantly higher than those of students with Spanish as a home

²⁷ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

language. There are no significant differences between the scores of students with Filipino as a home language and those of students with Armenian as a home language.

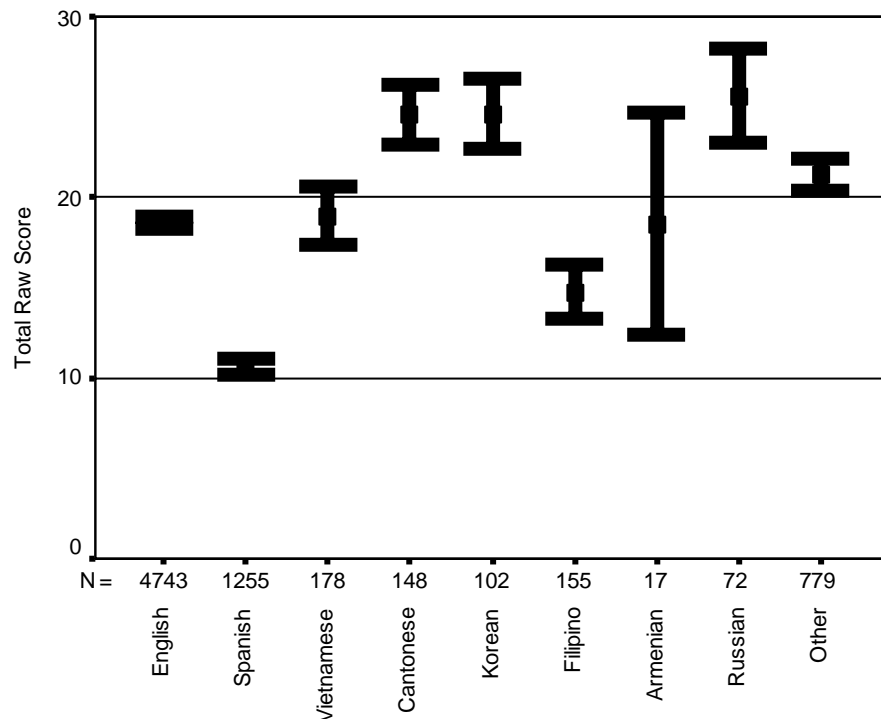
The scores of students with Armenian as a home language are significantly higher than the scores of students with Spanish as their home language. There are no other significant differences with any other language group.

The scores of students with Russian as a home language are significantly higher than those of students with English, Spanish, Vietnamese, Filipino, or “Other” as a home language. There are no other significant differences between the scores of students with Russian as a home language and any other language group.

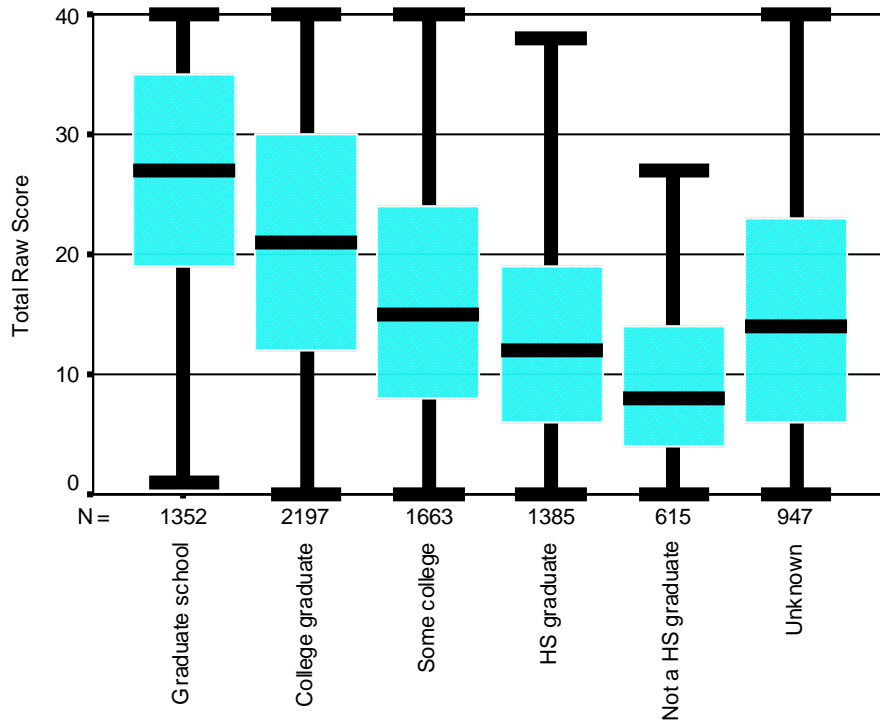
The scores of students with “Other” as a home language are significantly higher than those of students with English, Spanish or Filipino as a home language, and significantly lower than those of students with Cantonese, Korean, or Russian as a home language. There are no other significant differences between the students with “Other” as a home language and any other language group.

Other includes Albanian, Arabic, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Lahu, Lao, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 8 - Home Language



Grade 8 - Parent Education



Distribution of sampling means - Grade 8 - Parent Education

In this section, test scores are compared across groups of different levels of parent education²⁸. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate school education are significantly higher than those of students in all other Parent Education categories.

The scores of students whose parents have a college education are significantly lower than those whose parents have a graduate education, and significantly higher than all other groups.

The scores of students whose parents have some college education are significantly higher than those of students whose parents are not High School graduates and those whose parents are High School graduates, and significantly lower than those whose parents have a college education or a graduate education. There is not a significant

²⁸ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

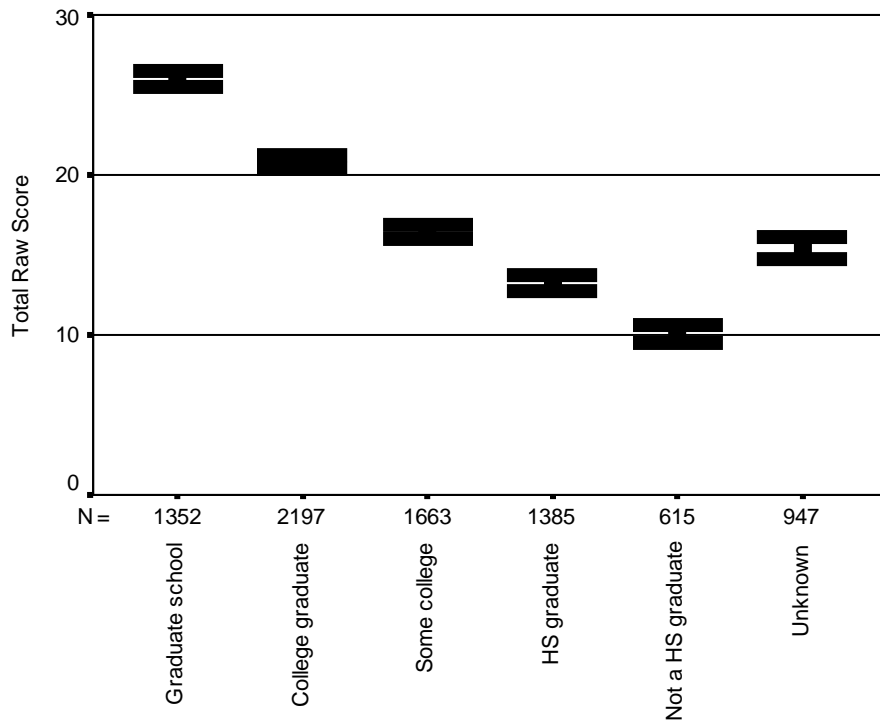
difference between scores of students whose parent education level is unknown and those whose parents have some college education.

The scores of students whose parents are High School graduates are significantly higher than those of students whose parents are not High School graduates, and significantly lower than the scores of students in all other Parent Education categories.

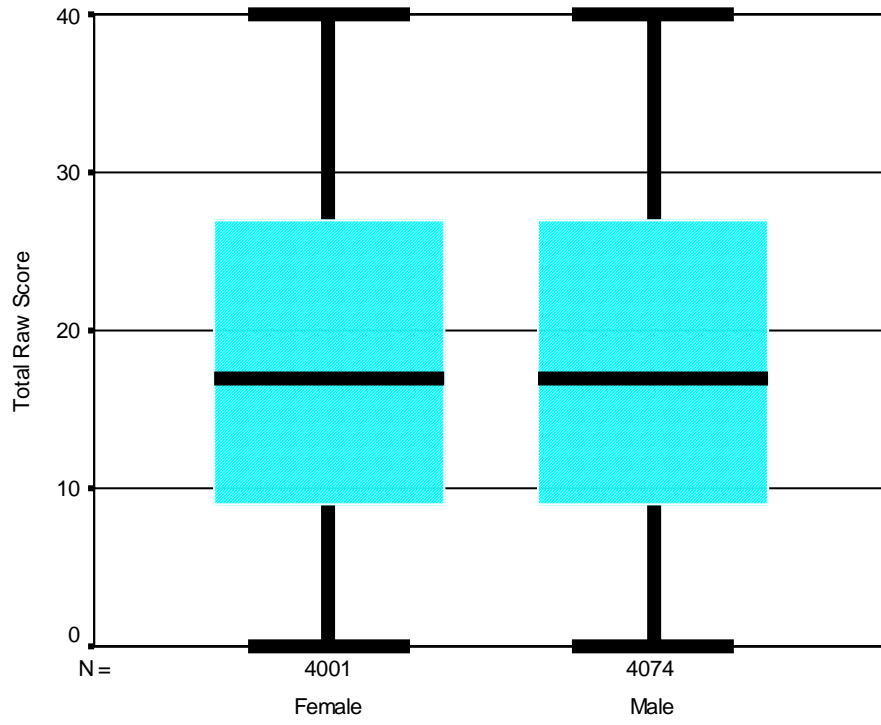
The scores of students whose parents are not High School graduates are significantly lower than the scores of students in all other Parent Education categories.

The scores of students whose parents' education is unknown are significantly higher than those of students whose parents are not High School graduates and those whose parents are High School graduates. The scores of students whose parents' education is unknown are significantly lower than those of students whose parents have a college degree or have a graduate education. There is not a significant difference between scores of students whose parent education level is unknown and those whose parents have some college education.

Distribution of sampling means - Grade 8 - Parent Education



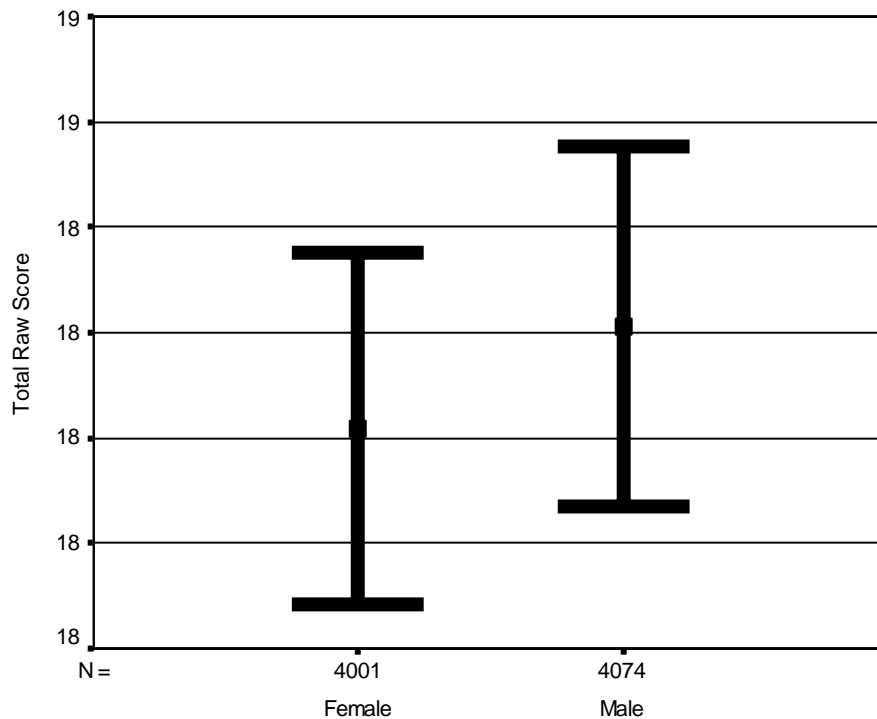
Grade 8 - Gender



Distribution of sampling means - Grade 8 - Gender

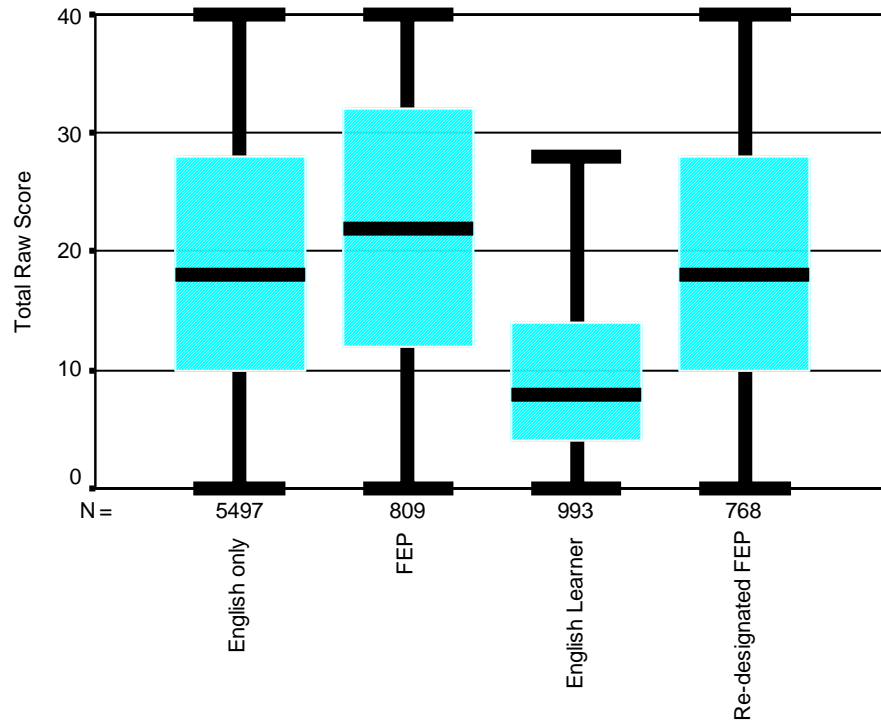
In this section, test scores are compared across gender²⁹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B, then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

There is no statistically significant difference in the scores of females and males.



²⁹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 8 - Language Fluency



Distribution of sampling means - Grade 8 - Language Fluency

In this section, test scores are compared across groups of different language fluency³⁰. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly lower than those of students described as Full English Proficiency (FEP) and significantly higher than those students described as English Learners. There is no significant difference between scores of students described as English Only, and those described as re-designated FEP.

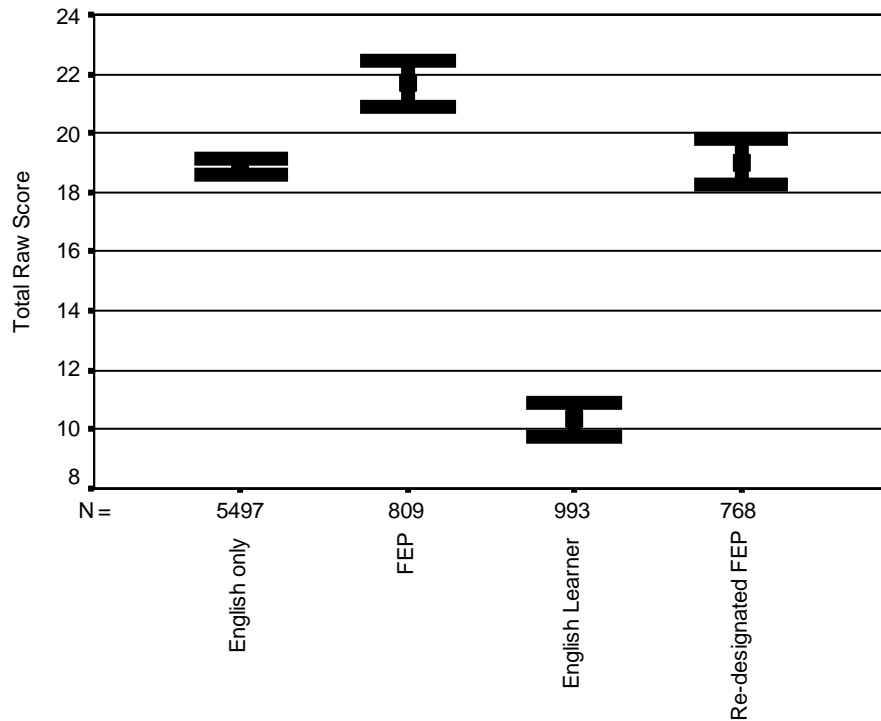
The scores of students in the FEP category are significantly higher than the scores of students in any other language fluency category.

The scores of students in the English Learner category are significantly lower than those of the students in all other category of language fluency.

³⁰ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

The scores of students in the Re-designated FEP category are significantly higher than those of students in the English Learner category, and lower than those of students in the FEP category. There is no significant difference between scores of students described as re-designated FEP, and those described as English Only.

Distribution of sampling means - Grade 8 - Language Fluency



8th grade

Task 1

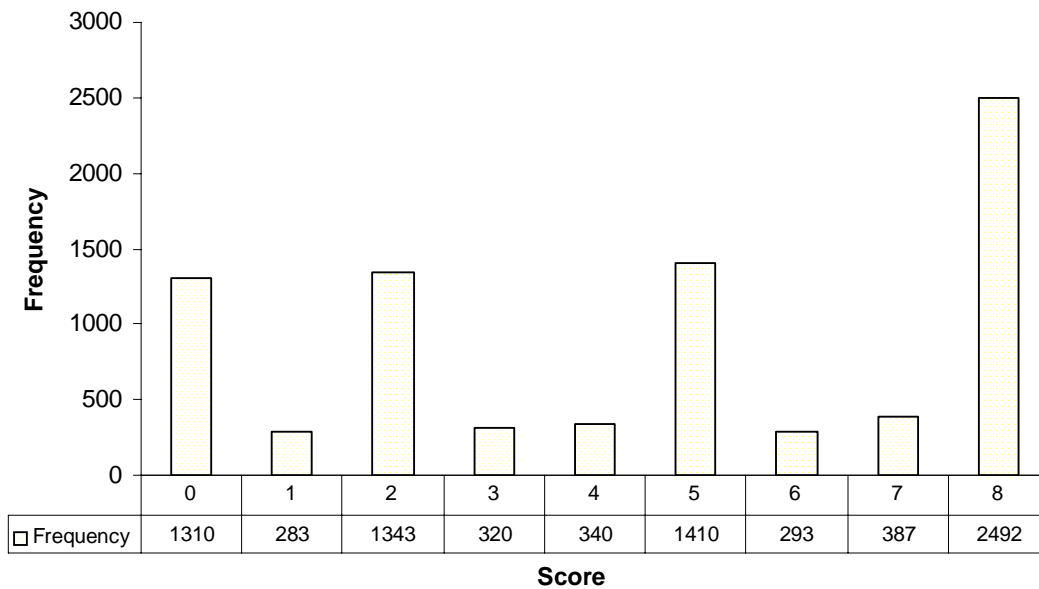
Pete's Numbers

Student Task	Use mathematical reasoning to solve number problems.
Core Idea 2 Mathematical Reasoning	<p>Employ forms of mathematical reasoning and justification appropriately to the solution of a problem.</p> <ul style="list-style-type: none"> • Extract pertinent information from situations and determine what additional information is needed • Verify and interpret results of a problem. • Invoke problem-solving strategies • Use mathematical language and representations including numerical tales and equations, simple algebraic equations, formulas, and graphs to make complex situations easier to understand
Core Idea 3 Algebra and Functions	<p>Understand relations and functions, analyze mathematical situations, and use models to solve problems involving quantity and change.</p> <ul style="list-style-type: none"> • Recognize and generate equivalent forms of simple algebraic expressions and solve linear equations

Frequency Distribution for each Task – Grade 8
Grade 8 – Pete’s Numbers

Pete's Numbers

Mean:4.49, S.D.: 2.99



Score:	0	1	2	3	4	5	6	7	8
% < =	16.0%	19.5%	35.9%	39.8%	44.0%	61.2%	64.8%	69.5%	100.0%
% > =	100.0%	84.0%	80.5%	64.1%	60.2%	56.0%	38.8%	35.2%	30.5%

The maximum score available for this task is 8 points.

The cut score for a level 3 response is 5 points.

Most students (about 84%) could find the numbers using the clues in part 1 of the task and about 80% could show the proof for picking those numbers. Many students (over 60%) could find the numbers and prove how they met the conditions for part 1 and find the numbers to fit the clues in part 2. About 30% of the students could use the clues to find the missing numbers in all 3 parts of the task and show how the numbers fit the clues. About 16% of the students scored no points on this task. 91% of the students with a score of zero attempted the task.

Based on teacher observations, this is what eighth grade students seemed to know and be able to do:

- Find numbers to fit number clues
- Use guess and check strategies to find missing numbers
- Write symbolic expressions or number sentences to fit problem statements

Areas of difficulty for eighth grade students, eighth grade students struggled with:

- Solving two equations for two unknowns
- Justifying solutions

Implications for Instruction:

Students at this grade level should be comfortable with the idea of constraints and finding numbers that meet multiple conditions. They should also have experience with the need to justify how a solution fits the conditions in the problem. Students should ask themselves, “Have I checked for each condition in the problem?” Just as students at lower grades need help transitioning from the concrete to numeric and symbolic representations, students in eighth grade need help transitioning from guess and check to the use of algebra as a problem-solving tool. What questions or experiences can help them make that transition? How can we help them connect the multiple guesses to the concept of a variable? How can we get them to see the connection between the guesses and a unique solution in equations with multiple variables? Could graphing help them to see some of these connections or relationships? Perhaps their experiences in Pre-Algebra need to be more focused on some of these big ideas around variables and justification.

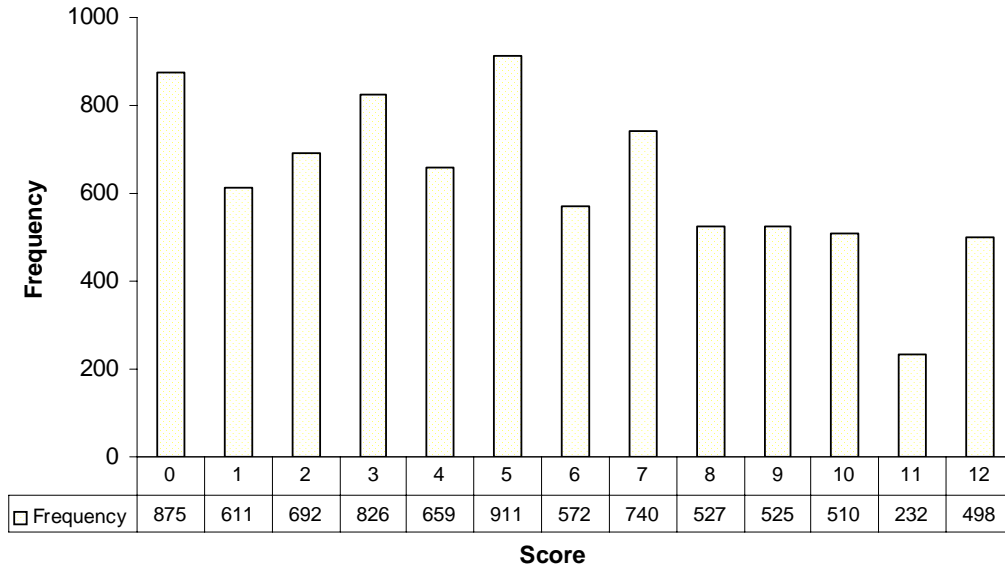
8th grade Task 2 Squares and Rectangles

Student Task	Use the properties of shapes to find similar shapes.
Core Idea 4 Geometry and Measurement	<p>Analyze characteristics and properties of two- and three-dimensional geometric shapes; develop mathematical arguments about geometric relationships; apply transformations and use symmetry to analyze mathematical situations; and apply appropriate techniques tools, and formulas to determine measurements.</p> <ul style="list-style-type: none"> • Understand relationships among the angles, side lengths, perimeter, and area of similar objects • Describe sizes, positions, and orientations of shapes under informal transformations such as flips, turns, slides, and scaling.
Core Idea 2 Mathematical Reasoning	<p>Employ forms of mathematical reasoning and justification appropriately to the solution of a problem.</p> <ul style="list-style-type: none"> • Formulate conjectures and test them for validity

Grade 8 – Squares and Rectangles

Squares and Rectangles

Mean: 5.24, S.D.: 3.56



Score:	0	1	2	3	4	5	6	7	8	9	10	11	12
% <=	10.7%	18.2%	26.6%	36.7%	44.8%	55.9%	62.9%	72.0%	78.4%	84.8%	91.1%	93.9%	100.0%
% >=	100.0%	89.3%	81.8%	73.4%	63.3%	55.2%	44.1%	37.1%	28.0%	28.0%	44.1%	8.9%	6.1%

The maximum score available for this task is 12 points.
The cut score for a level 3 response is 6 points.

Most students (more than 80%) could define a square and find 1 or 2 squares in part 3 of the task. More than half the students could define a square, draw 2 squares on the grid, and correctly give the coordinates for the vertices. 20% of the students could define a square, find and give the coordinates for 2 of the squares, draw a rectangle with sides in a given ratio and find the area of the rectangle. 6% of the students met all the demands of the task. 11% of the students scored no points on this task. Of those students, almost 94% attempted the task.

Based on teacher observations, this is what eighth grade students seemed to know and be able to do:

- Define a square.
- Locate squares on a grid, whose sides were located on grid lines.
- Give coordinates of points on a grid.
- Find area of a rectangle.

Areas of difficulty for eighth graders, eighth grade students struggled with:

- Defining a rectangle.
- Finding squares whose sides were not parallel to the sides of the grid.
- Finding the dimensions of a shape when given the ratio of the sides.
- Finding the total number of rectangles that would fit on a grid.

Implications for Instruction:

Students at this level need to know mathematical definitions of common geometric shapes like squares and rectangles and be able to analyze the properties of two-dimensional figures. More experiences with classifying and sorting, working with attributes, or logic problems might help them become more precise in the use of definitions. They should be able to find and to draw those shapes on a coordinate graph, including those shapes with sides not parallel to the sides of the grid. They also need to be able to use coordinates to locate points a graph. Students should be comfortable with the idea that not all points in a graph are located on the grid intersections. Points can lie anywhere on the continuum and represent fractional distances. Students at this grade level should also be comfortable working with simple ratios to find different sides. Their thinking should be flexible enough to see that the horizontal axis could be the width or height of a shape. Students should work with shapes in a variety of orientations, including those with sides not parallel to the side of the paper.

8th grade

Task 3

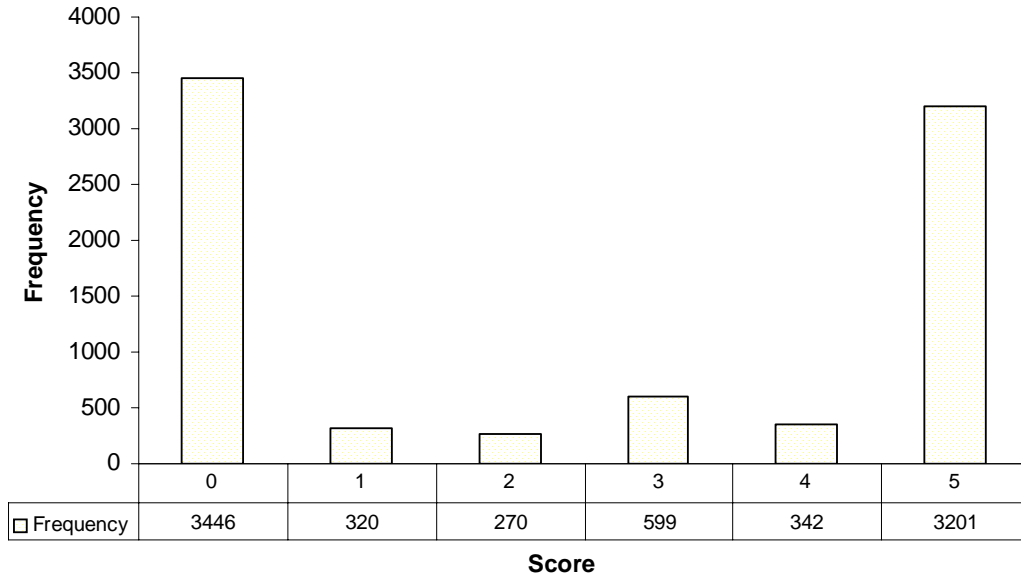
Sport Injuries

Student Task	Use real data to interpret a circle graph regarding sports injuries.
Core Idea 5 Data Analysis	Formulate questions that can be addressed with data and collect, organize, analyze, and display relevant data to answer them. <ul style="list-style-type: none"> • Use graphical representations of data
Core Idea 2 Mathematical Reasoning	Employ forms of mathematical reasoning and justification appropriately to the solution of a problem. <ul style="list-style-type: none"> • Extract pertinent information from situations and determine what additional information is needed • Invoke problem-solving strategies
Core Idea 1 Number and Operation	<ul style="list-style-type: none"> • Work flexibly with fractions, decimals, and percents to solve problems.

Grade 8 – Sports Injuries

Sports Injuries

Mean: 2.45; S.D.: 2.30



Score:	0	1	2	3	4	5
% < =	42.1%	46.1%	49.4%	56.7%	60.9%	100.0%
% > =	100.0%	57.9%	53.9%	50.6%	43.3%	39.1%

The maximum score available for this task is 5 points.
The cut score for a level 3 response is 2 points.

About half the students (51%) could correctly convert three or more of the percentages to degrees in a circle in part one of the task. Almost 40% of the students could convert percentages to degrees and degrees to percentages in the context of a circle graph. More than 40% of the students scored no points on this task. More than 80% of those students attempted the task.

Based on teacher observations, this is what eighth grade students seemed to know and be able to do:

- Estimate the percentage of a circle given the number of degrees or estimate the degrees when given a percentage.
- Convert percentages to degrees.

Areas of difficulty for eighth graders, eighth grade students struggled with:

- Calculating percentages when given the number of degrees.

- Converting percentages to decimals and vice versa.

Implications:

Students need to know that there are 360 degrees in a circle. At this grade level they should be proficient at calculating with percents. Students, who lack a conceptual understanding of percents, have difficulty remembering and applying procedural rules when solving problems or determining if their answers are reasonable. Some students have conceptual understanding and can do estimation, but students at this grade level should be fluent with these conversions. A few students still do not understand the relationship between decimals and percents.

8th grade

Task 4

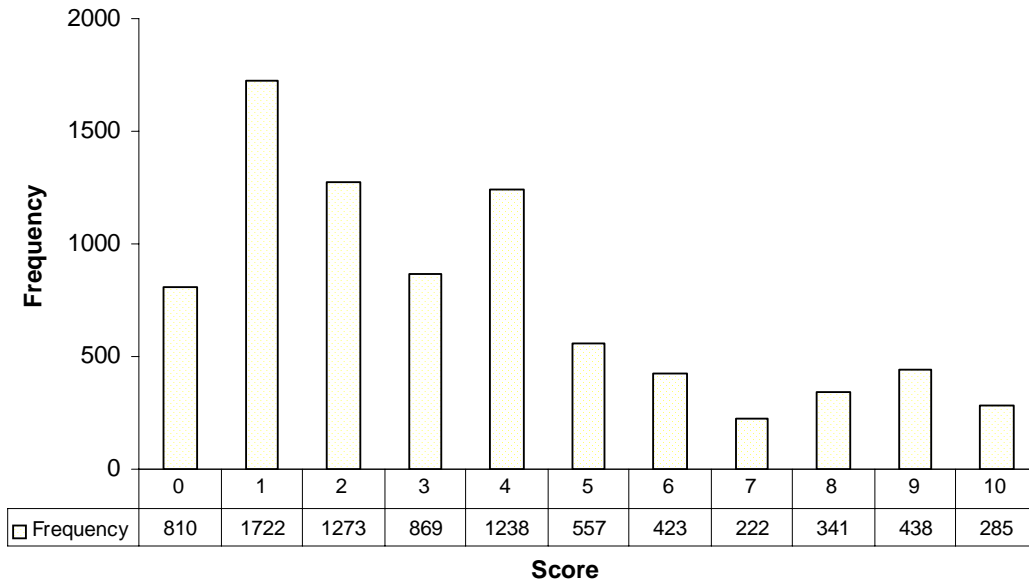
Dots and Squares

Student Task	Find and table number patterns in a geometric content. Find and use rules or formulas to answer questions.
Core Idea 3 Algebra and Functions	<p>Understand relations and functions, analyze mathematical situations, and use models to solve problems involving quantity and change.</p> <ul style="list-style-type: none"> • Use tables to analyze the nature of changes on quantities in linear relationships • Recognize and generate equivalent forms of simple algebraic expressions and solve linear equations. • Represent, analyze, and generalize a linear relationship (7th grade) • Use symbolic algebra to represent situations to solve problems (7th grade)
Core Idea 2 Mathematical Reasoning	<p>Employ forms of mathematical reasoning and justification appropriately to the solution of a problem.</p> <ul style="list-style-type: none"> • Use mathematical language and representations to make situations easier to understand

Grade 8 – Dots and Squares

Dots and Squares

Mean: 3.45, S.D.: 2.77



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	9.9%	31.0%	46.5%	57.2%	72.3%	79.1%	84.3%	87.0%	91.2%	96.5%	100.0%
% > =	100.0%	90.1%	69.0%	53.5%	42.8%	27.7%	20.9%	42.8%	13.0%	8.8%	3.5%

The maximum score available for this task is 10 points.

The cut score for a level 3 response is 6 points.

Most students (about 90%) could fill in the table with the correct perimeter for each square. Many students (about 70%) could also find the number of inside dots for a square. A little less than half the students (43%) could find the perimeter and the inside dots for the squares and the rectangles. About 21% could meet standards by filling in the tables for perimeter and dots and find the dimensions of square with 49 inside dots. Less than 5% of the students met all the demands of the task. Almost 10% of the students scored no points on this task. Of those more than 60% attempted some part of the task. *Based on teacher observations, this is what eighth grade students seemed to know and be able to do:*

- Find the perimeter of squares and rectangles.
- Find the number of inside dots for a square or rectangle.

Areas of difficulty for eighth graders, eighth grade students struggled with:

- Writing rules or formulas for geometric patterns.
- Using rules to work backwards.

- Understanding how to check a rule to see if it works for all the cases in the given information. (Making generalizations on too little information.)

Instructional Implications:

When students look at pattern problems, it is helpful to visualize what is changing and what is staying the same. As they progress through the grades this information could be used to help them write a rule or formula. At this grade level, they can no longer rely on drawing pictures or doing repeated addition to find the solutions to complex problems. Students at this grade level should be proficient at answering a variety of questions about patterns. They need to recognize that patterns can grow in more than one direction and be able to investigate those changes. Students should work with patterns with exponential growth as well as linear growth. Students should develop the habit of verifying their rules or formulas to see if they work for more than one example.

8th grade

Task 5

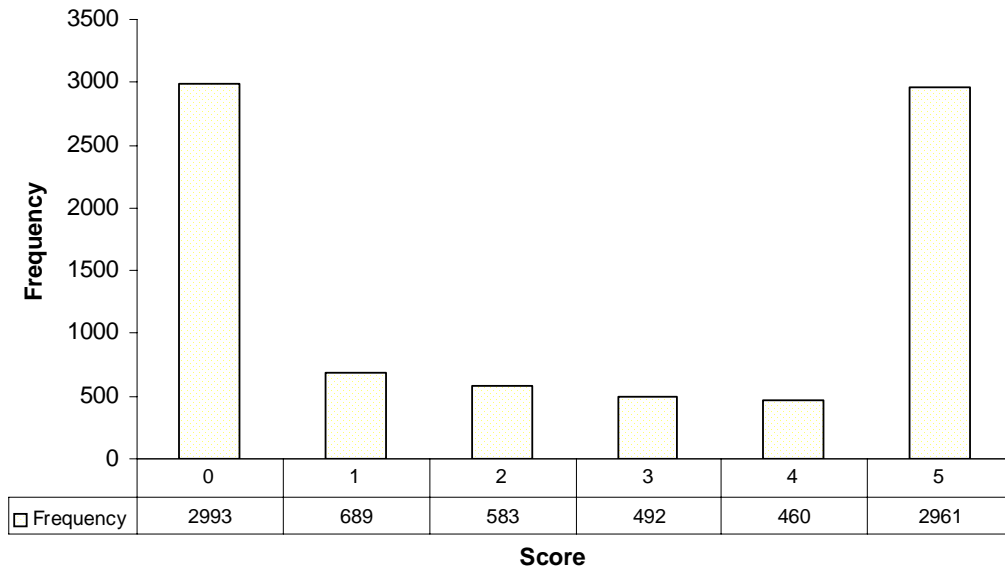
Number Pairs

Student Task	Identify number pairs on a coordinate grid.
Core Idea 3 Algebra and Functions	<p>Understand relations and functions, analyze mathematical situations, and use models to solve problems involving quantity and change.</p> <ul style="list-style-type: none"> • Explore relationships between symbolic expressions and graphs of lines • Relate and compare different forms of representations for relationships including words, tables, graphs in the coordinate plane and symbols (7th grade)

Grade 8 – Number Pairs

Number Pairs

Mean: 2.44, S.D.: 2.21



Score:	0	1	2	3	4	5
% < =	36.6%	45.0%	52.2%	58.2%	63.8%	100.0%
% > =	100.0%	63.4%	55.0%	47.8%	41.8%	36.2%

The maximum score available for this task is 5 points.

The cut score for a level 3 response is 3 points.

Many students (about 63%) can match one or more descriptions to a graph. A little more than half (55%) can match two or more descriptions to their graphs. 36% of the students can match all descriptions to their graphs and make a coordinate graph from a verbal description. 36% of the students scored no points on this task. About half of these students attempted the task.

Based on teacher observations, this is what eighth grade students seemed to know and be able to do:

- Match a description about the first number of a coordinate pair to its graph.
- Understand that graphs of functions go in straight lines.
- Recognize when a plot of discrete points is appropriate

Areas of difficulty for eighth graders, eighth grade students struggled with:

- Matching descriptions of coordinate pairs to their graphs.
- Making a graph from a verbal description.
- Choosing coordinate pairs to match a verbal description.

Instructional Implications:

Given a simple rule, students should be able to make an equation and use that equation to develop a table of values. Students should be comfortable using tables to graph simple equations. Students need experiences describing given graphs as well as experiences plotting points. Students should have some system of checking whether different coordinates fit the conditions of the equations.

Blue Print for Course One - MAC

Alignment of Tasks to NCTM Content and Process Standards

	Vacuum Cleaning	Snakes	Crisscross Numbers	Conference Tables	Number Towers
NUM			S		S
ALG		S	P	P	P
GEO	P			S	
MEAS	S				
DATA		P			
PS					
REAS	*	*	*	*	*
COMM			*	*	*
CONN		*		*	
REP	*		*	*	*

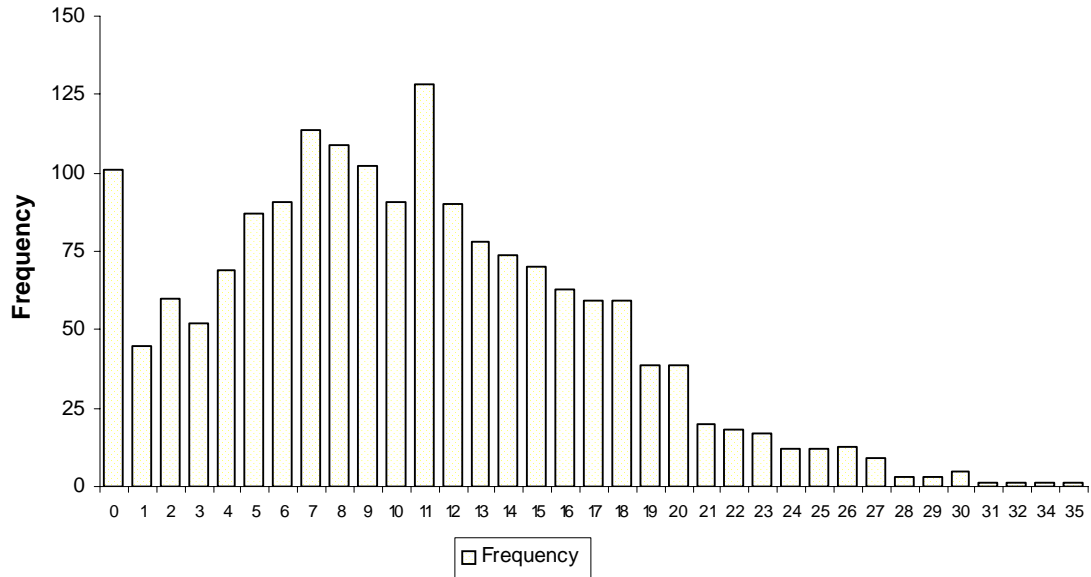
NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

- P denotes Primary NCTM Content Standard
 S denotes Secondary NCTM Content Standard
 • denotes NCTM Process Standard

Overall Frequency Distribution by Total Score

Grade 9

Mean=10.42; S.D.=6.47



Level Frequency Distribution Chart and Frequency Distribution

2003 - Numbers of students Course One: 1736
tested:

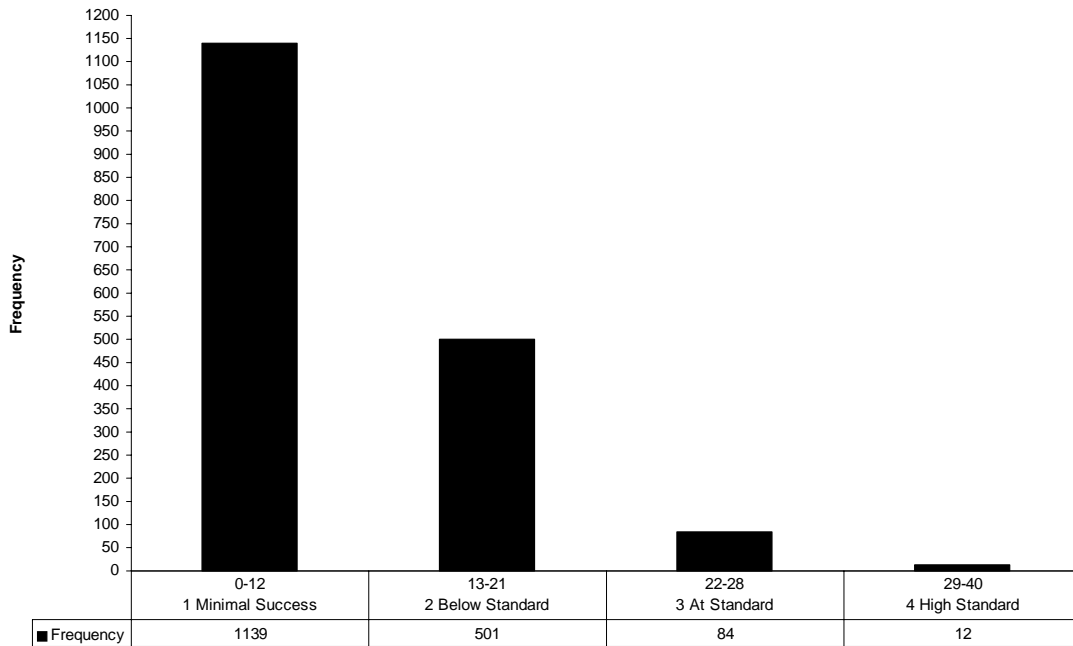
Course One 1999 and 2001

Level	% at ('99)	% at least ('99)	% at ('01)	% at least ('01)
1	37%	100%	22%	100%
2	43%	63%	62%	78%
3	13%	20%	15%	16%
4	7%	7%	1%	1%

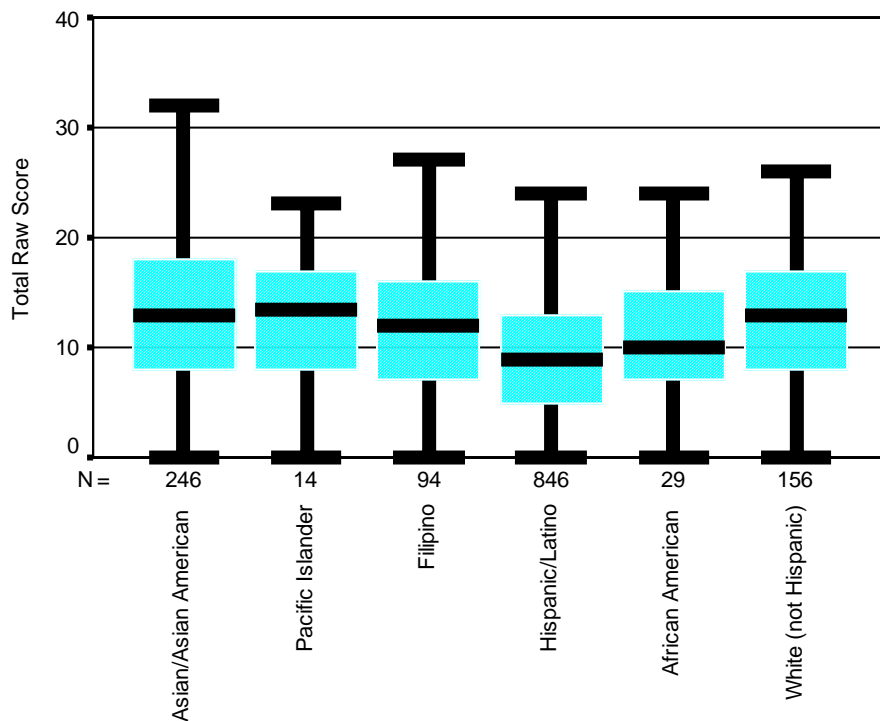
Course One 2002 - 2003

Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	18%	100%	66%	100%
2	61%	82%	29%	34%
3	20%	22%	5%	6%
4	2%	2%	1%	1%

Course One – Distribution by Performance Levels



Grade 9 – Ethnicity -The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values * within a category lie between the highest and lowest horizontal bars.



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Grade 9 - Ethnicity

In this section, test scores are compared across different ethnic groups³¹. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of Asian/Asian American students are significantly higher than those of Hispanic/Latino students. There are no significant differences between Asian/Asian American students and any other ethnic groups. There are no significant differences between Pacific Islander students and any other group.

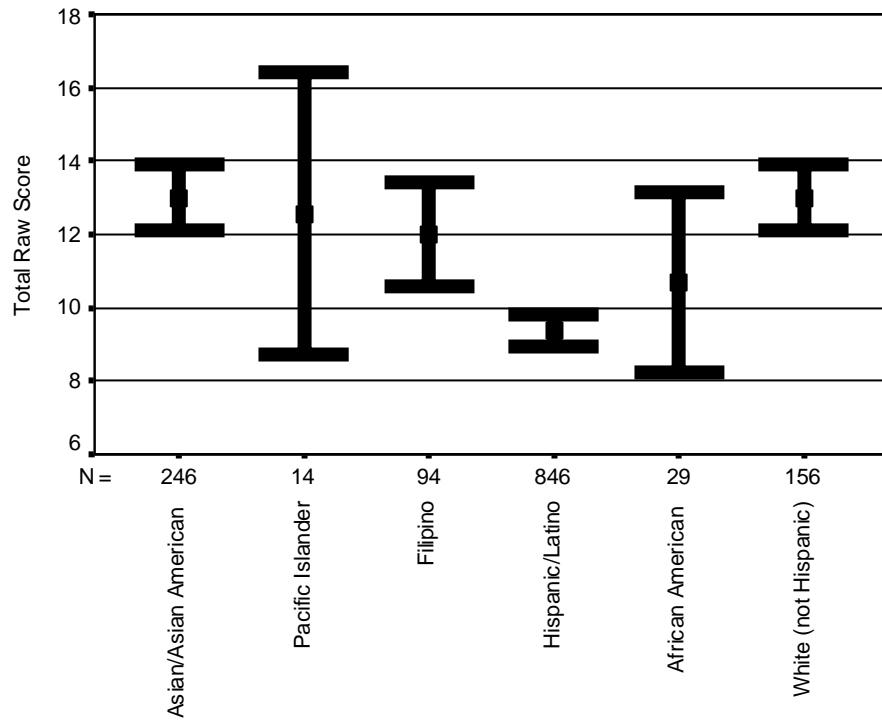
The scores of Filipino students are significantly higher than those of Hispanic/Latino students. There are no significant differences between Filipino students and any other ethnic group.

The scores of Hispanic/Latino students are significantly lower than those of Asian/Asian American, Filipino, and White students. There are no significant differences between Hispanic/Latino students and any other ethnic group. There are no significant differences between African American students and any other group.

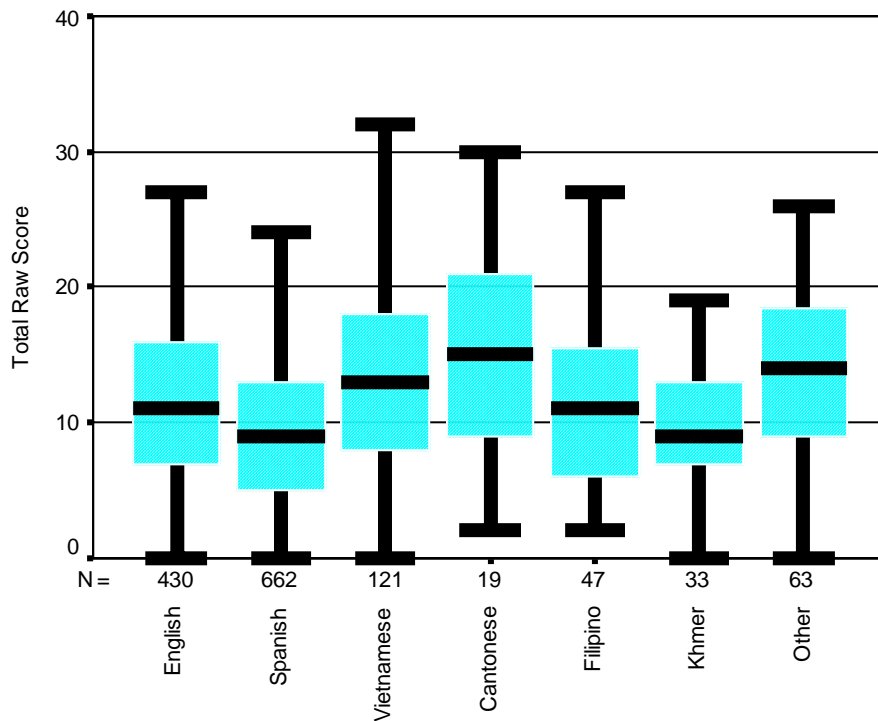
The scores of White students are significantly higher than those of Hispanic/Latino students. There are no significant differences between White students and any other ethnic group.

³¹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 9 - Ethnicity



Grade 9 - Home Language



Grade 9 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home³². One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English as a home language are significantly lower than those of students with Cantonese as a home language, and significantly higher than those of students with Spanish as a home language. There are no significant differences between students with English as a home language and any other language group.

The scores of students with Spanish as a home language are significantly lower than those of students with English, Vietnamese, Cantonese, and “Other” as a home language. There are no significant differences between students with Spanish as a home language and any other language group.

The scores of students with Vietnamese as a home language are significantly higher than those of students with Spanish as a home language. There are no significant differences between students with Vietnamese as a home language and any other language group.

The scores of students with Cantonese as a home language are significantly higher than those of students with English, Spanish, and Khmer as a home language. There are no significant differences between students with Cantonese as a home language and any other language group.

There are no significant differences between students with Filipino as a home language and any other language group.

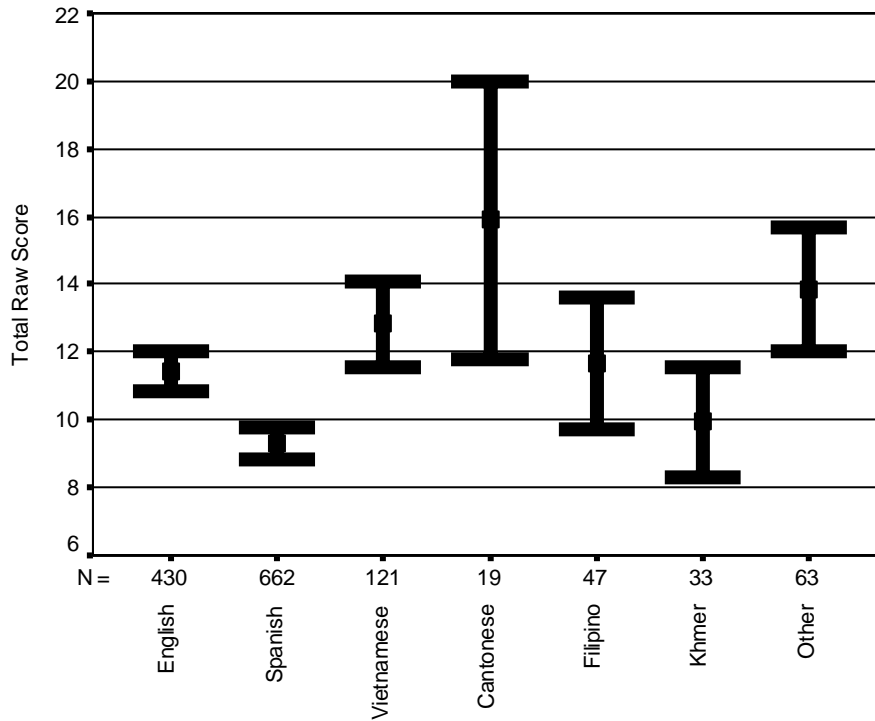
The scores of students with Khmer as a home language are significantly lower than those of students with Cantonese as a home language. There are no significant differences between students with Khmer as a home language and any other language group.

The scores of students with “Other” as a home language are significantly higher than those of students with Spanish as a home language. There are no significant differences between students with “Other” as a home language and any other language group.

Other includes Albanian, Arabic, Armenian, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Korean, Lahu, Lao, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Russian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

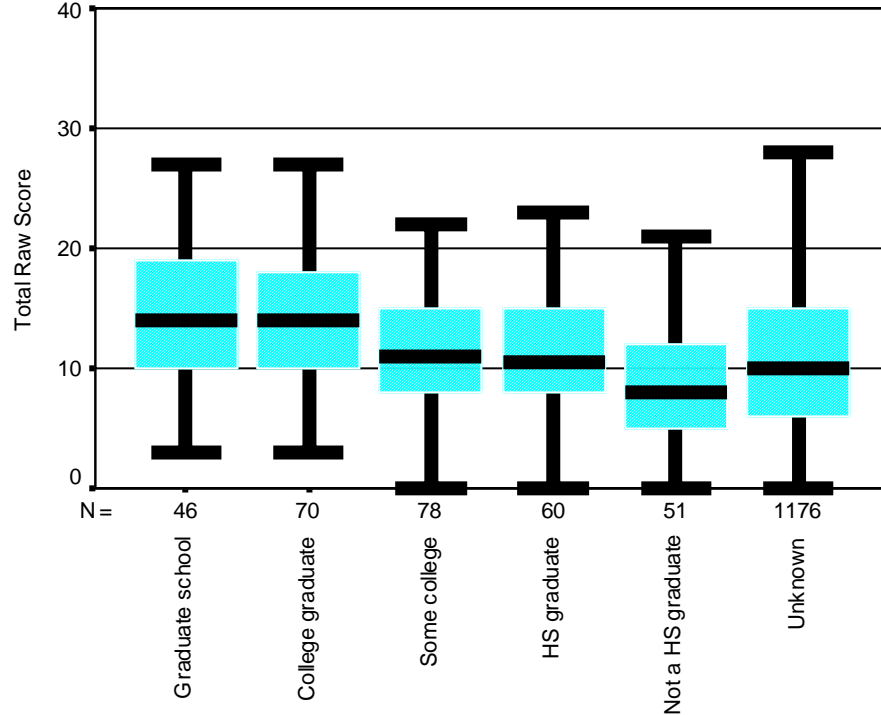
³² Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 9 - Home Language



Other includes Albanian, Arabic, Armenian, Assyrian, Burmese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Korean, Lahu, Lao, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Russian, Samoan, Serbo-Croatian, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Grade 9 - Parent Education



Grade 9 - Parent Education

In this section, test scores are compared across groups of different levels of parent education.³³ One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate degree are significantly higher than those whose parents are not High School graduates and those whose parents' education is unknown. There are no significant differences between students whose parents have a college education and any other group.

The scores of students whose parents have a college education are significantly higher than those whose parents are not High School graduates and those whose parents' education is unknown. There are no significant differences between students whose parents have a college education and any other group.

There are no significant differences between students whose parents have some college education and any other group.

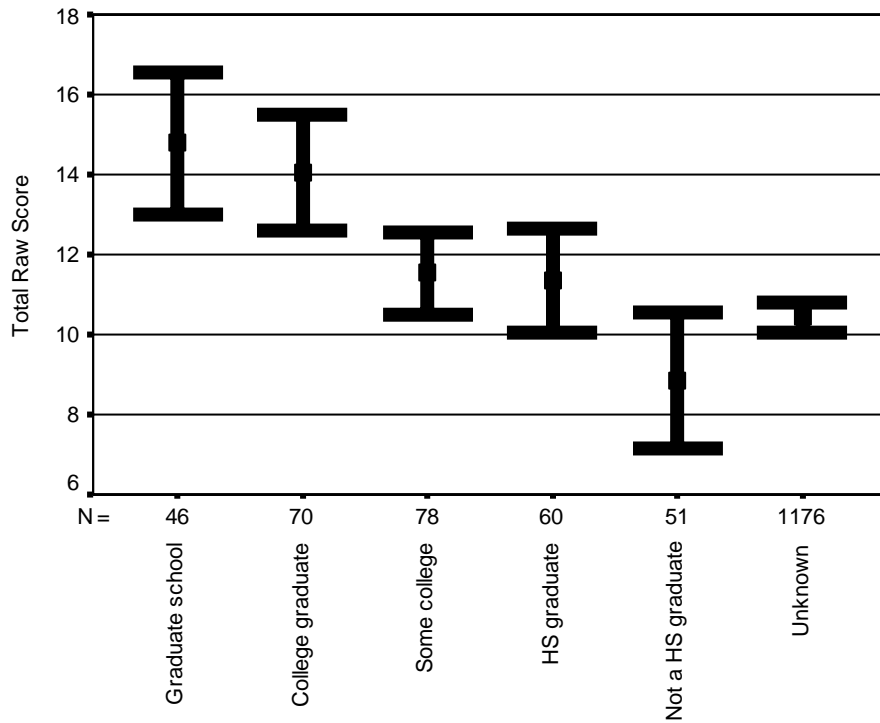
There are no significant differences between students whose parents are High School graduates and any other group.

The scores of students whose parents are not High School graduates are significantly lower than those whose parents have a college education and a graduate degree. There are no significant differences between students whose parents are not High School graduates and any other group.

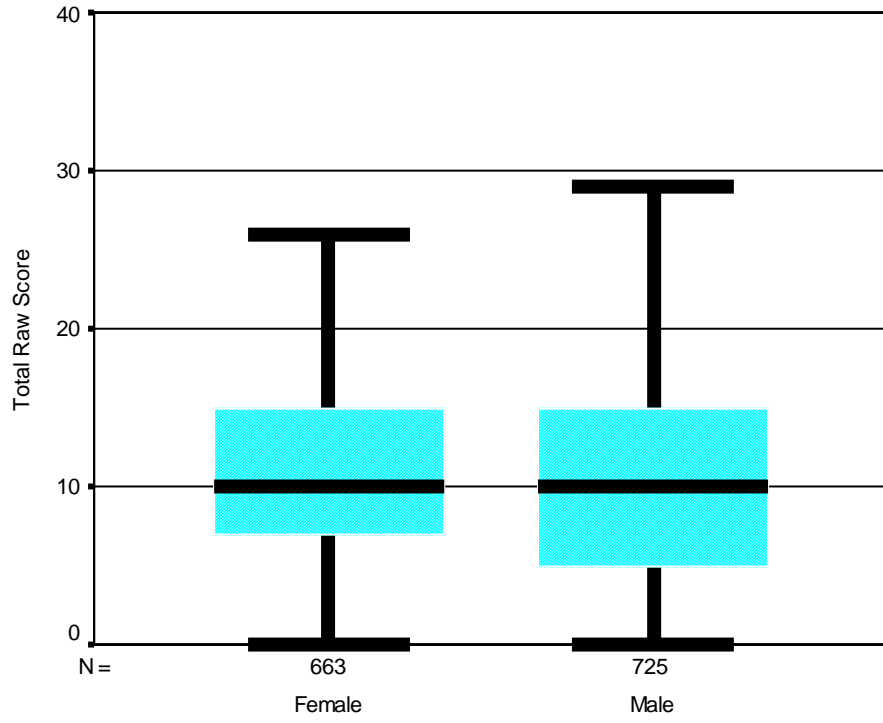
The scores of students whose parents' education is "unknown" are significantly lower than those whose parents are college graduates and have a graduate degree. There are no significant differences between "unknown" students and any other group.

³³ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means -Grade 9 - Parent Education



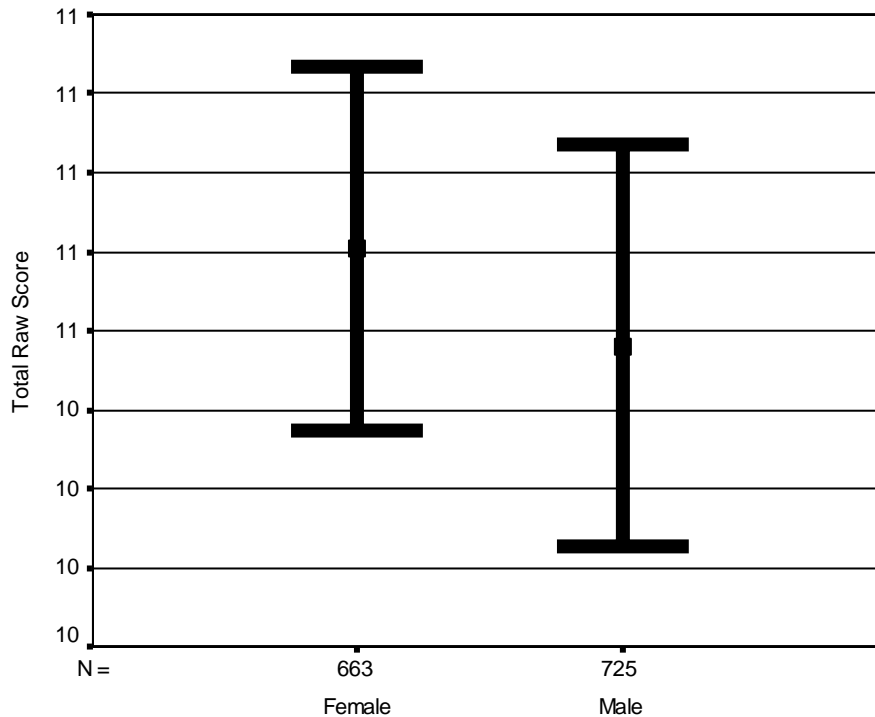
Grade 9 - Gender



Distribution of sampling means - Grade 9 – Gender

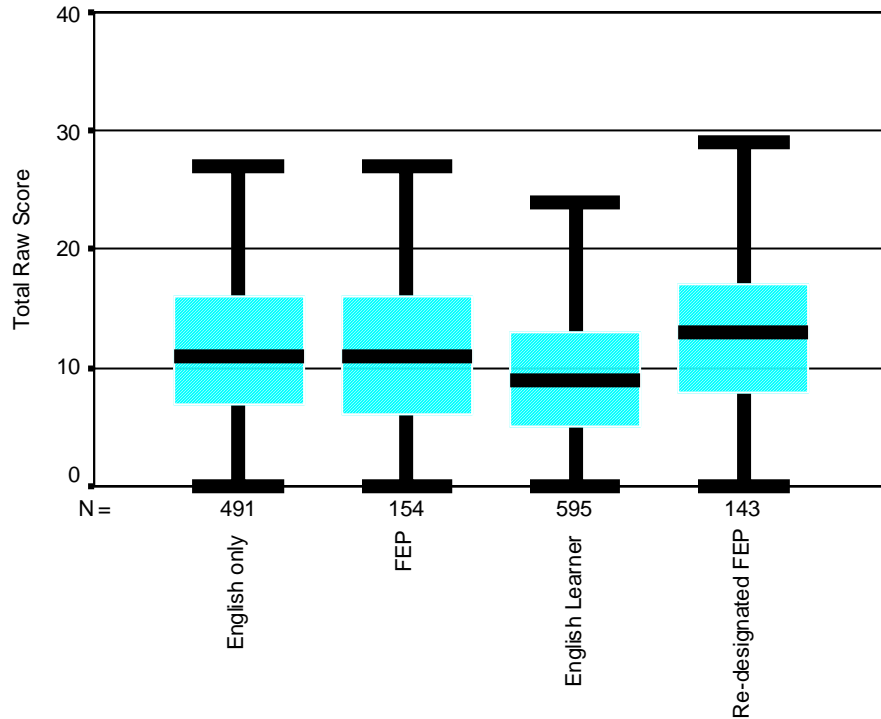
In this section, test scores are compared across gender.³⁴ One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

There is no significant difference between the scores of females and males.



³⁴ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 9 - Language Fluency



Language Fluency - Grade 9

In this section, test scores are compared across groups of different language fluency.³⁵ One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students with English Only are significantly higher than those students described as English Learners. There are no significant differences between students described as English Only and any other group.

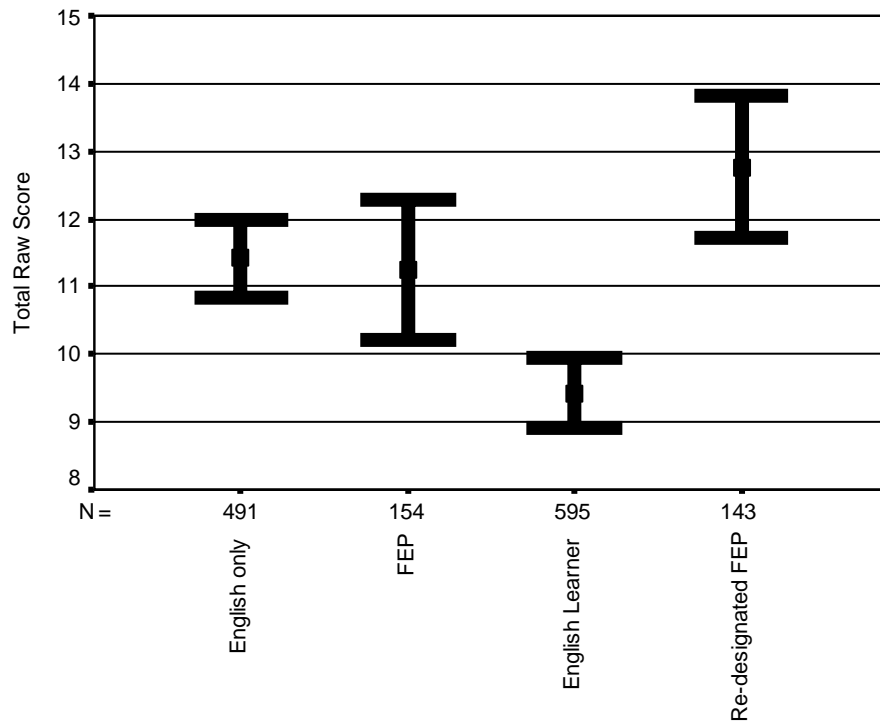
The scores of students in the FEP category are significantly higher than those of students in the English Learner category. There is no significant difference between the scores of students in the FEP category and any other group.

The scores of students in the English Learner category are significantly lower than all other groups.

The scores of students in the R-FEP category are significantly higher than those of students in the English Learner category. There is no significant difference between the scores of students in the FEP category and any other group.

³⁵ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 9 - Language Fluency



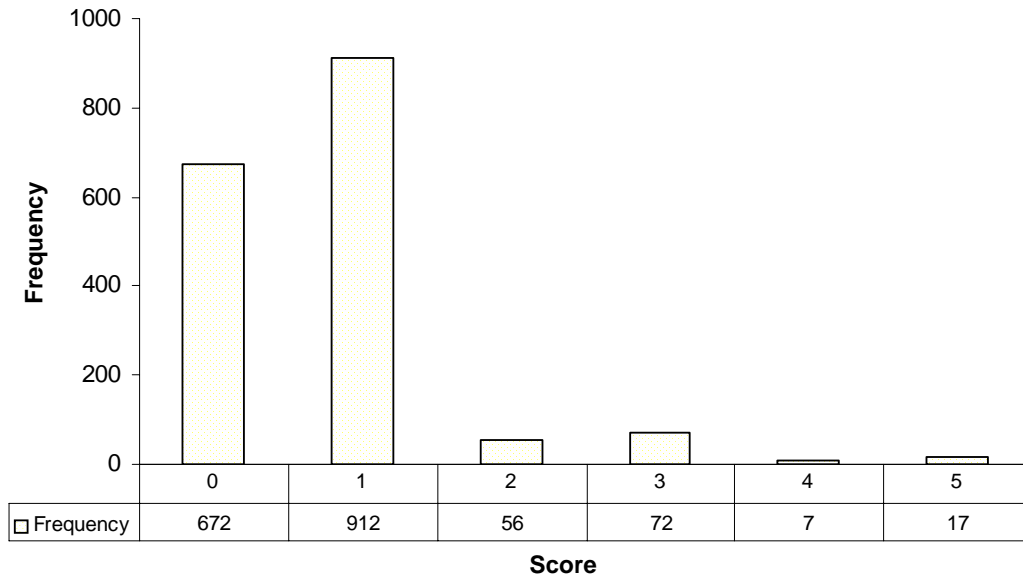
Vacuum Cleaning

Student Task	Given a radius, make an arc to show area covered by a vacuum cleaner. Use information about furniture to draw areas accessible to vacuum cleaner and uncleaned area on a scaled diagram.
Core Idea 4 Geometry/Meas.	Understand measurable attributes of objects; develop mathematical arguments about geometric relationships. Solve problems that involve measurement units and scale factors. Use geometric models to gain insights into, and answer questions in, other areas of mathematics.
Core Idea 2 Mathematical Reasoning	Show mathematical reasoning in a variety of ways including words, numbers, symbols, pictures, charts, graphs, tables, diagrams, and models. Draw reasonable conclusions about a situation being modeled.

Frequency Distribution for each Task – Grade 9
Grade 9 – Vacuum Cleaning

Vacuum Cleaning

Mean: .78, S.D.: .85



Score:	0	1	2	3	4	5
% < =	38.7%	91.2%	94.5%	98.6%	99.0%	100.0%
% > =	100.0%	61.3%	8.8%	5.5%	1.4%	1.0%

The maximum score available for this task is 5 points.
The cut score for a level 3 response is 3 points.

Many students (approximately 61%) could correctly identify the uncleaned area of the bedroom. Less than 10% of the students could meet the other demands of the task. Over 38 % of the students scored no points on this task. Approximately 80% of those students with a score of 0 attempted the problem.

Based on teacher observations, this is what Algebra One students seemed to know and be able to do:

- Understand the basic relationships between the approximate spaces cleaned by A and B, in order to identify the approximate location of the uncleaned area.
- Demonstrate some use of measurement in making their diagrams.
- Recognize that a cord from B will be able to clean some part of the space around the end of the bed.

Areas of difficulty for Algebra One students, Algebra One students struggled with:

- Identifying the cord as a radius and drawing arcs.
- Use of accurate placement of all end points.
- Identifying the change in radius caused by the bed.

Implications for Instruction:

Students need more experience with measurement and using appropriate mathematical tools. Students need to have practice making sense of diagrams and making their own diagrams as an aid to problem solving. Students need more experience applying mathematics in problem-solving contexts or in contexts different from the problem sets in standard textbooks. This allows teachers to check for deeper understanding and the ability of students to transfer knowledge.

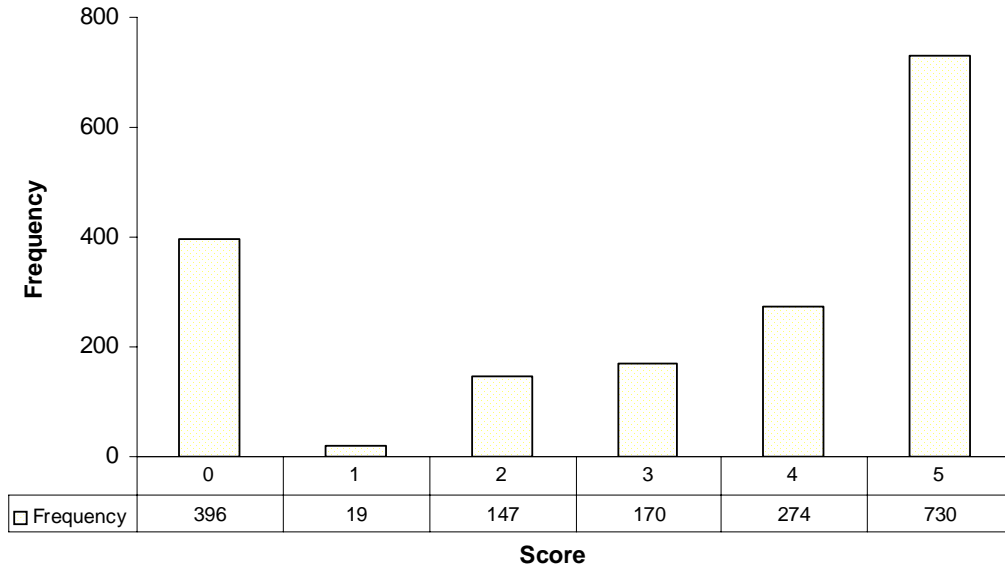
Snakes

Student Task	Read and interpret scatter plots. Locate points on a scatter plots to identify which scatter plot best fits the coordinates or values given.
Core Idea 5 Data Analysis	<ul style="list-style-type: none">• Understand the relationship between two sets of data, display such data in a scatterplot, and describe trends and shape of the plot including correlations and lines of best fit.• Make inferences based on the data and evaluate the validity of conclusions drawn.
Core Idea 2 Mathematical Reasoning	<ul style="list-style-type: none">• Show mathematical reasoning in a variety of ways, including words, numbers, symbols, pictures, charts, graphs, tables, diagrams and models.• Draw reasonable conclusions about a situation being modeled.

Grade 9 – Snakes

Snakes

Mean: 3.21, S.D.: 2.00



Score:	0	1	2	3	4	5
% < =	22.8%	23.9%	32.4%	42.2%	57.9%	100.0%
% > =	100.0%	77.2%	76.1%	67.6%	57.8%	42.1%

The maximum score available on this task is 5 points.

The cut score for a level 3 response is 3 points.

Most students (approximately 78%) could successfully identify the species of at least one of the snakes, usually 1 or 5. More than half the students could identify the species of 4 or more of the snakes. 42% of the students met all the demands of the task. Almost 23% of the students scored no points on this task. 52% of the students with scores of zero did not attempt the problem.

Based on teacher observations, this is what Algebra one students seemed to know and be able to do:

- Read and compare scatter plots.
- Correctly locate points on a grid.
- Interpret graph of best fit for information from a table.

Areas of difficulty for Algebra One students, Algebra One students struggled with:

- Attempting problems for which they were unsure of the solution.
- Interpreting relationships and fits for values that did not exactly fit the points on the grids.

Implications for Instruction:

Students at this grade level should be familiar not only with plotting points and making scatter plots, but should see the purpose and usefulness of scatter plots for looking at trends, making predictions, modeling situations, and solving problems. Students at this grade level also need to feel that their thinking is important and feel comfortable about showing their work and have the confidence in their mathematical thinking to attempt unfamiliar problems.

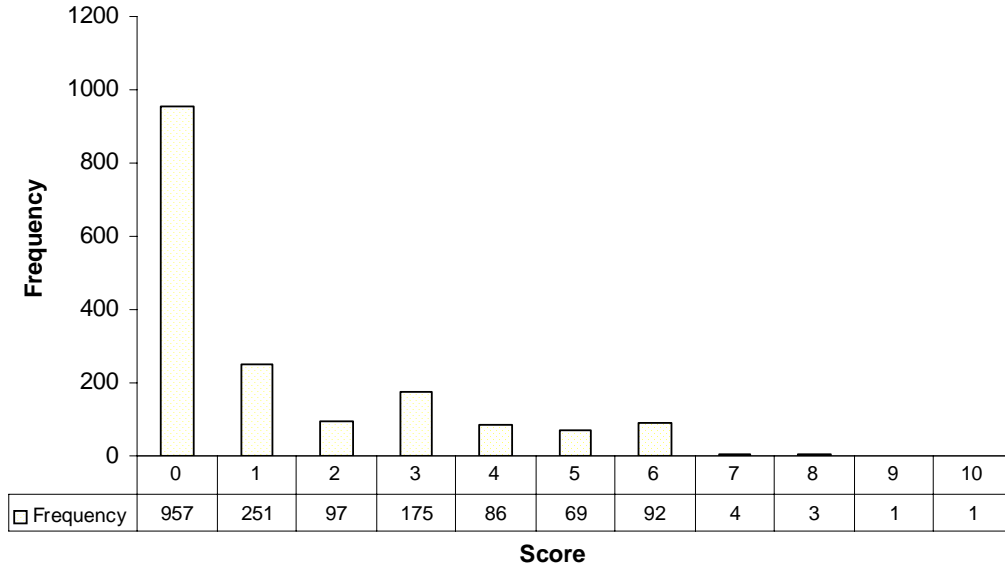
Crisscross Numbers

Student Task	Investigate number patterns on a hundreds chart. Describe rules or patterns in words or symbols. Use algebra to prove why the rules hold true for all cases.
Core Idea 3 Algebraic Properties and Representations	Represent and analyze mathematical situations and structures using algebraic symbols. Use symbolic algebra to represent and explain mathematical relationships. Judge the meaning, utility, and reasonableness of results of symbolic manipulations. Use symbolic expressions to represent relationships arising from various contexts.

Grade 9 – Crisscross Numbers

Crisscross Numbers

Mean: 1.31, S.D.: 1.89



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	55.1%	69.6%	75.2%	85.3%	90.2%	94.2%	99.5%	99.7%	99.9%	99.9%	100.0%
% > =	100.0%	44.9%	30.4%	24.8%	14.7%	9.8%	5.8%	0.5%	0.3%	0.1%	0.1%

The maximum score available on this task is 10 points.

The cut score for a level 3 response is 5 points.

Less than half the students (approximately 45%) could identify that adding the numbers on the diagonals of the crisscross gave you a sum of 70 or the same sum. About 25% of the students could make both of those observations. Less than 1% of the students used algebra appropriately to describe the relationships of the numbers in the crisscross or to prove why the rules held true. 55% of the students scored no points on this task. Of those students with a score of zero, approximately 59% made some attempt at solving the problem. Of the remaining 41% who didn't try this task, 33% made attempts at one or more of the later tasks on the test.

Based on teacher observations, this is what Algebra One students seemed to know and be able to do:

- Calculate sums of diagonal numbers on a grid.
- Recognize that the sums would remain the same if the design was repeated elsewhere on the grid.

Areas of difficulty for Algebra One students, Algebra One students struggled with:

- Understanding that specific numeric examples are not a proof or justification for a rule.
- Recognizing that the numbers in the crisscross could be described using one variable.
- Using algebra to prove why a rule works.

Implications for Instruction:

Part of learning algebra is gaining the appreciation of symbols to quantify relationships and using symbols and knowledge of the number system to prove conjectures. Students need to gain a deeper understanding of variables and their use to solve problems. Students need to be pushed to use their ability to manipulate symbols to solve problems. Students need more experiences with problem solving, looking for patterns, and making proofs. Discussions about how many variables are needed for given problems give students the opportunity to think more deeply about the purpose of variables and see their usefulness.

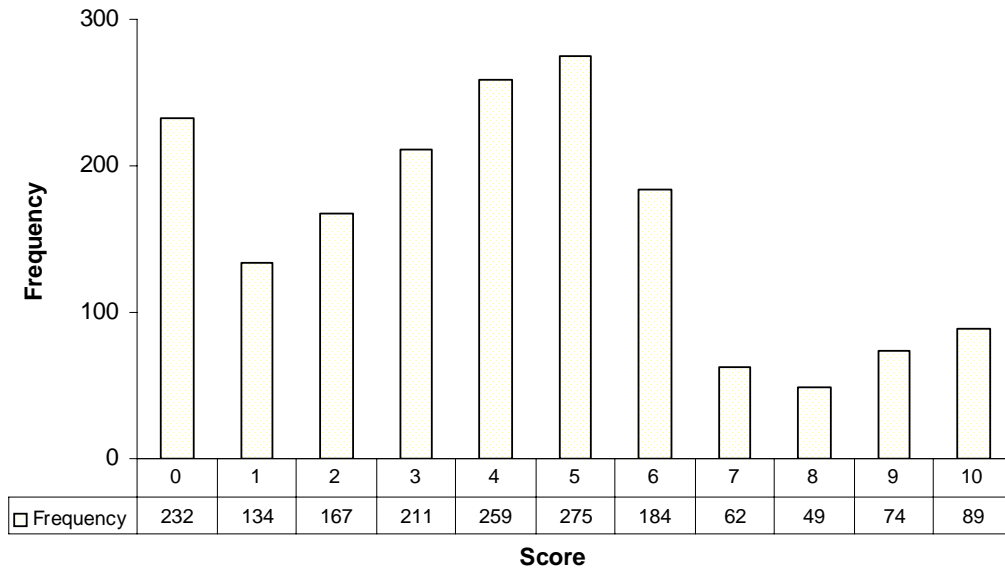
Conference Tables

Student Task	Find and extend patterns in a geometric context. Use inverse relationships to solve problems. Describe a rule or write a formula to explain how to find any number in the pattern.
Core Idea 1 Functions and Relations	Understand patterns, relations, and functions. <ul style="list-style-type: none"> • Generalize patterns using explicitly defined functions. • Understand relations and functions and select, convert flexibly among, and use various representations for them.
Core Idea 3 Algebraic Properties and Representations	Represent and analyze mathematical situations and structures using algebraic symbols. <ul style="list-style-type: none"> • Use symbolic algebra to represent and explain mathematical relationships. • Use symbolic expressions to represent relationships arising from various contexts.

Grade 9 – Conference Tables

Conference Tables

Mean: 4.03, S.D.: 2.76



Score:	0	1	2	3	4	5	6	7	8	9	10
% < =	13.4%	21.1%	30.7%	42.9%	57.8%	73.6%	84.2%	87.8%	90.6%	94.9%	100.0%
% > =	100.0%	86.6%	78.9%	69.3%	57.1%	42.9%	26.4%	15.8%	12.2%	9.4%	5.1%

The maximum score available for this task is 10 points.

The cut score for a level 3 response is 6 points.

Most students (87%) could count the number of tables in the diagram. Many students (about 79 %) could count the number of tables and use the diagram to find the number of people in the diagram for size 3 or count the number of tables in size three and extend the pattern for tables to size 7. More than half the students could find and extend the pattern for tables for size 7 and 13 and find the number of people in size three in the diagram. 16% of the students could correctly identify one or both rules for the geometric patterns. About 13% of the students scored zero points on this task. More than 80% of those students with a score of zero did not attempt the task.

Based on teacher observations, this is what Algebra One students seemed to know and be able to do?

- Count tables and people in a diagram.
- Find recursive rules to describe a pattern.
- Extend a pattern.

Areas of difficulty for Algebra One students, Algebra One students struggled with:

- Working backwards or using inverse relationships.
- Generating functional rules to describe patterns.

Instructional Implications:

Students at this grade level need to be able to go beyond counting and add-on strategies. They should be looking for and generating functional rules when looking at pattern problems. Students need frequent opportunities to generate their own algebraic expressions and use them to solve problems, including inverse operations. Students should be comfortable with use of symbolic notation and the ability to use appropriate variables in generating their rules.

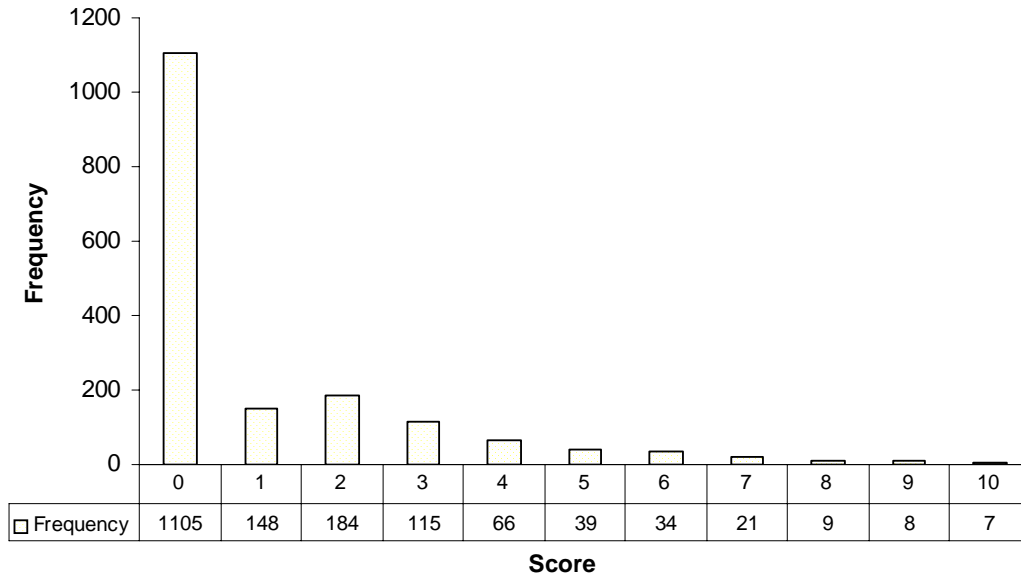
Number Towers

Student Task	Combine numbers and variables using addition or multiplication to fill in blanks in a number tower. Use symbol manipulation to prove why expressions from the number tower are equivalent to given expressions. Find values of unknowns in equations.
Core Idea 3 Algebraic Properties and Representations	Represent and analyze mathematical situations and structures using algebraic symbols. Understand the meaning of equivalent forms of expressions, equations, inequalities, or relations. Write equivalent forms of equations, inequalities and systems of equations and solve them. Use symbolic algebra to represent and explain mathematical relationships.
Core Idea 2 Mathematical Reasoning	Employ forms of mathematical reasoning and proof appropriate to the solution of the problem.

Grade 9 – Number Towers

Number Towers

Mean: 1.09, S.D.: 1.87



Score:	0	1	2	3	4	5	6	7	8	9	10
% <=	63.7%	72.2%	82.8%	89.4%	93.2%	95.4%	97.4%	98.6%	99.1%	99.6%	100.0%
% >=	100.0%	36.3%	27.8%	17.2%	10.6%	6.8%	4.6%	17.2%	1.4%	0.9%	0.4%

The maximum score available for this task is 10 points.

The cut score for a level 3 response is 5 points.

Some students (about 36%) could find one correct solution to satisfy the equations in part 3 or could fill in the correct symbolic expression for the multiplication tower. Less than 10% of the students met standard on this task. Some of these students managed to do this use guess and check strategies to solve for unknowns. More than 60% of the students scored zero on this task. About half of the students who scored zero attempted the task. *Based on teacher observations, this is what Algebra One students seemed to know and be able to do:*

- Use guess and check or substitution to solve for two unknowns in two equations.
- Write symbolic multiplication expressions.
- Simplify a multiplication expression using the distributive property.

Areas of difficulty for Algebra One students, Algebra One students struggled with:

- Writing symbolic addition expressions
- Checking their solutions to see if they fit all the conditions or parameters of the task
- Using symbolic expressions to form equivalent expressions
- Solving two equations for two unknowns
- Using factoring or the quadratic formula to find missing roots

Instructional Implications:

Students at this grade level need frequent opportunities to describe situations in symbolic notation. It is important that they not just use expressions provided for them for symbol manipulation. Students also need to develop their logic and have opportunities to use algebra to prove conjectures or rectify different solutions or formulae. Students should know that given two equations, the common variables stand for the same values. Students should be able to find positive and negative solutions for quadratic equations and have a variety of tools to help them.

Blue Print for Course Two -MAC

Alignment of Tasks to NCTM Content and Process Standards

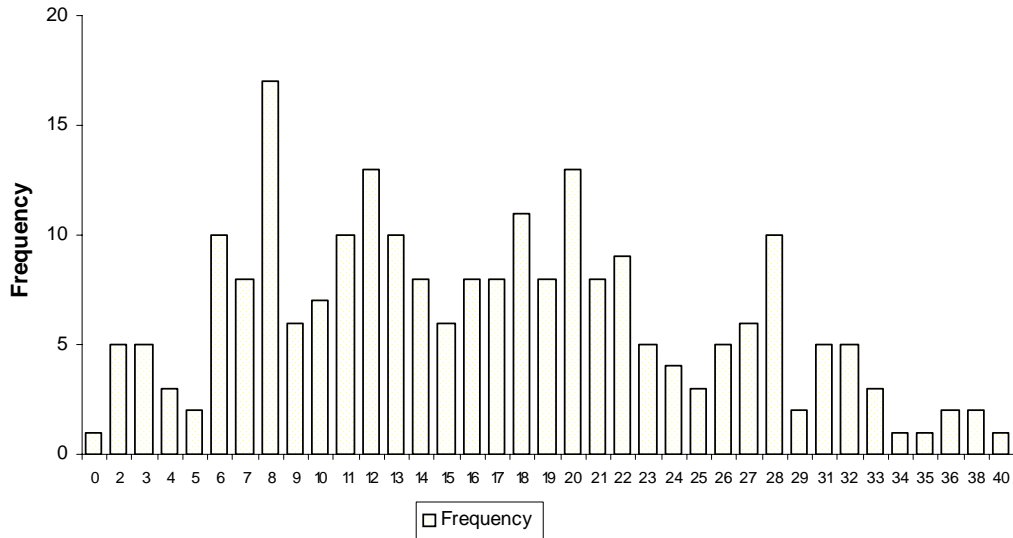
	Taxi Times	Number Patterns	Garden Chair	Lampshade	Rectangle & Square
NUM	S	S		S	S
ALG		P	S		
GEO			P	P	P
MEAS			P	P	S
DATA	P				
PS	*	*		*	
REAS	*	*	*	*	*
COMM	*	*	*	*	*
CONN	*	*	*		
REP		*			*

NUM = Number and Operations, ALG = Algebra, GEO = Geometry, MEAS = Measurement, DATA = Data Analysis and Probability, PS = Problem Solving, REAS = Reasoning and Proof, REP = Representation, CONN = connections, COMM = Communication

- P denotes Primary NCTM Content Standard
 S denotes Secondary NCTM Content Standard
 • denotes NCTM Process Standard

Overall Frequency Distribution by Total Score

Grade 10
Mean=16.66; S.D.=8.70



Level Frequency Distribution Chart and Frequency Distribution

2003 - Numbers of students Course Two: 231
tested:

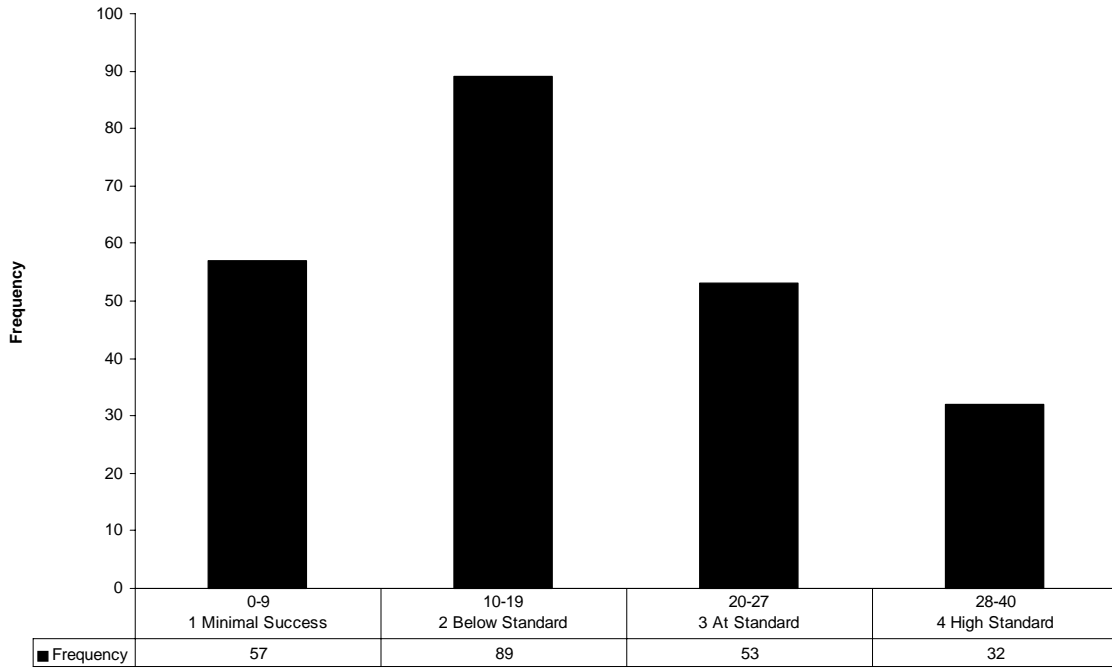
Course Two 2000 - 2001

Level	% at ('00)	% at least ('00)	% at ('01)	% at least ('01)
1	33%	100%	32%	100%
2	45%	67%	52%	68%
3	18%	22%	14%	16%
4	4%	4%	2%	2%

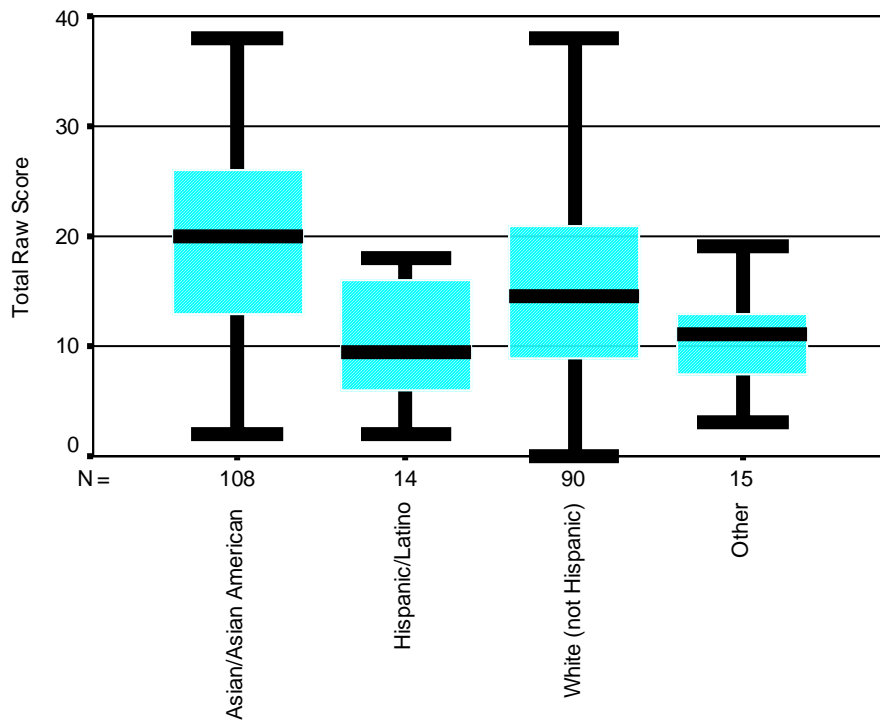
Course Two 2002 - 2003

Level	% at ('02)	% at least ('02)	% at ('03)	% at least ('03)
1	44%	100%	25%	100%
2	40%	56%	39%	75%
3	9%	16%	23%	37%
4	6%	6%	14%	14%

Course Two – Distribution by Performance Levels



Grade 10 – Ethnicity -The following figures show the distribution of raw scores with the median represented as a horizontal bar in the center of the box, the interquartile range (25 percentile to 75 percentile) represented by the box, and the extreme values * within a category lie between the highest and lowest horizontal bars.



*extremes are cases with values more than 3 box lengths from the upper or lower edge of the box.

Grade 10 - Ethnicity

In this section, test scores are compared across different ethnic groups³⁶. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of Asian/Asian American students are significantly higher than all other groups.

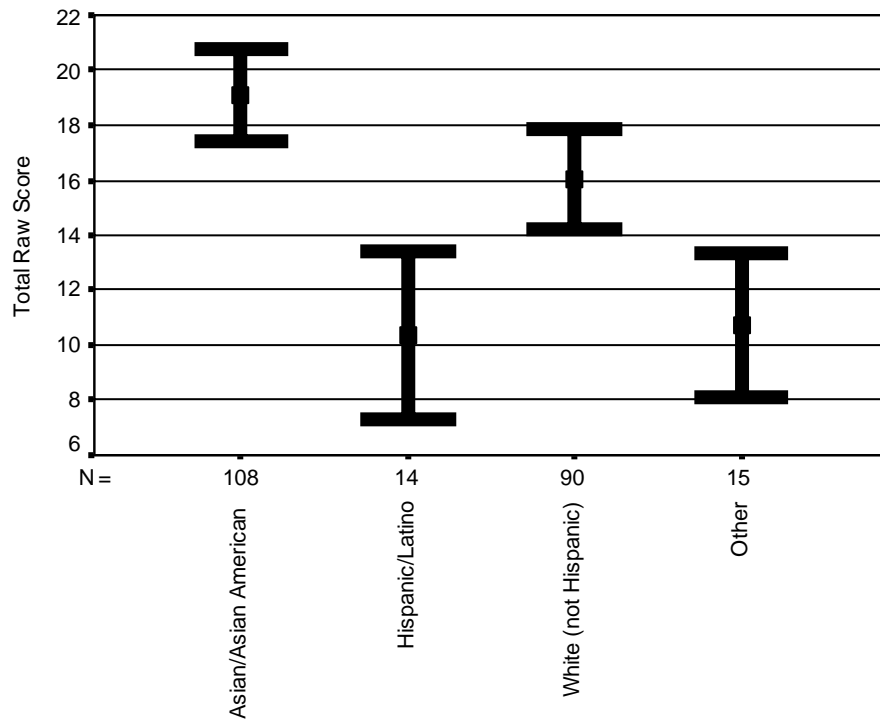
The scores of Hispanic/Latino students are significantly lower than those of Asian/Asian American students. There are no significant differences between Hispanic/Latino students and any other ethnic group.

The scores of White students are significantly lower than those of Asian/Asian American students. There are no significant differences between White students and any other ethnic group.

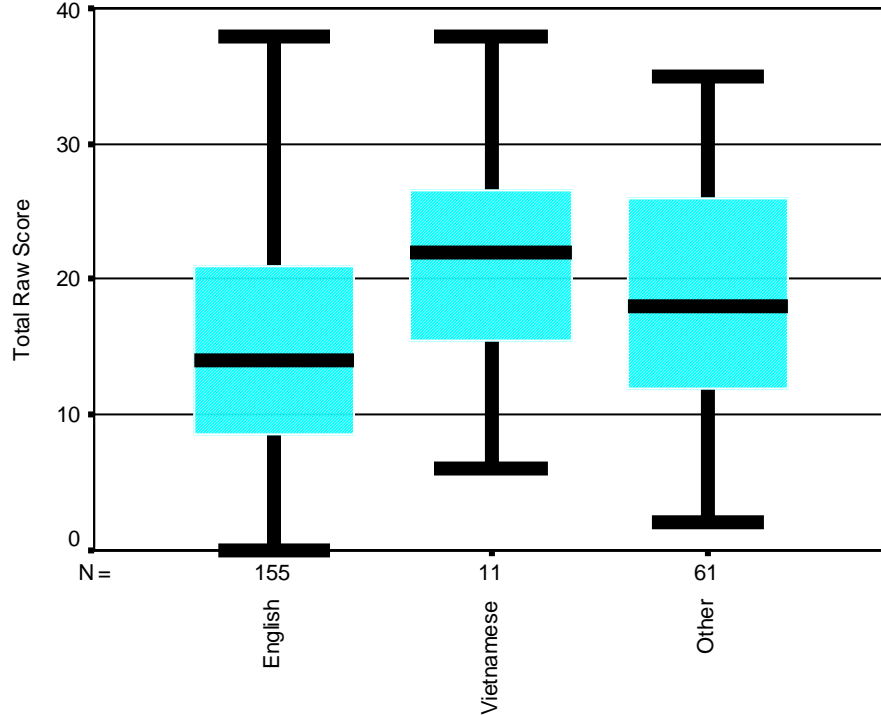
The scores of “Other” students are significantly lower than those of Asian/Asian American students. There are no significant differences between “Other” students and any other ethnic group.

³⁶ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons across the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 10 - Ethnicity



Grade 10 - Home Language

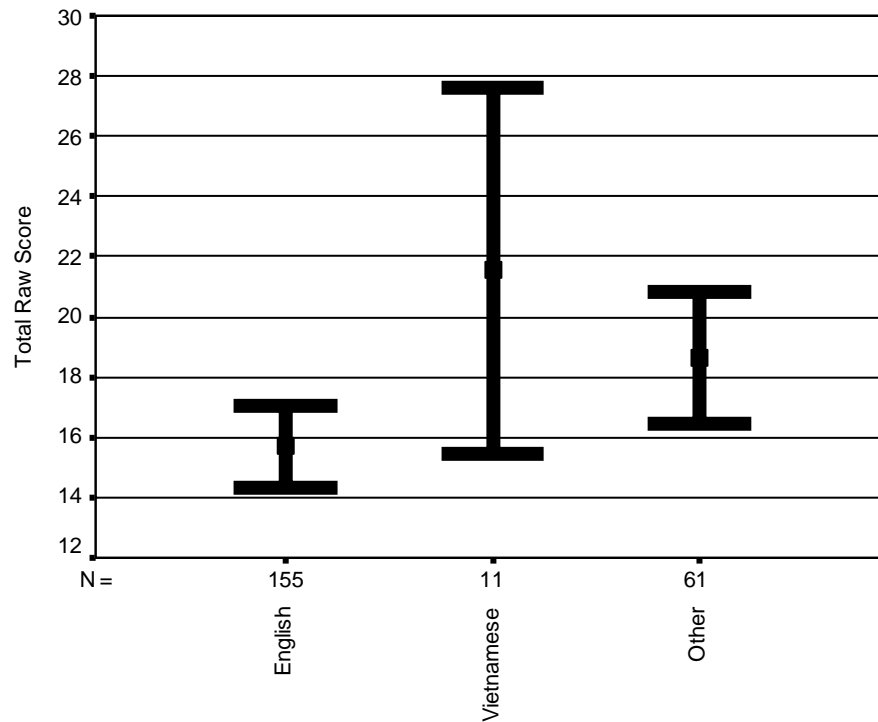


Other includes Albanian, Arabic, Armenian, Assyrian, Burmese, Cantonese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, Filipino, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Korean, Lao, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Russian, Samoan, Serbo-Croatian, Spanish, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

Distribution of sampling means - Grade 10 - Home Language

In this section, test scores are compared across groups of students who speak different languages at home³⁷. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups. For Grade 10, there are no overall significant differences across Home Language categories.

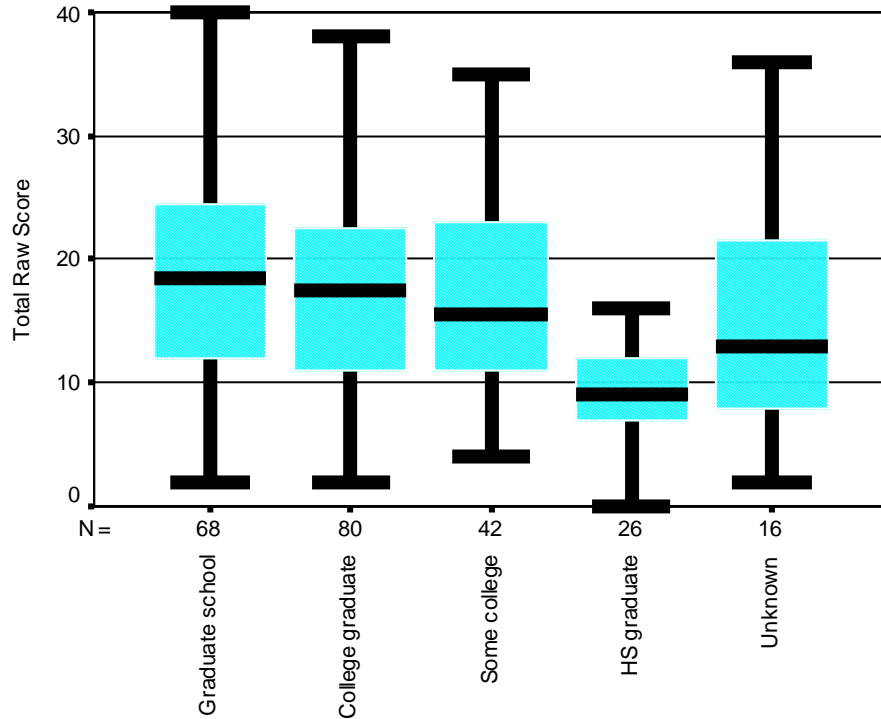
Grade 10 – Home Language



Other includes Albanian, Arabic, Armenian, Assyrian, Burmese, Cantonese, Cebuano, Chaldean, Chaozhou, Croatian, Dutch, Farsi, Filipino, French, German, Greek, Gujarati, Hebrew, Hindi, Hmong, Hungarian, Ilocano, Indonesian, Italian, Japanese, Khmer, Khmu, Korean, Lao, Lahu, Mandarin, Mien, Pashto, Polish, Portuguese, Punjabi, Rumanian, Russian, Samoan, Serbo-Croatian, Spanish, Taiwanese, Thai, Tongan, Turkish, Ukrainian, Urdu, and other non-English languages.

³⁷ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons across the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Grade 10 - Parent Education



Grade 10 - Parent Education

In this section, test scores are compared across groups of different levels of parent education.³⁸ One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

The scores of students whose parents have a graduate degree are significantly higher than those whose parents are High School graduates. There are no significant differences between students whose parents have a graduate degree and any other group.

The scores of students whose parents have a college education are significantly higher than those whose parents are High School graduates. There are no significant differences between students whose parents have a college education and any other group.

The scores of students whose parents have some college education are significantly higher than those whose parents are High School graduates. There are no significant differences between students whose parents have some college education and any other group.

The scores of students whose parents are High School graduates are significantly lower than those whose parents have some college education, a college education, and a

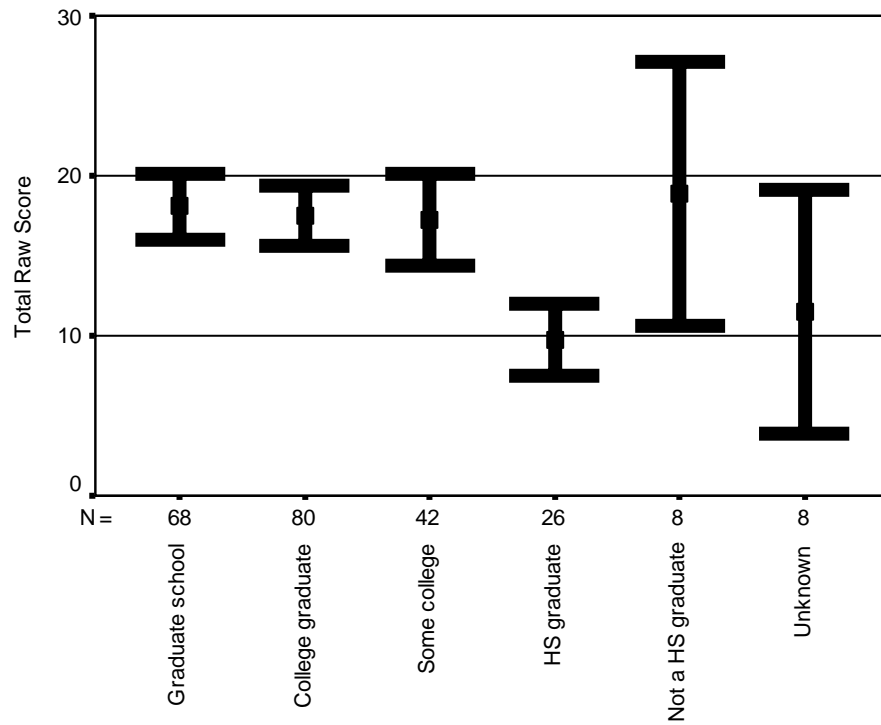
graduate degree. There are no significant differences between students whose parents are High School graduates and any other group.

There are no significant differences between students whose parents are not High School graduates and any other group.

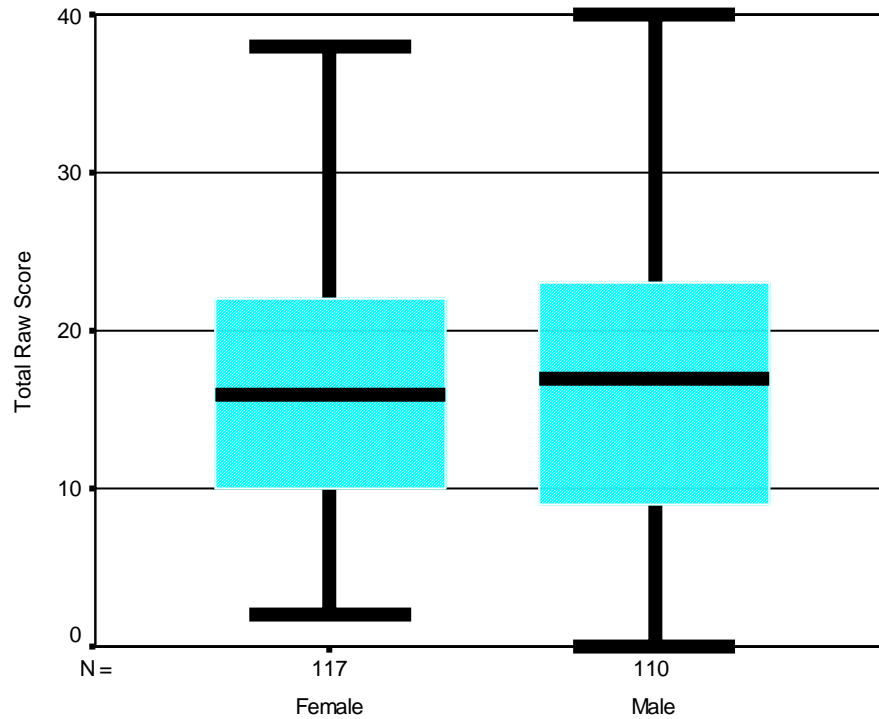
There are no significant differences between “Unknown” students and any other group.

³⁸ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey’s honestly significant difference comparison. All differences were significant at the .05 level.

Distribution of sampling means - Grade 10 - Parent Education



Grade 10 - Gender

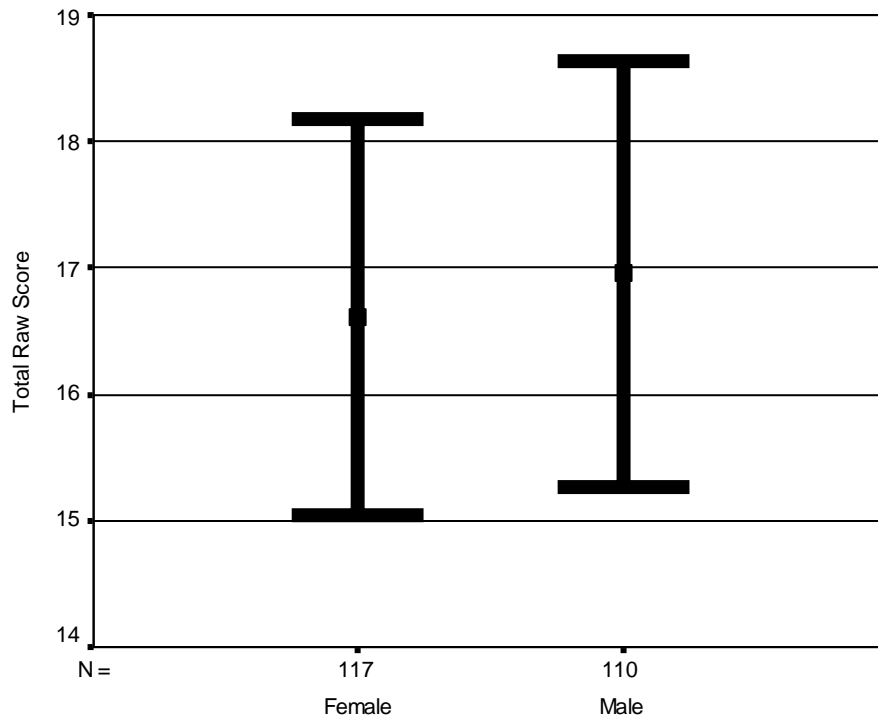


Distribution of sampling means - Grade 10 - Gender

In this section, test scores are compared across gender.³⁹ One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups.

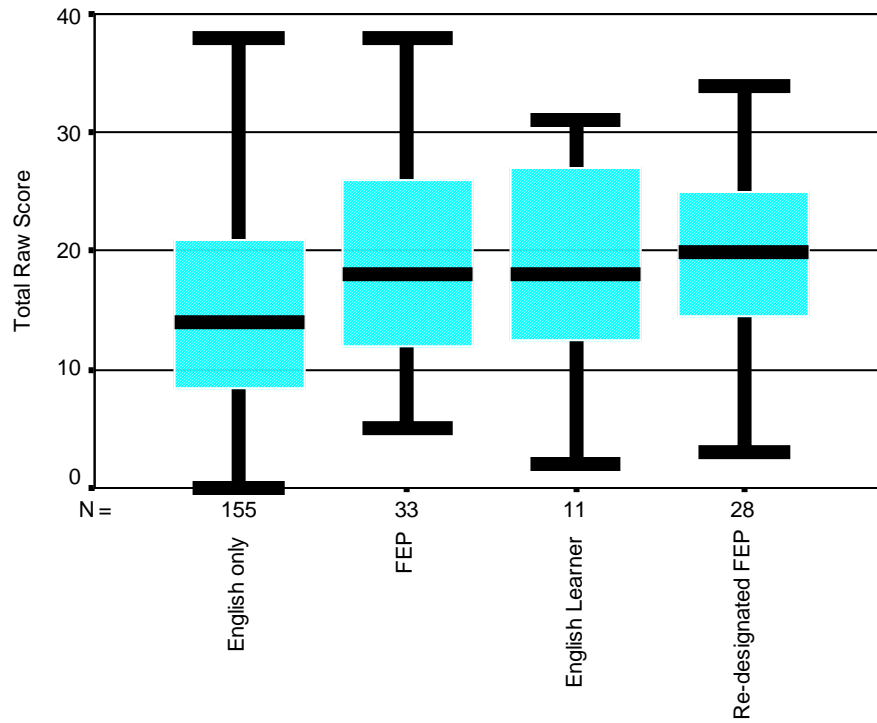
There is no significant difference between the scores of females and males.

Grade 10 – Distribution of sampling means – Gender



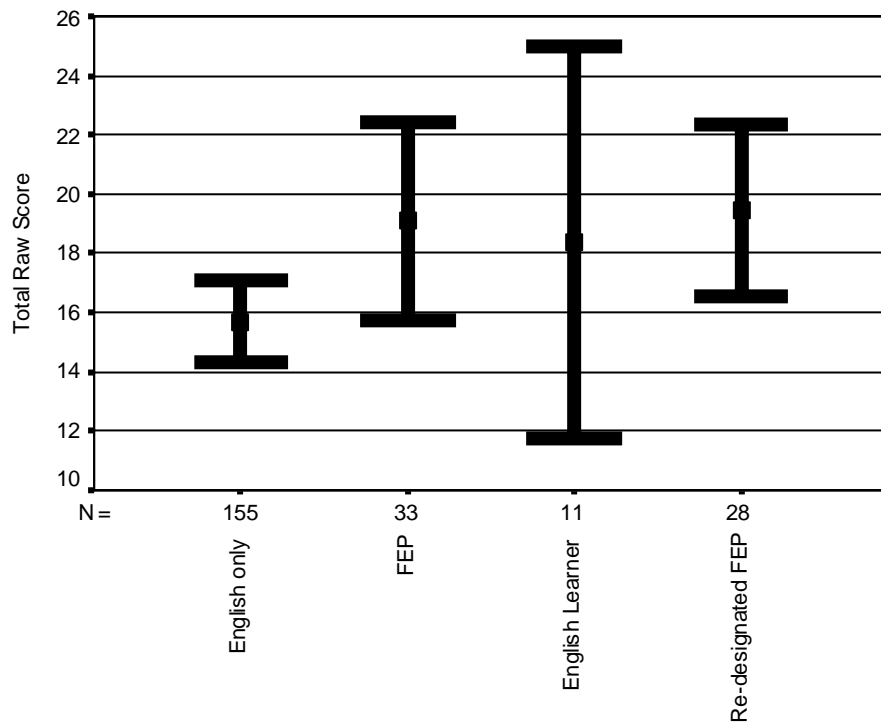
³⁹ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

Language Fluency – Grade 10



Distribution of sampling means- Language Fluency - Grade 10

In this section, test scores are compared across groups of different language fluency.⁴⁰ One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores from group A are below and do not overlap the scores from group B then group A is significantly lower than B. When two scores overlap, then there is no significant difference between the groups. For Grade 10, there are no overall significant differences across language fluency categories.



⁴⁰ Overall comparisons were made using an Analysis of Variance (ANOVA) test. Comparisons between the groups were made using Tukey's honestly significant difference comparison. All differences were significant at the .05 level.

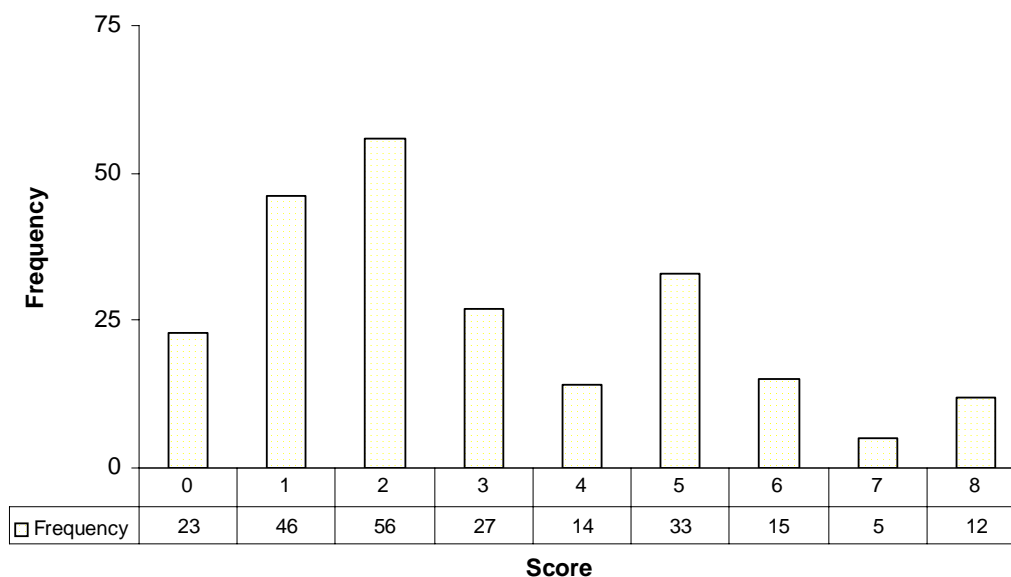
Taxi Times

Student Task	Use a scatterplot to find lowest average speed and average speed for a taxi drive. Draw and use line of best fit to plan fare rates for the taxi driver.
Data Analysis	Select and use appropriate statistical methods to analyze data. Find, use, and interpret measures of center. Understand the relationship between two sets of data (bivariate), display such data in a scatterplot, and describe trends and shape of the plot including correlations (positive, negative, and no) and lines of best fit.
Algebraic Properties and Representations	Recognize and use equivalent graphical and algebraic representations of lines with their geometric characteristics, such as slope.

Frequency Distribution for each Task – Grade 10
Grade 10 – Taxi Times

Taxi Times

Mean: 2.95, S.D.: 2.19



Score:	0	1	2	3	4	5	6	7	8
% < =	10.0%	29.9%	54.1%	65.8%	71.9%	86.1%	92.6%	94.8%	100.0%
% > =	100.0%	90.0%	70.1%	45.9%	34.2%	28.1%	13.9%	7.4%	5.2%

The maximum score available on this task is 8 points.

The cut score for a level 3 response is 4 points.

Most students (about 90%) could draw a line of best fit. Almost half the students could also identify the point on the scatterplot with the lowest average speed and do some correct calculation for using the rate/distance formula. Less than 20% of the students find both of the average speeds and relate speed to fare. 10% of the students scored no points on this task.

Based on teacher observations, this is what Course Two students seemed to know and be able to do:

- Draw a line of best fit
- Find the outlier or lowest average speed on a scatterplot

Areas of difficult for Course Two students:

- Understanding the relationship between slope and speed on a time/ distance graph
- Applying distance formula to numbers in a graph
- Converting from miles per minute to miles per hour
- Understanding the difference in meaning of points lying on the line of best fit and those not on the line
- Understanding average in a graphing situation

Instructional Implications:

1. Distance/Time graphs are excellent examples of problems to focus on the concept of slope in a real context. The average speed is distance divide by time or the change in y's divided by the change in x's, which of course is the slope of the line. If you start with a horizontal ray from the origin it would have zero slope and zero speed. As you rotate the ray counterclockwise the slope increase and thus the average speed increases, so the first point you hit in the graph would have the lowest average speed. As you continue rotating the average speed increase until it is vertical or undefined. This is the powerful idea behind slope that has many applications.
2. Students had trouble identifying the scale of the graph. On the vertical scale the interval distance was 1/4 in length and the horizontal scale interval was 1/2 in length. This contributed to some misreadings of the coordinates and ultimately incorrect calculations. Students should have experience with graph that have different vertical and horizontal scales.
3. Students had trouble converting from miles per minutes to miles per hour. Some work on dimensional analysis might be helpful, i.e. $\text{mile}/\text{min} \times \text{min}/\text{hour} = \text{mile}/\text{hour}$
4. Students had trouble with the concept of a line that best fits the data. Students need experience with scatterplots and using best fit type lines to determine trends and estimate functional relationships between the axes.
5. The last question requires students to attack the problem from an inverse perspective in relationship to the first sets of questions. Students need experience with determining and developing an inverse approach (undoing or working backwards) when confronted with a problem in context.
6. Baker's Choice, by Key Curriculum Press, is a good replacement unit for slope, inequalities, optimizing outcomes. This could be an addition to the Algebra One Program.

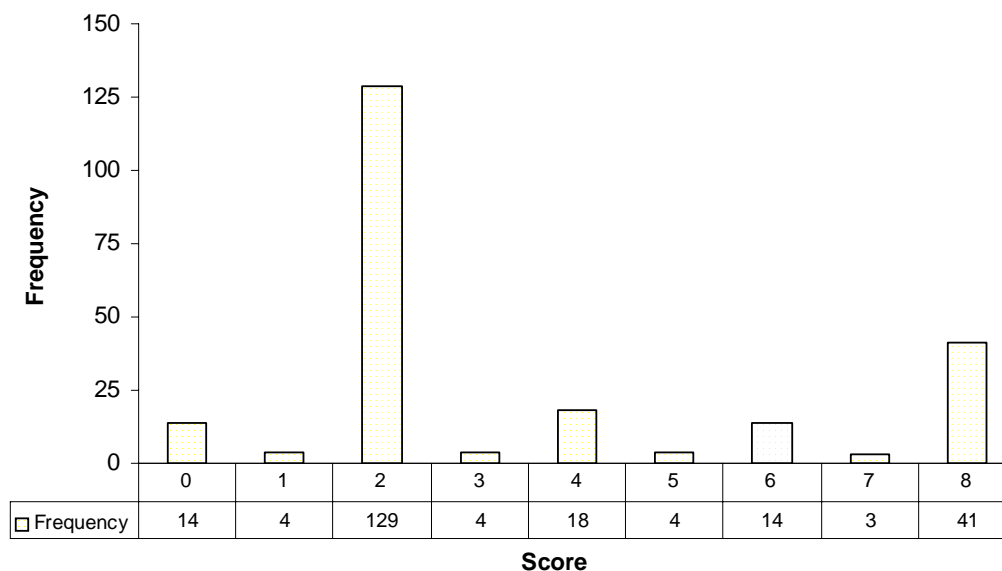
Number Patterns

Student Task	Explore patterns on a hundreds chart. Use algebra to prove why patterns hold true for all cases.
Mathematical Reasoning and Proof	Show mathematical reasoning in solutions in a variety of ways, including words, numbers, symbols, pictures, charts, and models.
Algebraic Properties and Representations	Use symbolic algebra to represent and explain mathematical relationships. Use symbolic expressions to represent relationships arising from various contexts.

Grade 10 – Number Patterns

Number Patterns

Mean: 3.46, S.D.: 2.52



Score:	0	1	2	3	4	5	6	7	8
% ≤	6.1%	7.8%	63.6%	65.4%	73.2%	74.9%	81.0%	82.3%	100.0%
% ≥	100.0%	93.9%	92.2%	36.4%	34.6%	26.8%	25.1%	19.0%	17.7%

The maximum score available on this task is 8 points.

The cut score for a level 3 responses is 4.

Most students (about 90%) were able to do the arithmetic calculations and describe the number pattern. About 50% of the students made some attempt to use algebra, with out 40 % being able to do at least one of the equations needed to show why the pattern would work for all cases. About 18% of the students could meet all the demands of the task. About 6% of the students scored 0 on this task.

Based on teacher observations, this is what Course Two students seemed to know and be able to do:

- Follow directions to complete a set of operations.
- Test those operations on different numeric examples.
- Describe the pattern formed by the calculations.

Areas of difficulty for Course Two students:

- Using algebraic to express relationships of numbers in a sequence.
- Using equations to prove a generalization or rule.
- Squaring binomials.

Implications for Instruction:

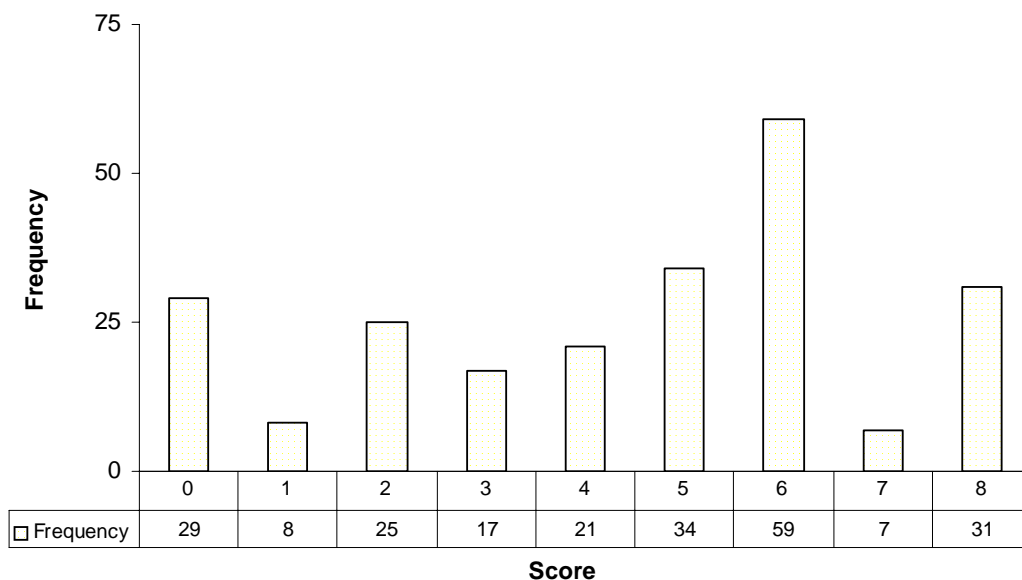
Students at this grade level should be comfortable with the idea of a mathematical proof or justification. They should have many opportunities to explain their thinking, make hypotheses, and test them using both algebra and geometric theorems. Students should have enough problems to solve to enable them to see why proof by example is not enough to complete an argument. There are many interesting problems that seem to have one pattern for the first 5 or 6 cases, but if explored further prove to have a different pattern. Mark Driscoll's Fostering Algebraic Thinking or NCTM's Navigations in Algebra might be good reference sources. As part of their algebra curriculum, they should be given problems which require them to determine the number of variables needed to solve the problem and why using fewer variables is often more efficient or necessary. The use of three variables by many students shows a profound lack of understanding around the concept of variable and how it is used to define relationships.

Garden Chair

Student Task	Students analyze a diagram to determine angle sizes, use Pythagorean Theorem to find the hypotenuse of a right triangle, make a mathematical argument for why two triangles are similar, and use proportional reasoning to find the length of one side of a triangle.
Geometry and Measurement	Analyze characteristics and properties of two-dimensional geometric shapes; develop mathematical arguments about geometric relationships, and apply appropriate techniques, tools, and formulas to determine measurements.
Mathematical Reasoning and Proofs	Use synthetic, coordinate, and /or transformational geometry in direct or indirect proof of geometric relationships. Establish the validity of geometric conjectures using deduction; prove theorems, and critique arguments made by others.

Garden Chair

Mean: 4.39, S.D.: 2.51



Score:	0	1	2	3	4	5	6	7	8
% <=	12.6%	16.0%	26.8%	34.2%	43.3%	58.0%	83.5%	86.6%	100.0%
% >=	100.0%	87.4%	84.0%	73.2%	65.8%	56.7%	42.0%	16.5%	13.4%

The maximum score available on this task is 8 points.

The cut score needed for a level 3 response is 5.

Most students (about 87%) could use Pythagorean theorem to solve for an unknown side. About half the students could find angle X and use Pythagorean theorem. About 17% of the students could also prove similarity and use that property for proportional reasoning. About 13% of the students scored zero points on the task. *Based on teacher observations, this is what Course Two students seemed to know and be able to do:*

- Use Pythagorean Theorem to solve for a hypotenuse.
- Use properties of a triangle and measure of a straight line to find a missing angle.
- Show that vertical angles are equal.

Areas of difficulty for Course Two students:

- Converting 3 feet 6 inches to 3.5 feet.
- Comparing all three angles to make a case for similarity.
- Use similarity and proportional reasoning to solve for a missing side of a triangle.

Implications for Instruction:

Students need to be asked a variety of questions to check for understanding around a concept. So of those questions need to push students beyond recall and procedure, into analysis. For example, students learning about relationships of angles need a variety of complicated diagrams to push them to use multiple definitions or relationships to find the missing value. Students learning a concept, such as similarity, need to see the application of the rules and tests for similarity, as a useful tool for solving problems. Students need to make connections between procedures and definitions to applications. Students at this grade level should frequently be given challenging problems to integrate a variety of skills and procedures, where the steps are not broken down into smaller bits and that require longer chains of reasoning. A good resource for problems might be NCTM's [Navigating Through Geometry](#) or some of Key Curriculum Press's IMP materials.

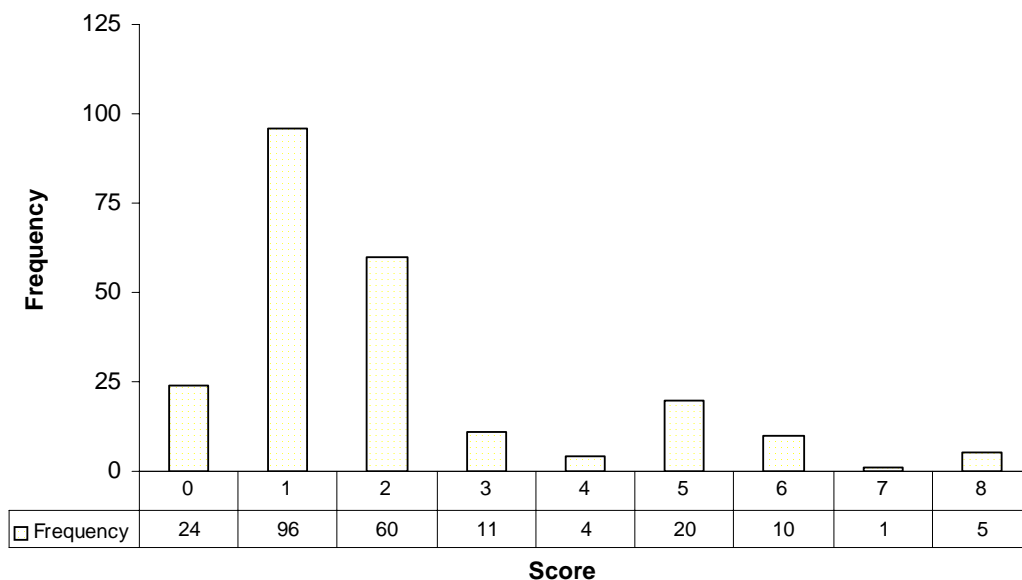
Lampshade

Student Task	Students compare and analyze two diagrams of a lampshade to determine circumference of the shade, circumference and diameter of the pattern, fabric size for making the shade and percentage of fabric used.
Geometry and Measurement	Analyze characteristics and properties of two- and three-dimensional geometric shapes; develop mathematical arguments about geometric relationships; and apply appropriate techniques, tools, and formulas to determine measurements. <ul style="list-style-type: none">• Understand and use formulas for the area, surface area, and volume of geometric figures, including spheres and cylinders.• Draw and construct representations of two- and three-dimensional geometric objects using a variety of tools.

Grade 10 – Lampshade

Lampshade

Mean: 2.04, S.D.: 1.82



Score:	0	1	2	3	4	5	6	7	8
% ≤	10.4%	51.9%	77.9%	82.7%	84.4%	93.1%	97.4%	97.8%	100.0%
% ≥	100.0%	89.6%	48.1%	22.1%	17.3%	15.6%	6.9%	2.6%	2.6%

The maximum score available on this task is 8 points.
The cut score needed for a level 3 response is 3 points.

Most students (about 90%) could use the formula to find the circumference of the lampshade. About half the students could also find the percentage of the fabric used to make the shade. This may have been based on errors in previous calculations. About 16% of the students could find the circumference of the lampshade, the diameter of the pattern and the small circle, find the area of the total fabric, and make some correct calculation of percentages. More than 10% of the students scored no points on this task.

Based on teacher observations, this is what Course Two students seemed to be know and be able to do:

- Use the formula to find circumference of a circle.
- Calculate percentages.

Areas of difficulty for Course Two students:

- Analyzing and synthesizing diagrams
- Keeping track of what information results from their calculations
- Working multi-step problems
- Using inverse operations on formulas

Implications for Instruction:

Students in middle grades frequently compute area, circumference and diameter of circles and interpret simple diagrams. As students move through the grades the types of thinking and reasoning need to get progressively more complex. What opportunities are provided in the curriculum to go beyond the reasoning of simple calculations and push students to this level? Students should frequently be given complex diagrams to interpret. New Elementary Mathematics Course 2 and 3, by Dr. Wong Khoon Yoong and Sin Kwai Meng, Pan Pacific Publications, might be a good source of problems. Students also need some strategy to help them make sense of problem situations. Students who were successful seemed to make their own diagrams. Perhaps labeling answers as they work through multi-step problems would also be a helpful tool; as students seemed to lose track of where they were in the process. A study in Focus Magazine (Volume 10, Number 1, 2003) shows that high school teachers who worked at not breaking down problems into smaller parts for students quadrupled test scores. It is important for students to struggle with organizing their information and doing longer chains of reasoning.

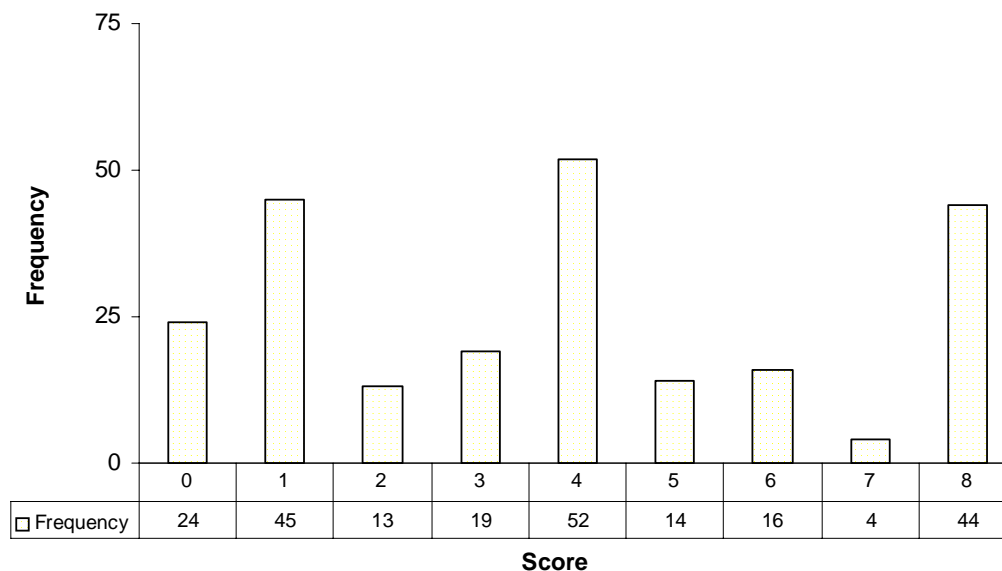
Rectangle and Square

Student Task	Find missing lengths in a diagram using Pythagorean Theorem and similarity. Use understanding of area and properties of a square to solve a spatial puzzle.
Geometry and Measurement.	<ul style="list-style-type: none">• Analyze characteristics and properties of two-dimensional geometric shapes; develop mathematical arguments about geometric relationships; apply transformations a; and apply appropriate techniques, tools, and formulas to determine measurements.• Understand relationships among the angles, side lengths, perimeter, and areas, and volumes of similar figures.• Create and critique inductive and deductive arguments concerning geometric ideas and relationships, such as congruence, similarity, and the Pythagorean relationship.

Grade 10 – Rectangle and Square

Rectangle and Square

Mean: 3.82, S.D.: 2.69



Score:	0	1	2	3	4	5	6	7	8
% <=	10.4%	29.9%	35.5%	43.7%	66.2%	72.3%	79.2%	81.0%	100.0%
% >=	100.0%	89.6%	70.1%	64.5%	56.3%	33.8%	33.8%	20.8%	19.0%

The maximum score available for this task is 8 points.

The cut score needed for a level 3 response is 4.

Most students (about 90%) could use Pythagorean theorem and solve for side BF. More than half the students could use similarity and Pythagorean theorem to find all the missing dimensions in the diagram. Almost 21% of the students could use information about the properties of a square to solve the puzzle and explain their reasoning in part 2. About 10% of the students scored zero points on this task.

Based on teacher observations, this is what Course Two students seemed to know and be able to do:

- Use Pythagorean theorem
- Use properties of similarity

Areas of difficulty for Course Two students:

- Drawing and rotating geometric shapes, spatial visualization
- Using geometric properties to reason or to make an argument about a square

Implications for Instruction:

Students need to have many experiences working with similar figures, particularly when the shapes have different orientations. Students need to be able to understand and represent translations, reflections, and rotations of objects in the plane by using sketches. While spatial visualization is difficult for students, facility with this can be improved with practice. Some research says that lack of spatial visualization skills prevents students from being successful in more advanced math courses. Students at this grade level need many opportunities to use geometric relationships to solve problems and explain their thinking.

Appendix 1

Star Data Comparisons

Comparison of CAT/6 Math Percentile ranks to MARS levels

Comparison of CST Math Performance Levels to MARS levels

Correlations of STAR scores with MARS Scores

CST Cluster Level Comparison on MAC tasks

Comparison of CAT/6 Math percentile ranks to MARS levels

Grade 3

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	5.6%	5.1%	4.3%	0.8%
	Below Average (2nd)	1.8%	5.4%	8.7%	3.3%
	Above Average (3rd)	0.6%	2.8%	10.6%	9.9%
	Top (4th)	0.2%	1.1%	9.3%	30.1%

Grade 4

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	5.8%	6.2%	5.0%	0.7%
	Below Average (2nd)	1.2%	4.8%	10.2%	3.2%
	Above Average (3rd)	0.2%	1.9%	10.5%	7.6%
	Top (4th)	0.0%	0.8%	10.8%	31.0%

Grade 5

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	7.4%	7.6%	2.4%	0.3%
	Below Average (2nd)	1.8%	9.2%	6.7%	1.5%
	Above Average (3rd)	0.5%	6.9%	16.2%	10.1%
	Top (4th)	0.1%	0.9%	7.5%	20.8%

Grade 6

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	11.2%	5.1%	1.0%	0.1%
	Below Average (2nd)	5.4%	8.2%	2.9%	0.1%
	Above Average (3rd)	3.3%	10.3%	10.8%	1.6%
	Top (4th)	0.8%	4.7%	18.1%	16.4%

Grade 7

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	13.8%	4.2%	0.2%	0.0%
	Below Average (2nd)	12.1%	13.0%	1.1%	0.1%
	Above Average (3rd)	4.2%	14.1%	4.5%	1.0%
	Top (4th)	0.8%	7.9%	10.5%	12.5%

Grade 8

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	14.5%	3.1%	0.4%	0.1%
	Below Average (2nd)	9.6%	7.3%	1.8%	0.4%
	Above Average (3rd)	5.4%	11.5%	7.9%	3.5%
	Top (4th)	0.9%	4.5%	8.9%	20.2%

Grade 9

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	26.4%	3.1%	0.2%	0.0%
	Below Average (2nd)	23.2%	8.7%	0.6%	0.2%
	Above Average (3rd)	11.2%	11.6%	1.6%	0.1%
	Top (4th)	2.3%	7.1%	3.4%	0.5%

Grade 10

		MARS LEVEL			
		1	2	3	4
CAT/6 Math Percentile Rank	Bottom (1st quartile)	2.3%	0.5%	0.0%	0.0%
	Below Average (2nd)	6.4%	1.8%	0.0%	0.0%
	Above Average (3rd)	7.3%	4.1%	0.9%	0.0%
	Top (4th)	8.6%	30.9%	22.7%	14.5%

Comparison of CST Math performance levels to MARS levels

Grade 3

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	2.3%	0.9%	0.4%	0.0%
	Below Basic	4.2%	5.8%	5.3%	1.0%
	Basic	1.3%	5.4%	10.7%	4.3%
	Proficient	0.4%	2.1%	12.4%	14.5%
	Advanced	0.0%	0.3%	4.2%	24.4%

Grade 4

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	2.2%	0.9%	0.5%	0.0%
	Below Basic	3.7%	5.7%	5.6%	0.8%
	Basic	1.2%	5.5%	13.3%	4.7%
	Proficient	0.1%	1.5%	13.6%	15.6%
	Advanced	0.0%	0.1%	3.5%	21.3%

Grade 5

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	4.2%	2.8%	0.8%	0.1%
	Below Basic	4.3%	10.8%	5.2%	1.0%
	Basic	1.1%	8.0%	11.7%	4.0%
	Proficient	0.2%	2.9%	12.6%	14.4%
	Advanced	0.1%	0.2%	2.6%	13.1%

Grade 6

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	2.7%	0.6%	0.1%	0.0%
	Below Basic	11.1%	7.0%	1.4%	0.1%
	Basic	5.8%	13.0%	7.3%	0.4%
	Proficient	1.2%	7.1%	18.5%	5.6%
	Advanced	0.1%	0.5%	5.6%	12.1%

Grade 7

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	5.1%	1.0%	0.0%	0.0%
	Below Basic	13.9%	6.6%	0.3%	0.0%
	Basic	10.2%	18.7%	2.6%	0.4%
	Proficient	1.7%	12.1%	10.5%	5.2%
	Advanced	0.0%	0.8%	3.0%	7.9%

Grade 8

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	7.0%	1.7%	0.3%	0.0%
	Below Basic	14.1%	7.0%	2.2%	0.5%
	Basic	7.8%	10.7%	6.3%	2.9%
	Proficient	1.4%	6.7%	9.0%	12.3%
	Advanced	0.0%	0.3%	1.4%	8.4%

Grade 9

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	17.1%	3.1%	0.2%	0.0%
	Below Basic	35.8%	10.4%	1.1%	0.1%
	Basic	9.4%	12.1%	1.1%	0.2%
	Proficient	0.5%	4.7%	2.7%	0.2%
	Advanced	0.2%	0.4%	0.5%	0.3%

Grade 10

		MARS LEVEL			
		1	2	3	4
CST Math Performance Levels	Far Below Basic	0.9%	0.0%	0.5%	0.0%
	Below Basic	14.2%	4.6%	0.0%	0.0%
	Basic	8.2%	9.1%	0.0%	0.0%
	Proficient	1.4%	16.4%	9.6%	1.4%
	Advanced	0.0%	7.3%	13.7%	12.8%

Grade 3
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho	1.000	.695**	1.000**	.695**	.838**	.785**	.813**	.763**	.681**	.646**
	Correlation
	N	14180	14158	14180	14158	14140	14128	14140	14128	14180	14180
CAT/6 Reading raw score	Spearman's rho	.695**	1.000	.695**	1.000**	.698**	.839**	.675**	.815**	.610**	.580**
	Correlation
	N	14158	14187	14159	14187	14147	14143	14147	14143	14187	14187
CAT/6 Math percentile rank	Spearman's rho	1.000**	.695**	1.000	.696**	.838**	.785**	.813**	.763**	.681**	.646**
	Correlation
	N	14180	14159	14182	14160	14140	14128	14140	14128	14182	14182
CAT/6 Reading percentile rank	Spearman's rho	.695**	1.000**	.696**	1.000	.698**	.838**	.675**	.815**	.610**	.579**
	Correlation
	N	14158	14187	14160	14188	14147	14143	14147	14143	14188	14188
CST Math raw score	Spearman's rho	.838**	.698**	.838**	.698**	1.000	.812**	.968**	.790**	.726**	.686**
	Correlation
	N	14140	14147	14140	14147	14178	14147	14178	14147	14178	14178
CST English Language raw score	Spearman's rho	.785**	.839**	.785**	.838**	.812**	1.000	.786**	.973**	.697**	.659**
	Correlation
	N	14128	14143	14128	14143	14147	14186	14147	14183	14186	14186
CST Math Performance Standards	Spearman's rho	.813**	.675**	.813**	.675**	.968**	.786**	1.000	.765**	.705**	.673**
	Correlation
	N	14140	14147	14140	14147	14178	14147	14178	14147	14178	14178
CST English Performance Standards	Spearman's rho	.763**	.815**	.763**	.815**	.790**	.973**	.765**	1.000	.678**	.643**
	Correlation
	N	14128	14143	14128	14143	14147	14183	14147	14183	14183	14183
MARS raw score	Spearman's rho	.681**	.610**	.681**	.610**	.726**	.697**	.705**	.678**	1.000	.937**
	Correlation
	N	14180	14187	14182	14188	14178	14186	14178	14183	14568	14568
MARS Performance Level	Spearman's rho	.646**	.580**	.646**	.579**	.686**	.659**	.673**	.643**	.937**	1.000
	Correlation
	N	14180	14187	14182	14188	14178	14186	14178	14183	14568	14568

** . Correlation is significant at the .01 level (2-tailed).

Grade 4
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho Correlation	1.000	.703**	1.000**	.703**	.825**	.762**	.803**	.733**	.722**	.675**
	N	9086	9074	9086	9074	9076	9075	9071	9063	9086	9086
CAT/6 Reading raw score	Spearman's rho Correlation	.703**	1.000	.703**	1.000**	.710**	.869**	.691**	.838**	.634**	.591**
	N	9074	9082	9074	9082	9071	9070	9066	9063	9082	9082
CAT/6 Math percentile rank	Spearman's rho Correlation	1.000**	.703**	1.000	.704**	.825**	.763**	.803**	.732**	.721**	.675**
	N	9086	1.000	1.000	9087	9076	9087	9071	9063	9099	9099
CAT/6 Reading percentile rank	Spearman's rho Correlation	.703**	1.000**	.704**	1.000	.710**	.870**	.691**	.838**	.633**	.591**
	N	9074	9082	9087	9095	9071	9082	9066	9063	9095	9095
CST Math raw score	Spearman's rho Correlation	.825**	.710**	.825**	.710**	1.000	.789**	.968**	.757**	.744**	.695**
	N	9076	9071	9076	9071	9087	9078	9082	9068	9087	9087
CST English Language raw score	Spearman's rho Correlation	.762**	.869**	.763**	.870**	.789**	1.000	.767**	.965**	.683**	.637**
	N	9075	9070	9087	9082	9078	9179	9073	9077	9179	9179
CST Math Performance Standards	Spearman's rho Correlation	.803**	.691**	.803**	.691**	.968**	.767**	1.000	.737**	.723**	.678**
	N	9071	9066	9071	9066	9082	9073	9082	9068	9082	9082
CST English Performance Standards	Spearman's rho Correlation	.733**	.838**	.732**	.838**	.757**	.965**	.737**	1.000	.671**	.627**
	N	9063	9063	9063	9063	9068	9077	9068	9077	9077	9077
MARS raw score	Spearman's rho Correlation	.722**	.634**	.721**	.633**	.744**	.683**	.723**	.671**	1.000	.936**
	N	9086	9082	9099	9095	9087	9179	9082	9077	9370	9370
MARS Performance Level	Spearman's rho Correlation	.675**	.591**	.675**	.591**	.695**	.637**	.678**	.627**	.936**	1.000
	N	9086	9082	9099	9095	9087	9179	9082	9077	9370	9370

** . Correlation is significant at the .01 level (2-tailed).

Grade 5
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho	1.000	.712**	1.000**	.712**	.846**	.771**	.821**	.746**	.747**	.714**
	Correlation
	N	13976	13956	13976	13956	13957	13955	13957	13955	13976	13976
CAT/6 Reading raw score	Spearman's rho	.712**	1.000	.711**	1.000**	.699**	.854**	.679**	.823**	.644**	.613**
	Correlation
	N	13956	13968	13956	13968	13952	13955	13952	13955	13968	13968
CAT/6 Math percentile rank	Spearman's rho	1.000**	.711**	1.000	.712**	.846**	.771**	.821**	.746**	.746**	.713**
	Correlation
	N	13976	13956	13980	13960	13957	13955	13957	13955	13980	13980
CAT/6 Reading percentile rank	Spearman's rho	.712**	1.000**	.712**	1.000	.699**	.854**	.679**	.823**	.643**	.613**
	Correlation
	N	13956	13968	13960	13972	13952	13955	13952	13955	13972	13972
CST Math raw score	Spearman's rho	.846**	.699**	.846**	.699**	1.000	.775**	.971**	.749**	.746**	.712**
	Correlation
	N	13957	13952	13957	13952	13971	13958	13971	13958	13971	13971
CST English Language raw score	Spearman's rho	.771**	.854**	.771**	.854**	.775**	1.000	.752**	.965**	.705**	.672**
	Correlation
	N	13955	13955	13955	13955	13958	13978	13958	13977	13978	13978
CST Math Performance Standards	Spearman's rho	.821**	.679**	.821**	.679**	.971**	.752**	1.000	.728**	.725**	.693**
	Correlation
	N	13957	13952	13957	13952	13971	13958	13971	13958	13971	13971
CST English Performance Standards	Spearman's rho	.746**	.823**	.746**	.823**	.749**	.965**	.728**	1.000	.680**	.649**
	Correlation
	N	13955	13955	13955	13955	13958	13977	13958	13977	13977	13977
MARS raw score	Spearman's rho	.747**	.644**	.746**	.643**	.746**	.705**	.725**	.680**	1.000	.957**
	Correlation
	N	13976	13968	13980	13972	13971	13978	13971	13977	14361	14361
MARS Performance Level	Spearman's rho	.714**	.613**	.713**	.613**	.712**	.672**	.693**	.649**	.957**	1.000
	Correlation
	N	13976	13968	13980	13972	13971	13978	13971	13977	14361	14361

** . Correlation is significant at the .01 level (2-tailed).

Grade 6
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho Correlation	1.000	.736**	1.000**	.736**	.871**	.776**	.843**	.750**	.753**	.727**
	N	9541	9515	9541	9515	9507	9519	9507	9519	9541	9541
CAT/6 Reading raw score	Spearman's rho Correlation	.736**	1.000	.736**	1.000**	.762**	.859**	.737**	.829**	.685**	.662**
	N	9515	9560	9515	9560	9526	9540	9526	9540	9560	9560
CAT/6 Math percentile rank	Spearman's rho Correlation	1.000**	.736**	1.000	.736**	.871**	.776**	.844**	.750**	.753**	.727**
	N	9541	9515	9541	9515	9507	9519	9507	9519	9541	9541
CAT/6 Reading percentile rank	Spearman's rho Correlation	.736**	1.000**	.736**	1.000	.762**	.859**	.738**	.829**	.685**	.662**
	N	9515	9560	9515	9560	9526	9540	9526	9540	9560	9560
CST Math raw score	Spearman's rho Correlation	.871**	.762**	.871**	.762**	1.000	.813**	.967**	.784**	.801**	.773**
	N	9507	9526	9507	9526	9555	9534	9555	9534	9555	9555
CST English Language raw score	Spearman's rho Correlation	.776**	.859**	.776**	.859**	.813**	1.000	.784**	.966**	.736**	.711**
	N	9519	9540	9519	9540	9534	9572	9534	9572	9572	9572
CST Math Performance Standards	Spearman's rho Correlation	.843**	.737**	.844**	.738**	.967**	.784**	1.000	.756**	.772**	.746**
	N	9507	9526	9507	9526	9555	9534	9555	9534	9555	9555
CST English Performance Standards	Spearman's rho Correlation	.750**	.829**	.750**	.829**	.784**	.966**	.756**	1.000	.710**	.688**
	N	9519	9540	9519	9540	9534	9572	9534	9572	9572	9572
MARS raw score	Spearman's rho Correlation	.753**	.685**	.753**	.685**	.801**	.736**	.772**	.710**	1.000	.964**
	N	9541	9560	9541	9560	9555	9572	9555	9572	9837	9837
MARS Performance Level	Spearman's rho Correlation	.727**	.662**	.727**	.662**	.773**	.711**	.746**	.688**	.964**	1.000
	N	9541	9560	9541	9560	9555	9572	9555	9572	9837	9837

** . Correlation is significant at the .01 level (2-tailed).

Grade 7
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho Correlation	1.000	.738**	1.000**	.739**	.878**	.793**	.849**	.766**	.771**	.736**
	N	11342	11313	11342	11314	11304	11318	11302	11310	11342	11342
CAT/6 Reading raw score	Spearman's rho Correlation	.738**	1.000	.738**	1.000**	.733**	.863**	.710**	.834**	.649**	.617**
	N	11313	11364	11314	11364	11323	11341	11321	11332	11364	11364
CAT/6 Math percentile rank	Spearman's rho Correlation	1.000**	.738**	1.000	1.000**	.878**	.793**	.849**	.766**	.771**	.735**
	N	11342	11314	11345	11317	11304	11321	11302	11312	11345	11345
CAT/6 Reading percentile rank	Spearman's rho Correlation	.739**	1.000**	.739**	1.000	.733**	.863**	.711**	.835**	.649**	.617**
	N	11314	11364	11317	11367	11324	11344	11322	11334	11367	11367
CST Math raw score	Spearman's rho Correlation	.878**	.733**	.878**	.733**	1.000	.794**	.965**	.766**	.784**	.751**
	N	11304	11323	11304	11324	11368	11342	11366	11333	11368	11368
CST English Language raw score	Spearman's rho Correlation	.793**	.863**	.793**	.863**	.794**	1.000	.770**	.964**	.708**	.675**
	N	11318	11341	11321	11344	11342	11455	11340	11385	11455	11455
CST Math Performance Standards	Spearman's rho Correlation	.849**	.710**	.849**	.711**	.965**	.770**	1.000	.744**	.758**	.724**
	N	11302	11321	11302	11322	11366	11340	11366	11333	11366	11366
CST English Performance Standards	Spearman's rho Correlation	.766**	.834**	.766**	.835**	.766**	.964**	.744**	1.000	.686**	.655**
	N	11310	11332	11312	11334	11333	11385	11333	11385	11385	11385
MARS raw score	Spearman's rho Correlation	.771**	.649**	.771**	.649**	.784**	.708**	.758**	.686**	1.000	.951**
	N	11342	11364	11345	11367	11368	11455	11366	11385	11696	11696
MARS Performance Level	Spearman's rho Correlation	.736**	.617**	.735**	.617**	.751**	.675**	.724**	.655**	.951**	1.000
	N	11342	11364	11345	11367	11368	11455	11366	11385	11696	11696

** . Correlation is significant at the .01 level (2-tailed).

Grade 8
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho	1.000	.676**	1.000**	.676**	.700**	.753**	.758**	.727**	.789**	.761**
	Correlation
	N	7936	7895	7936	7897	7832	7907	7832	7907	7936	7936
CAT/6 Reading raw score	Spearman's rho	.676**	1.000	.675**	1.000**	.533**	.832**	.581**	.803**	.637**	.606**
	Correlation
	N	7895	7945	7899	7945	7841	7920	7841	7920	7945	7945
CAT/6 Math percentile rank	Spearman's rho	1.000**	.675**	1.000	.676**	.699**	.752**	.757**	.726**	.786**	.758**
	Correlation
	N	7936	7899	7947	7908	7835	7911	7837	7912	7947	7947
CAT/6 Reading percentile rank	Spearman's rho	.676**	1.000**	.676**	1.000	.534**	.832**	.581**	.804**	.636**	.606**
	Correlation
	N	7897	7945	7908	7954	7842	7923	7844	7924	7954	7954
CST Math raw score	Spearman's rho	.700**	.533**	.699**	.534**	1.000	.621**	.948**	.600**	.666**	.639**
	Correlation
	N	7832	7841	7835	7842	7887	7862	7887	7862	7887	7887
CST English Language raw score	Spearman's rho	.753**	.832**	.752**	.832**	.621**	1.000	.674**	.964**	.715**	.682**
	Correlation
	N	7907	7920	7911	7923	7862	7973	7863	7973	7973	7973
CST Math Performance Standards	Spearman's rho	.758**	.581**	.757**	.581**	.948**	.674**	1.000	.651**	.721**	.694**
	Correlation
	N	7832	7841	7837	7844	7887	7863	7889	7864	7889	7889
CST English Performance Standards	Spearman's rho	.727**	.803**	.726**	.804**	.600**	.964**	.651**	1.000	.693**	.661**
	Correlation
	N	7907	7920	7912	7924	7862	7973	7864	7974	7974	7974
MARS raw score	Spearman's rho	.789**	.637**	.786**	.636**	.666**	.715**	.721**	.693**	1.000	.966**
	Correlation
	N	7936	7945	7947	7954	7887	7973	7889	7974	8159	8159
MARS Performance Level	Spearman's rho	.761**	.606**	.758**	.606**	.639**	.682**	.694**	.661**	.966**	1.000
	Correlation
	N	7936	7945	7947	7954	7887	7973	7889	7974	8159	8159

** . Correlation is significant at the .01 level (2-tailed).

Grade 9
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho	1.000	.576**	.994**	.569**	.633**	.610**	.616**	.592**	.594**	.529**
	Correlation
	N	1326	1300	1326	1300	1269	1317	1269	1317	1326	1326
CAT/6 Reading raw score	Spearman's rho	.576**	1.000	.564**	.999**	.442**	.793**	.432**	.766**	.445**	.375**
	Correlation
	N	1300	1337	1300	1337	1278	1320	1278	1320	1337	1337
CAT/6 Math percentile rank	Spearman's rho	.994**	.564**	1.000	.561**	.633**	.600**	.616**	.583**	.591**	.525**
	Correlation
	N	1326	1300	1326	1300	1269	1317	1269	1317	1326	1326
CAT/6 Reading percentile rank	Spearman's rho	.569**	.999**	.561**	1.000	.440**	.792**	.431**	.764**	.441**	.371**
	Correlation
	N	1300	1337	1300	1337	1278	1320	1278	1320	1337	1337
CST Math raw score	Spearman's rho	.633**	.442**	.633**	.440**	1.000	.496**	.927**	.484**	.502**	.465**
	Correlation
	N	1269	1278	1269	1278	1310	1294	1310	1294	1310	1310
CST English Language raw score	Spearman's rho	.610**	.793**	.600**	.792**	.496**	1.000	.485**	.956**	.522**	.431**
	Correlation
	N	1317	1320	1317	1320	1294	1357	1294	1357	1357	1357
CST Math Performance Standards	Spearman's rho	.616**	.432**	.616**	.431**	.927**	.485**	1.000	.473**	.488**	.469**
	Correlation
	N	1269	1278	1269	1278	1310	1294	1310	1294	1310	1310
CST English Performance Standards	Spearman's rho	.592**	.766**	.583**	.764**	.484**	.956**	.473**	1.000	.501**	.416**
	Correlation
	N	1317	1320	1317	1320	1294	1357	1294	1357	1357	1357
MARS raw score	Spearman's rho	.594**	.445**	.591**	.441**	.502**	.522**	.488**	.501**	1.000	.849**
	Correlation
	N	1326	1337	1326	1337	1310	1357	1310	1357	1481	1481
MARS Performance Level	Spearman's rho	.529**	.375**	.525**	.371**	.465**	.431**	.469**	.416**	.849**	1.000
	Correlation
	N	1326	1337	1326	1337	1310	1357	1310	1357	1481	1481

** . Correlation is significant at the .01 level (2-tailed).

Grade 10
Correlations of STAR scores with MARS scores

		CAT/6 Math raw score	CAT/6 Reading raw score	CAT/6 Math percentile rank	CAT/6 Reading percentile rank	CST Math raw score	CST English Language raw score	CST Math Performance Standards	CST English Performance Standards	MARS raw score	MARS Performance Level
CAT/6 Math raw score	Spearman's rho Correlation	1.000	.466**	.997**	.460**	.742**	.494**	.707**	.481**	.671**	.643**
	N	220	219	220	219	217	219	217	219	220	220
CAT/6 Reading raw score	Spearman's rho Correlation	.466**	1.000	.462**	.997**	.593**	.864**	.570**	.833**	.536**	.520**
	N	219	219	219	219	216	218	216	218	219	219
CAT/6 Math percentile rank	Spearman's rho Correlation	.997**	.462**	1.000	.459**	.750**	.497**	.715**	.477**	.677**	.649**
	N	220	219	220	219	217	219	217	219	220	220
CAT/6 Reading percentile rank	Spearman's rho Correlation	.460**	.997**	.459**	1.000	.598**	.862**	.576**	.828**	.541**	.525**
	N	219	219	219	219	216	218	216	218	219	219
CST Math raw score	Spearman's rho Correlation	.742**	.593**	.750**	.598**	1.000	.621**	.963**	.594**	.824**	.786**
	N	217	216	217	216	219	217	219	217	219	219
CST English Language raw score	Spearman's rho Correlation	.494**	.864**	.497**	.862**	.621**	1.000	.592**	.945**	.560**	.551**
	N	219	218	219	218	217	220	217	220	220	220
CST Math Performance Standards	Spearman's rho Correlation	.707**	.570**	.715**	.576**	.963**	.592**	1.000	.572**	.796**	.766**
	N	217	216	217	216	219	217	219	217	219	219
CST English Performance Standards	Spearman's rho Correlation	.481**	.833**	.477**	.828**	.594**	.945**	.572**	1.000	.508**	.504**
	N	219	218	219	218	217	220	217	220	220	220
MARS raw score	Spearman's rho Correlation	.671**	.536**	.677**	.541**	.824**	.560**	.796**	.508**	1.000	.956**
	N	220	219	220	219	219	220	219	220	232	232
MARS Performance Level	Spearman's rho Correlation	.643**	.520**	.649**	.525**	.786**	.551**	.766**	.504**	.956**	1.000
	N	220	219	220	219	219	220	219	220	232	232

** . Correlation is significant at the .01 level (2-tailed).

CST Cluster Levels are compared to MAC data. Each task is correlated with the CST Cluster that more closely corresponds. In some cases the CST Cluster does not correspond to any of the tasks and no correlations are reported in those cases. The correlation coefficients below indicate the relationship between a particular CST Cluster and MAC Task. For example, Cluster 1 includes place value, fractions, and decimals items and it has a correlation of .541 with Task 1 which also includes the same type of item. A correlation of 1.00 signifies perfect agreement, a correlation of -1.00 signifies perfect disagreement, and a correlation of 0.00 shows no relation between the scores at all.

Grade 3
CST Cluster Level comparison to MAC tasks

		TASK1	TASK2	TASK3	TASK4	TASK5
CST Cluster1	Spearman's rho	.541**				
	Correlation	.				
	N	14178				
CST Cluster 2	Spearman's rho		.340**			
	Correlation		.			
	N		14178			
CST Cluster 3	Spearman's rho				.340**	
	Correlation				.	
	N				14178	
CST Cluster 4	Spearman's rho					.437**
	Correlation					.
	N					14178
CST Cluster 5	Spearman's rho			.355**		
	Correlation			.		
	N			14178		

** . Correlation is significant at the .01 level (2-tailed).

Cluster 1 – Number property: Place value, fractions and decimals and division

Cluster 2 – Number Operations: Addition, subtraction, multiplication

Cluster 3 – Algebra and functions

Cluster 4 – Measurement and Geometry

Cluster 5 – Statistics, data analysis, and probability

Grade 4
CST Cluster Level comparison to MAC tasks

		TASK1	TASK2	TASK3	TASK4	TASK5
CST Cluster1	Spearman's rho				.465**	
	Correlation				.	
	N				9082	
CST Cluster 2	Spearman's rho		.525**			
	Correlation		.			
	N		9082			
CST Cluster 3	Spearman's rho			.467**		
	Correlation			.		
	N			9082		
CST Cluster 4	Spearman's rho	.490**				
	Correlation	.				
	N	9081				
CST Cluster 5	Spearman's rho					.326**
	Correlation					.
	N					9080

** . Correlation is significant at the .01 level (2-tailed).

Cluster 1 – Number property: Decimals, fractions, and negative numbers

Cluster 2 – Number Operations: Operations and factoring

Cluster 3 – Algebra and functions

Cluster 4 – Measurement and Geometry

Cluster 5 – Statistics, data analysis, and probability

Grade 5
CST Cluster Level comparison to MAC tasks

		TASK1	TASK2	TASK3	TASK4	TASK5
CST Cluster1	Spearman's rho		.608**			
	Correlation N		. 13971			
CST Cluster 2	Spearman's rho	.466**				
	Correlation N	. 13970				
CST Cluster 3	Spearman's rho			.507**		
	Correlation N			. 13971		
CST Cluster 4	Spearman's rho					.430**
	Correlation N					. 13970
CST Cluster 5	Spearman's rho				.454**	
	Correlation N				. 13971	

** . Correlation is significant at the .01 level (2-tailed).

Cluster 1 – Number property: Estimation, percents, and factoring

Cluster 3 – Algebra and functions

Cluster 5 – Statistics, data analysis, and probability

Cluster 2 – Number Operations: Operations with fractions and decimals

Cluster 4 – Measurement and Geometry

Grade 6
CST Cluster Level comparison to MAC tasks

		TASK1	TASK2	TASK3	TASK4	TASK5
CST Cluster 1	Spearman's rho					
CST Cluster 2	Spearman's rho					.526**
	Correlation					.
	N					9555
CST Cluster 3	Spearman's rho		.648**			
	Correlation		.			
	N		9555			
CST Cluster 4	Spearman's rho			.416**		
	Correlation			.		
	N			9555		
CST Cluster 5	Spearman's rho	.463**			.587**	
	Correlation	.			.	
	N	9555			9555	

** . Correlation is significant at the .01 level (2-tailed).

- Cluster 1 – Number property: Ratios, proportions, percentages, negative fractions
- Cluster 2 – Number Operations: Operations and problem solving with fractions
- Cluster 3 – Algebra and functions
- Cluster 4 – Measurement and Geometry
- Cluster 5 – Statistics, data analysis, and probability

CST Cluster Level comparison to MAC tasks

		TASK1	TASK2	TASK3	TASK4	TASK5
CST Cluster1	Spearman's rho	.536**				
	Correlation	.				
	N	11366				
CST Cluster 2	Spearman's rho					
CST Cluster 3	Spearman's rho					.522**
	Correlation					.
	N					11366
CST Cluster 4	Spearman's rho		.538**			
	Correlation		.			
	N		11366			
CST Cluster 5	Spearman's rho			.577**		
	Correlation			.		
	N			11364		
CST Cluster 6	Spearman's rho				.310**	
	Correlation				.	
	N				11345	

** . Correlation is significant at the .01 level (2-tailed).

Cluster 1 – Number property: Rational numbers

Cluster 2 – Number Operations: Exponents, powers, and roots

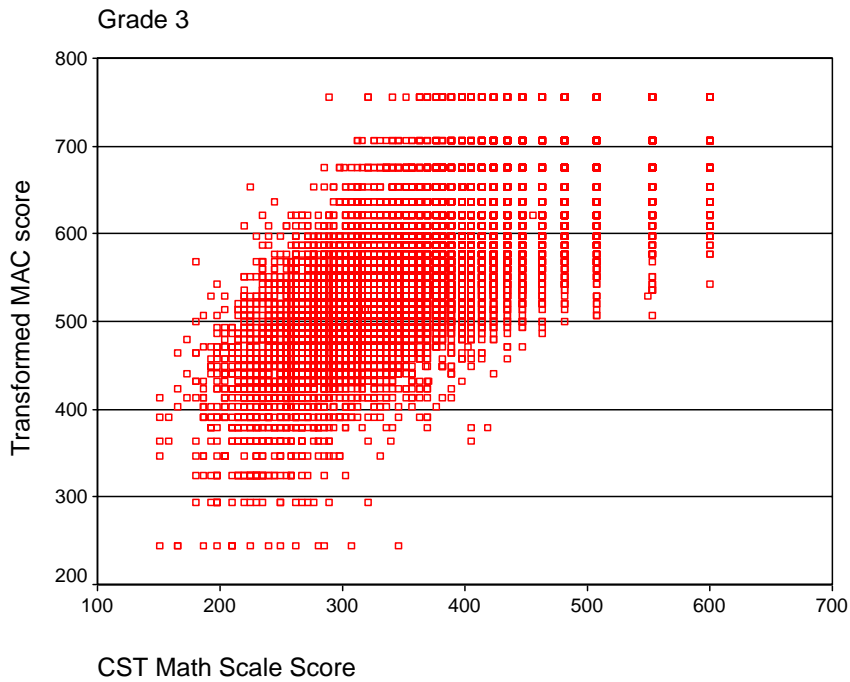
Cluster 3 – Algebra and functions: Quantitative relationships and evaluating expressions

Cluster 4 –Algebra and functions: Multi-step problems, graphing, and functions

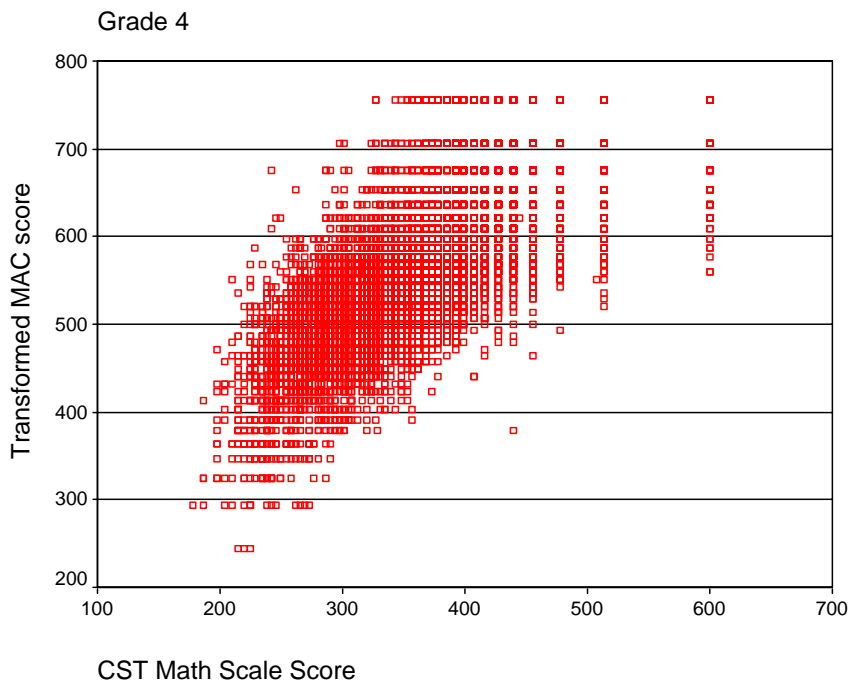
Cluster 5 – Measurement and Geometry

Cluster 6 – Statistics, data analysis, and probability

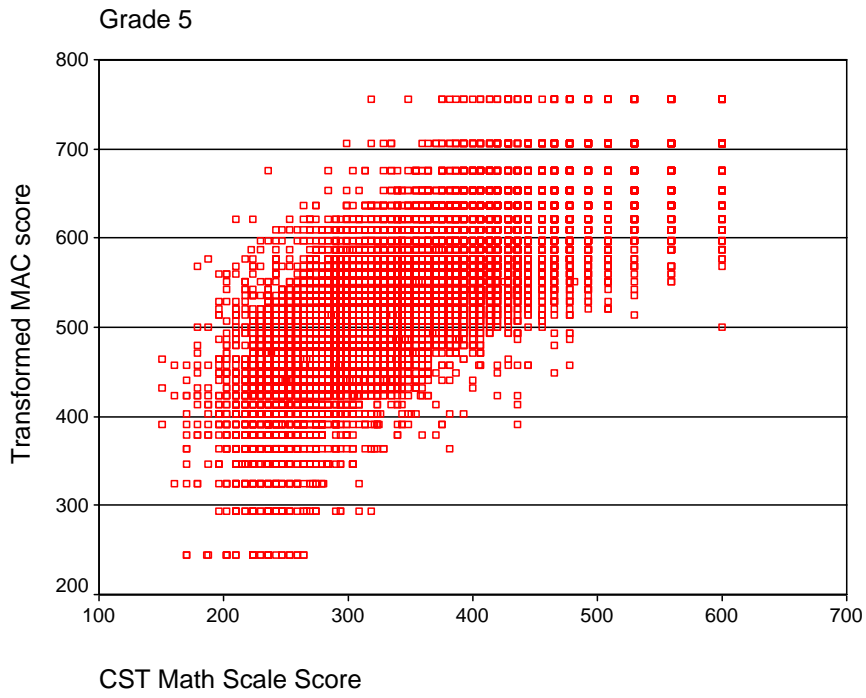
The following are scatterplots showing the relationship between CST Math scale scores and a transformed MAC score. The relationship between the variables is noted as a correlation. Two asterisks indicate statistical significance at the .01 level. In general, strong correlations are represented by a straight line and weak correlations are shown as “clouds”. Grade 3 shows a strong relationship ($r = .711$) between the CST Math Scale Score and the overall transformed MAC score. The reason the scatterplot looks more like a cloud than a straight line is that the scatterplot represents more than 14,000 students, so the cloud shape might be driven by a few outlier students.



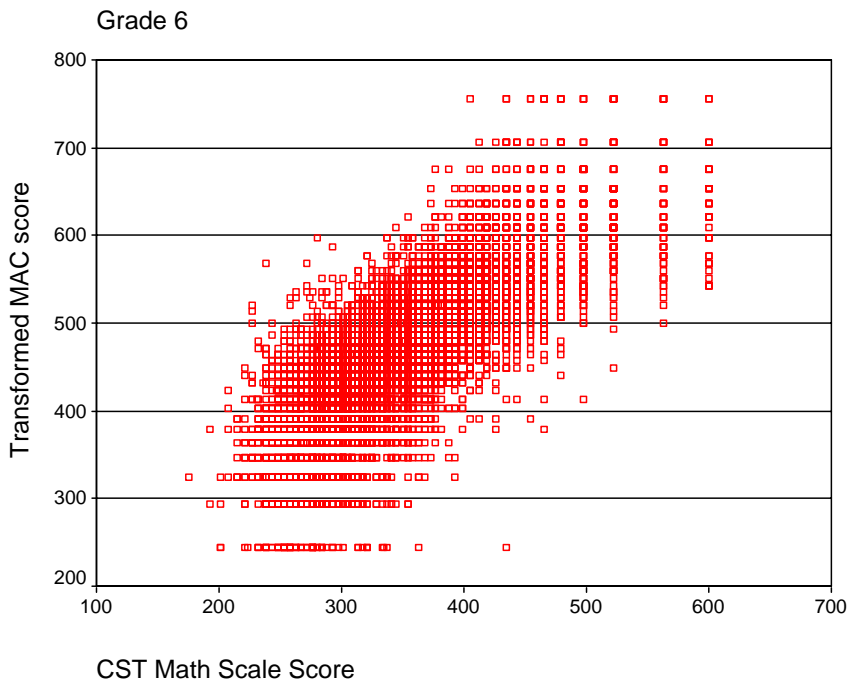
Correlation $r = .711^{}$**



Correlation $r = .710^{}$**

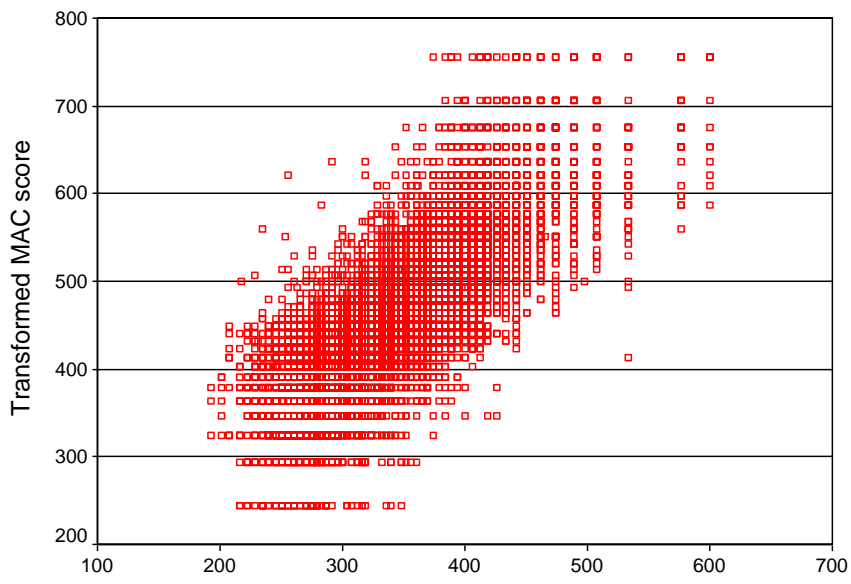


Correlation $r = .725^{}$**



Correlation $r = .777^{}$**

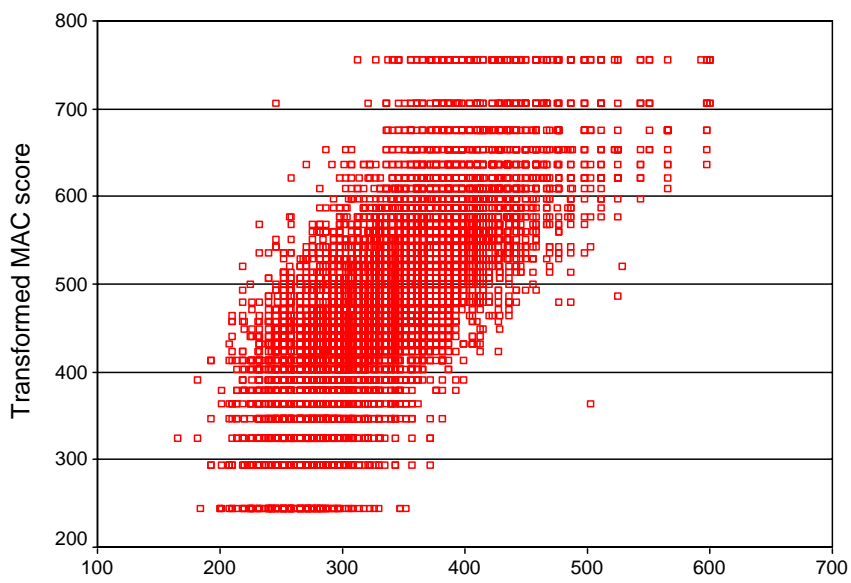
Grade 7



CST Math Scale Score

Correlation $r = .776^{**}$

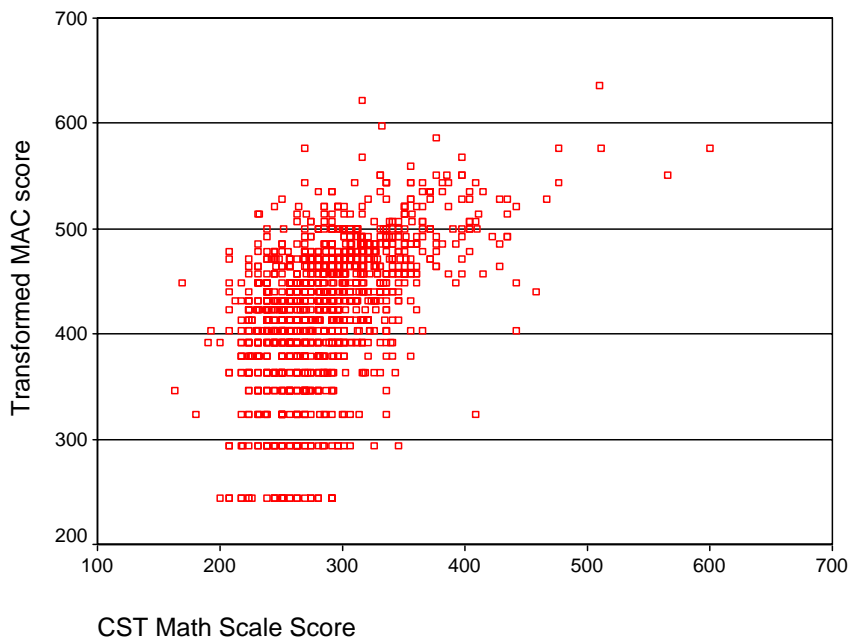
Grade 8



CST Math Scale Score

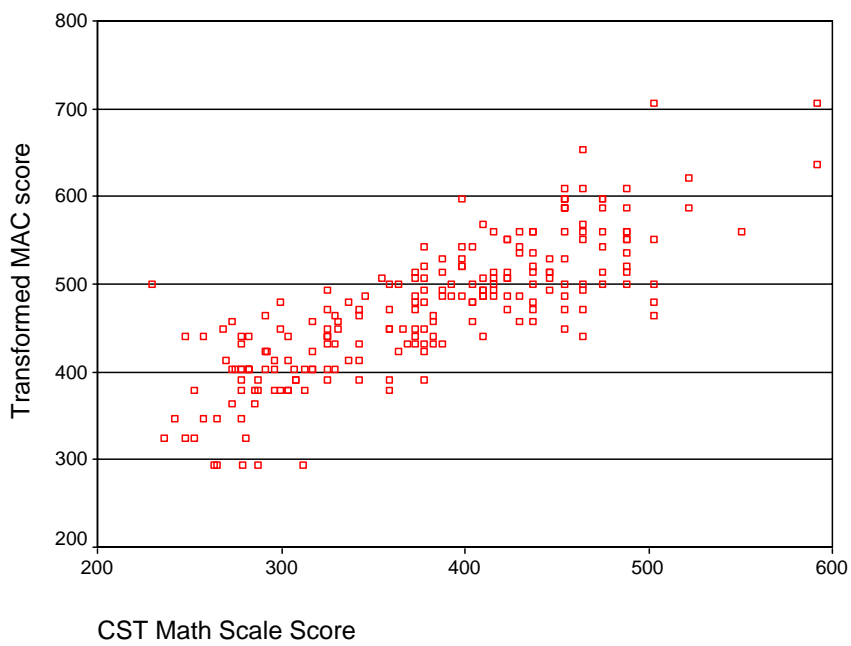
Correlation $r = .723^{**}$

Grade 9



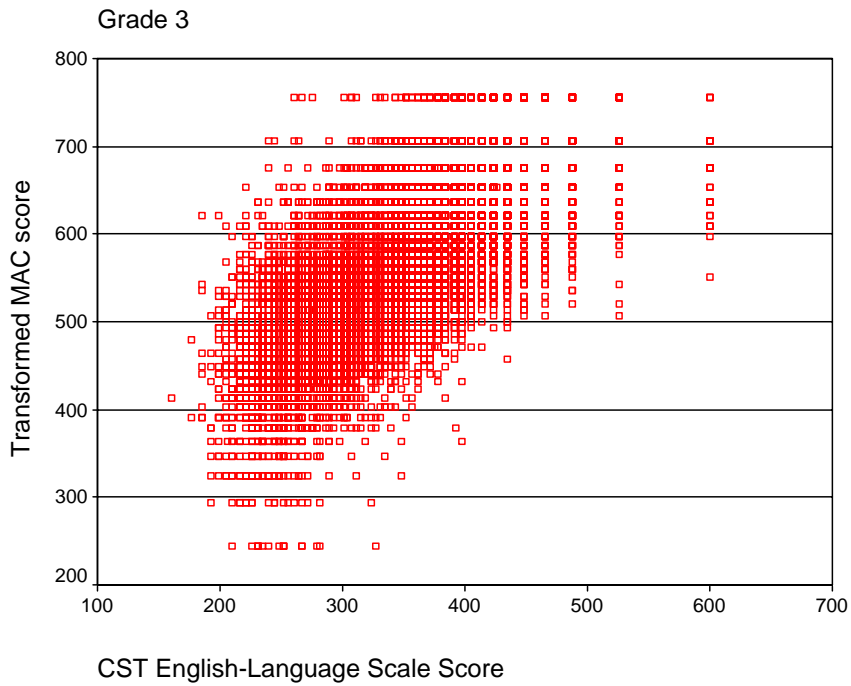
Correlation $r = .514^{**}$

Grade 10

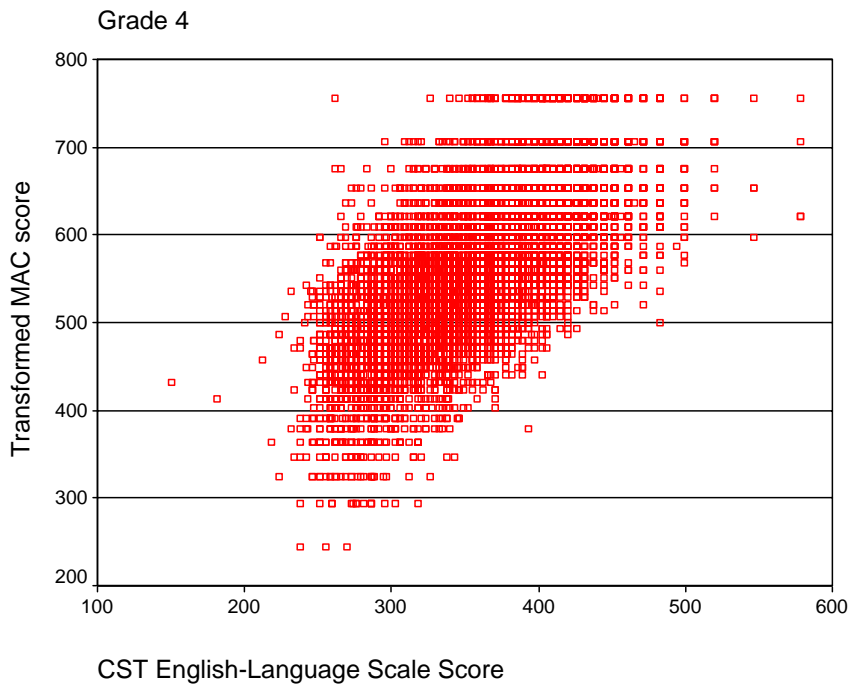


Correlation $r = .812^{**}$

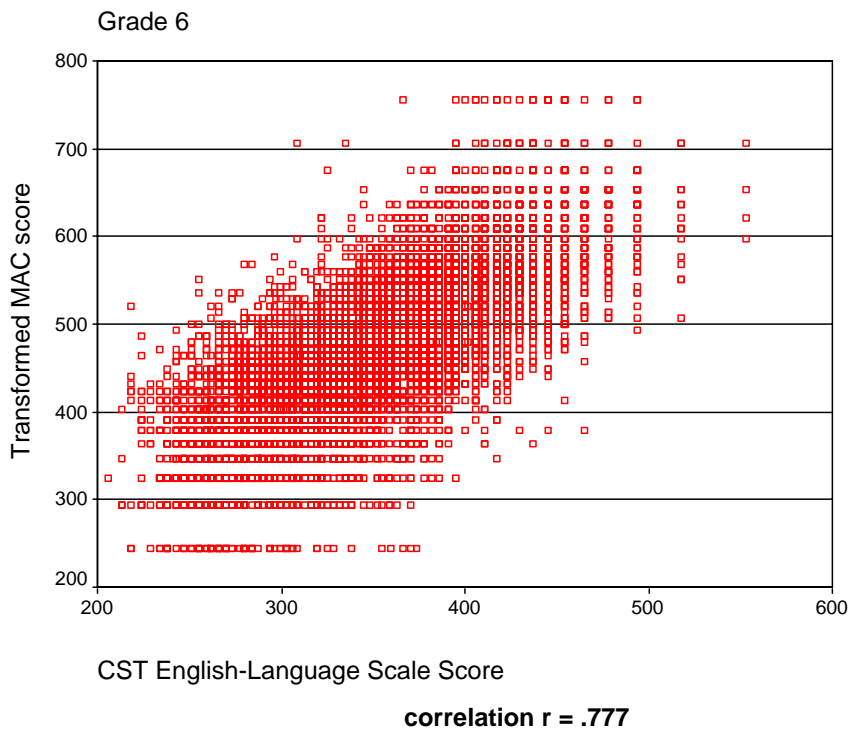
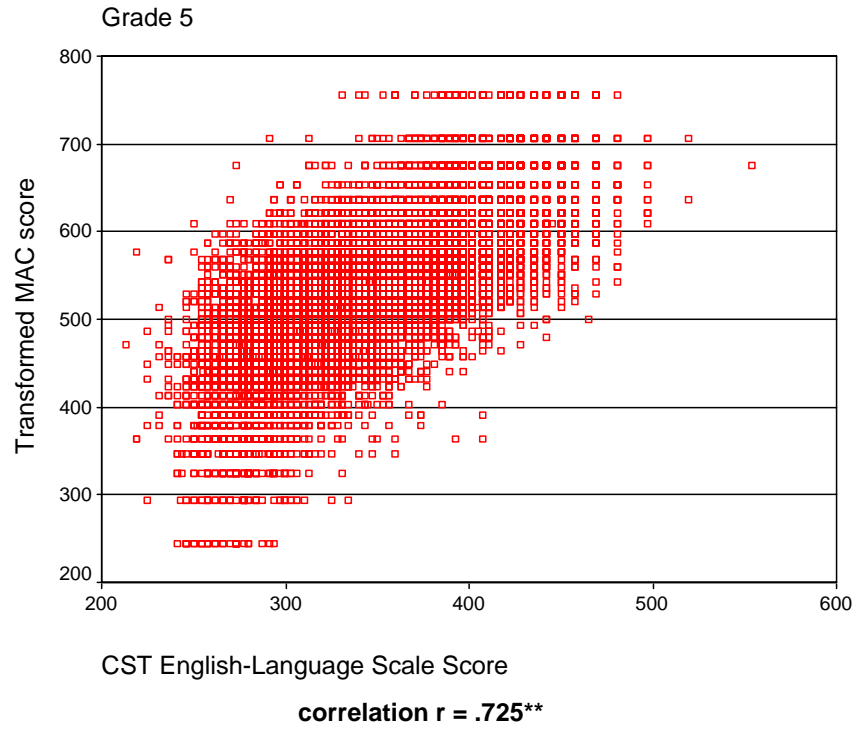
Scatterplots showing the relationship between CST English-Language Arts scale scores and a transformed MAC score.

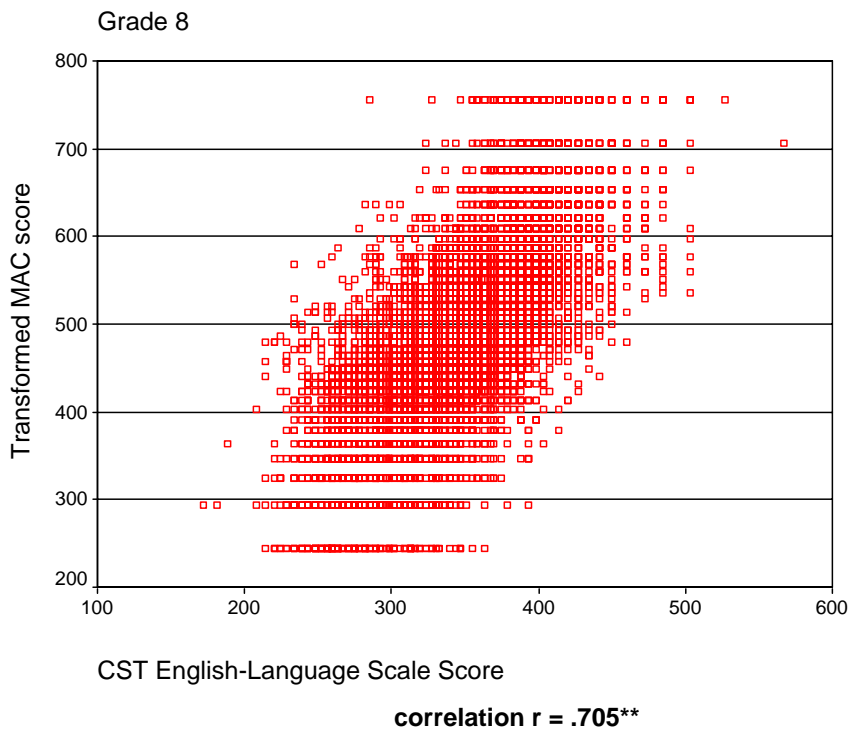
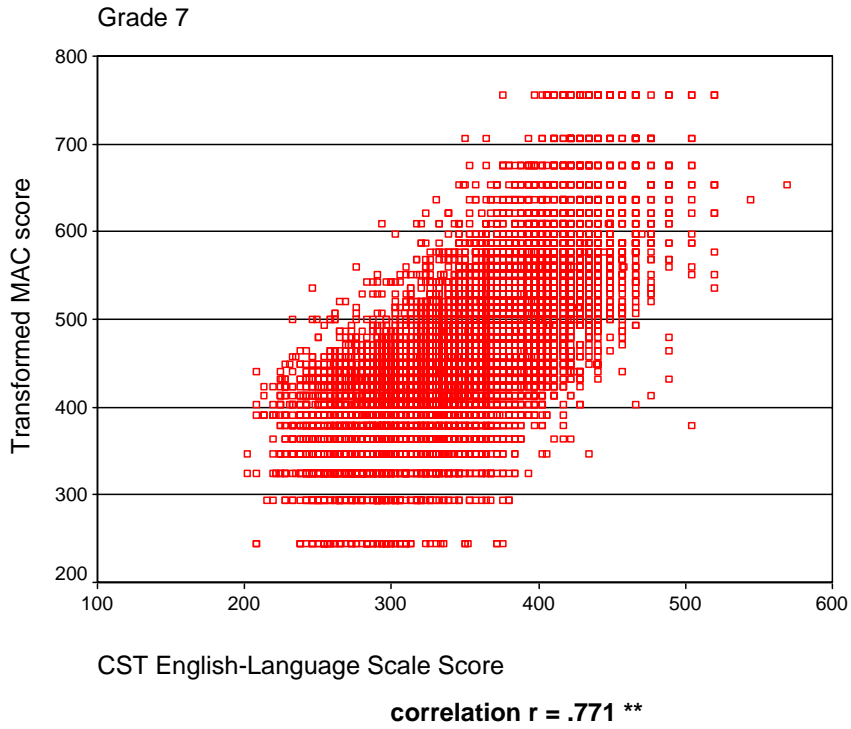


correlation $r = .711^{}$**

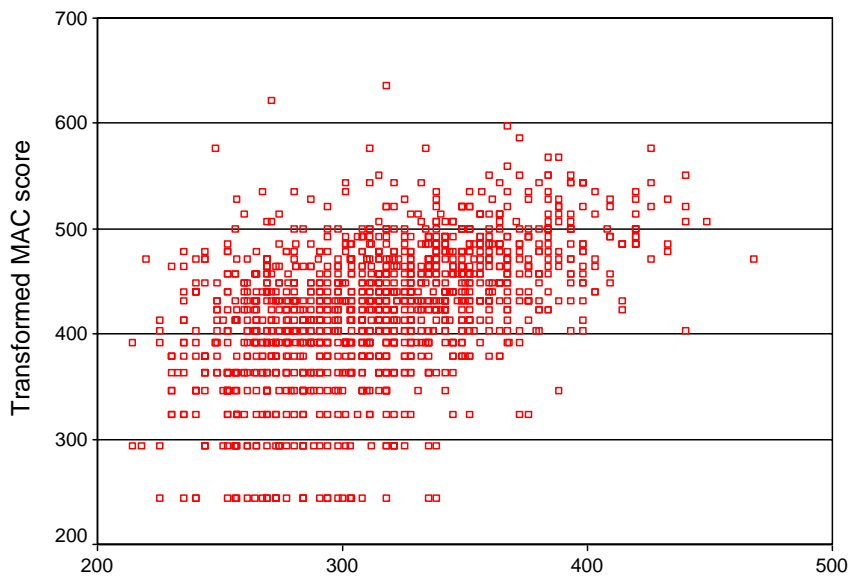


correlation $r = .708^{}$**





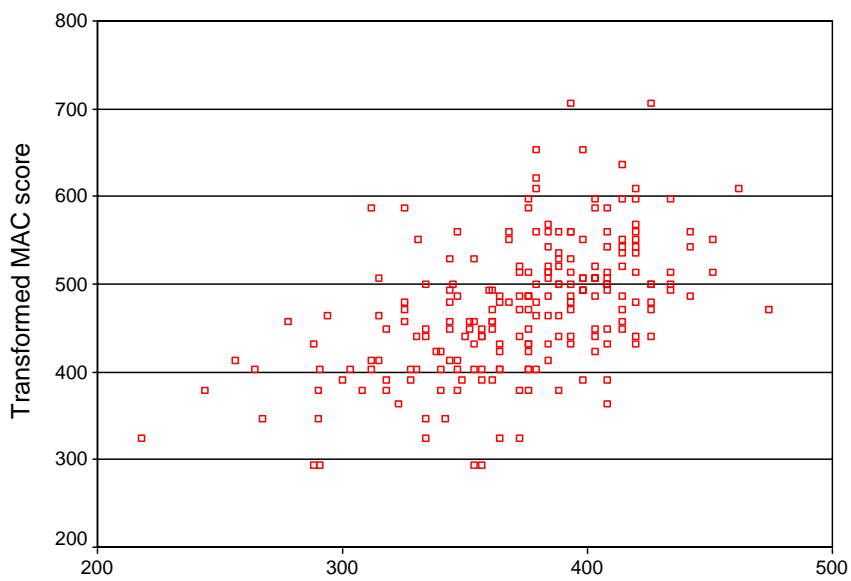
Grade 9



CST English-Language Scale Score

correlation $r = .513^{}$**

Grade 10



CST English-Language Scale Score

correlation $r = .810^{}$**

Appendix 2

Scoring Reliability

Scoring Reliability by Grade

Grade	N	Point Boundaries			
		0	+/- 1	+/- 2	+/- 3
Grade 3	770	56%	83%	94%	98%
Grade 4	467	50%	82%	94%	98%
Grade 5	737	62%	84%	93%	98%
Grade 6	536	68%	91%	98%	99%
Grade 7	617	70%	87%	98%	99%
Grade 8	447	68%	89%	96%	98%

Grade	Correlation
Grade 3	0.990
Grade 4	0.989
Grade 5	0.991
Grade 6	0.996
Grade 7	0.995
Grade 8	0.995

Grade	Mean Difference
Grade 3	-0.04
Grade 4	-0.06
Grade 5	-0.01
Grade 6	0.03
Grade 7	-0.04
Grade 8	-0.11

All differences are statistically significant.

Scoring Discrepancy by Task
Grade 3

		Difference between original and audited scores									
		-4	-3	-2	-1	0	1	2	3	4	Total
Even and Odd	Count		1	20	693	55	1				770
	% within Even and Odd		0.1%	2.6%	90.0%	7.1%	0.1%				100.0%
Addition Trains	Count		2	1	4	756	5	1	1		770
	% within Addition Trains		0.3%	0.1%	0.5%	98.2%	0.6%	0.1%	0.1%		100.0%
Vending Machines	Count			6	25	712	23	3	1		770
	% within Vending Machines			0.8%	3.2%	92.5%	3.0%	0.4%	0.1%		100.0%
Cherie's Shapes	Count	3	6	28	45	631	41	12	3	1	770
	% within Cherie's Shapes	0.4%	0.8%	3.6%	5.8%	81.9%	5.3%	1.6%	0.4%	0.1%	100.0%
Patchwork Quilt	Count	1	7	36	31	596	79	16	4		770
	% within Patchwork Quilt	0.1%	0.9%	4.7%	4.0%	77.4%	10.3%	2.1%	0.5%		100.0%

Scoring Discrepancy by Task
Grade 4

		Difference between original and audited scores									
		-4	-3	-2	-1	0	1	2	3	4	Total
Shapes with Straws	Count		1	2	44	358	40	15	6	1	467
	% within Shapes with Straws		0.2%	0.4%	9.4%	76.7%	8.6%	3.2%	1.3%	0.2%	100.0%
Number Trains	Count			6	23	412	23	3			467
	% within Number Trains			1.3%	4.9%	88.2%	4.9%	0.6%			100.0%
Hexagon Desks	Count	3	4	5	2	434	2	14	1	2	467
	% within Hexagon Desks	0.6%	0.9%	1.1%	0.4%	92.9%	0.4%	3.0%	0.2%	0.4%	100.0%
Flower Arranging	Count		1	6	37	406	14	3			467
	% within Flower Arranging		0.2%	1.3%	7.9%	86.9%	3.0%	0.6%			100.0%
Traveling to School	Count			6	43	404	11	3			467
	% within Traveling to School			1.3%	9.2%	86.5%	2.4%	0.6%			100.0%

Scoring Discrepancy by Task
Grade 5

		Difference between original and audited scores										
		-4	-3	-2	-1	0	1	2	3	4	5	Total
Number Story Time	Count		4	23	21	633	32	22	2			737
	% within Number Story Time		0.5%	3.1%	2.8%	85.9%	4.3%	3.0%	0.3%			100.0%
Raspberry Cake	Count		11	7	23	654	25	12	4			737
	% within Raspberry Cake		1.5%	0.9%	3.1%	88.7%	3.4%	1.6%	0.5%			100.0%
Buttons	Count			5	28	667	34	2			1	737
	% within Buttons			0.7%	3.8%	90.5%	4.6%	0.3%			0.1%	100.0%
Winter Sports	Count		1	1	25	678	26	5		1		737
	% within Winter Sports		0.1%	0.1%	3.4%	92.0%	3.5%	0.7%		0.1%		100.0%
Juan's Shapes	Count	1	11	3	12	693	6	2	9			737
	% within Juan's Shapes	0.1%	1.5%	0.4%	1.6%	94.0%	0.8%	0.3%	1.2%			100.0%

Scoring Discrepancy by Task
Grade 6

		Difference between original and audited scores								
		-3	-2	-1	0	1	2	3	4	Total
Baseball Players	Count		1	17	499	18	1			536
	% within Baseball Players		0.2%	3.2%	93.1%	3.4%	0.2%			100.0%
Gym	Count		1	19	501	15				536
	% within Gym		0.2%	3.5%	93.5%	2.8%				100.0%
Square Elk	Count			7	517	10	1	1		536
	% within Square Elk			1.3%	96.5%	1.9%	0.2%	0.1866		100.0%
Spinners	Count	2	5	29	457	28	12	3		536
	% within Spinners	0.4%	0.9%	5.4%	85.3%	5.2%	2.2%	0.6%		100.0%
Rabbit Costumes	Count	1	4	19	492	10	8	1	1	536
	% within Rabbit Costumes	0.2%	0.7%	3.5%	91.8%	1.9%	1.5%	0.2%	0.2%	100.0%

Scoring Discrepancy by Task
Grade 7

		Difference between original and audited scores									
		-4	-3	-2	-1	0	1	2	3	4	Total
Mixing Paints	Count			8	3	595	4	5	2		617
	% within Mixing Paints			1.3%	0.5%	96.4%	0.6%	0.8%	0.3%		100.0%
Hexagons	Count			2	15	586	9	5			617
	% within Hexagons			0.3%	2.4%	95.0%	1.5%	0.8%			100.0%
Pattern	Count			4	21	574	12	3	1	2	617
	% within Pattern			0.6%	3.4%	93.0%	1.9%	0.5%	0.162 1	0.3%	100.0%
Fair Game?	Count	2		6	34	548	18	6	3		617
	% within Fair Game?	0.3%		1.0%	5.5%	88.8%	2.9%	1.0%	0.5%		100.0%
Yogurt	Count			20	18	552	10	16	1		617
	% within Yogurt			3.2%	2.9%	89.5%	1.6%	2.6%	0.2%		100.0%

Scoring Discrepancy by Task
Grade 8

		Difference between original and audited scores										
		-4	-3	-2	-1	0	1	2	3	4	5	Total
Pete's Numbers	Count		1	1	15	417	10	2	1			447
	% within Pete's Numbers		0.2%	0.2%	3.4%	93.3%	2.2%	0.4%	0.2%			100.0%
Squares and Rectangles	Count	1	2	14	22	387	18	1	1	1		447
	% within Squares and Rectangles	0.2%	0.4%	3.1%	4.9%	86.6%	4.0%	0.2%	0.2%	0.2%		100.0%
Sports Injuries	Count		2	1	6	435	3					447
	% within Sports Injuries		0.4%	0.2%	1.3%	97.3%	0.7%					100.0%
Dots and Squares	Count		2	7	29	388	18	3				447
	% within Dots and Squares		0.4%	1.6%	6.5%	86.8%	4.0%	0.7%				100.0%
Number Pairs	Count		1	3	3	426	9	3			2	447
	% within Number Pairs		0.2%	0.7%	0.7%	95.3%	2.0%	0.7%			0.4%	100.0%

Appendix 3

Status of the Data

Status of the Data - This table shows the number of students in each grade for each district.

		Grade								Total
		3	4	5	6	7	8	9	10	
District ID	10	1439		1274		1145				3858
	11	876		904		956				2736
	12	66	71	85	61	73	79			435
	13	232	234	254	257	292	285			1554
	15	792	778	810	691	757	383			4211
	16	277	267	301	267	320	281			1713
	17	40	58	63	45	22	17			245
	18							1170		1170
	19	1359		1437						2796
	20	314	332	327	377	371	386			2107
	22	253	266	337	287	306	308			1757
	24	491	435	476	457	470	401			2730
	25	616	596	618	657	567	585			3639
	26	526	489	486	414	370	382			2667
	29	761		759		730				2250
	30	73	64	101	64	72	71			445
	32	229	248	209	253	173				1112
	33	2512	2453	2322	2367	1939	1579			13172
	35	1060	1053	1068	977	944	894			5996
	36	472	499	458	519	532	507			2987
	38	50	48	55	48	37	42			280
	40	766			755		429	197	232	2379
	41	770		420		726		114		2030
	42	144	151	157	132	253	242			1079
	43	197	1076	242	1	351	982			2849
	45	253	252	286	301	290	305			1687
	47			912	907					1819
Total		14568	9370	14361	9837	11696	8158	1481	232	69703

Understanding the t-test: If we had a single large pot containing student scores, we could select samples of scores at random, and we could call one, for example, “boys' scores” and another “girls' scores”, and compare them. There would almost always be a difference in the means; small differences would be more common than large differences. Differences in the scores would be larger if the range of scores in the pot were large (say 1 to 100) rather than small (say 1 to 5). For a given range of scores, small differences would be more likely when the sample sizes were large. For any particular sample size, we could count the number of times we got differences of a certain magnitude. Large differences would be rarer. Suppose we have a particular result we want to know about (there were 30 boys and 20 girls; the mean difference was 1.5 and the range of scores was 6, how likely is it that this was just a statistical fluctuation?). We could simulate drawing two sets of scores (of 30 and 20) over and over again from a pot of scores with a range of 6. We could count the relative occurrence of a difference in the means as big as 1.5.

The t-test provides a way of calculating the rarity of an observed difference between two sample means, if all the scores were actually drawn from the same pot, without having to do the simulation. Statistical significance is assessed in terms of the probability of getting a difference as big or bigger than the one observed, if the scores were actually drawn from the same pot. If an event would happen very rarely, the idea of “all scores coming from the same pot” is thrown away, and an alternative view- that the scores came from different pots- is accepted. By convention, if the observed difference would occur less than 1 in 20 times, we conclude that there is a “statistically significant difference between the means”. So statistically significant differences occur when the number “Significance 2-tailed” is smaller than 0.05.

The **correlation coefficient** provides a measure of the agreement between two sets of scores. Imagine the situation where 2 people grade a pile of scripts, independently. We could create a scatter plot of scores with the scores from the first marker along the y-axis, and the scores of the second marker along the x-axis. If all the points lie on a straight line with a positive slope, then there is perfect agreement about the rank order of students. If the points look like a fat cloud, there is little agreement. The correlation coefficient quantifies this intuitive idea. For linear relationships, a correlation of +1.00 signifies perfect agreement; a correlation of -1.00 signifies perfect disagreement. A correlation of 0.00 shows no relation between the scores at all. Typically, test designers use correlations to measure the agreement between scorers, and talk about “scorer reliability” (similarly, if students take parallel forms of the same test, it is common to talk about “test-retest reliability”-again using the correlation coefficient.). 'Statistical significance' is judged in ways analogous to the t-test. If two sets of scores were drawn from the same pot, in pairs, the correlation coefficient can be calculated. We could do this over and over again - the calculated correlation coefficient would almost never be exactly zero, but it would usually be small. When we have a real correlation coefficient, we could see how often a correlation that big would arise as the result of a random process. If it would occur less than one in 20 times (or one in 100 times) we conclude that the result is statistically significant at the 5% level (or the 1% level). The Spearman's rho correlation coefficient is used to correlate continuous and rank ordered variables.

Understanding the F-test: The variability in a given set of data can be calculated by looking at the amount by which the individual data points vary from the mean. In practice, this involves taking the squared values of these differences so that positive and negative differences can be taken into account. This measure is known as the variance of the sample. If, as in the t-test, sample scores were selected at random from the same large pot, there would be a difference in the variance of the two samples, small differences being more common than large differences. In the same way as the t-test provides a measure of the probability that the mean of the two samples are in fact significantly different statistically, the F-test provides a measure of the statistical significance of the recorded difference in the variance of the two samples. To inspect the reliability of trends in samples of data recorded at different times, the variance of each sample is compared to that of the whole. If the measured difference in variance is shown to be unlikely to occur by chance (By convention, less than 1 in 20), then the variation in data and therefore any observed trend, is said to be statistically significant.

Director's Note

The MAC Project is made possible through the generous contributions of the Robert Noyce Foundation. David Foster has been instrumental in his clear vision for mathematics, identifying the work of MARS and how it connects to that vision, and being able to pull together the diverse groups, which make up the Mathematics Assessment Collaborative. None of this would have materialized without his leverage and encouragement.

The Santa Clara Valley Mathematics Project, led by Dr. Joanne Rossi Becker, has been instrumental in the success of the Mathematics Assessment Collaborative. Dr. Becker and the Math Project provide support to M.A.C. in various capacities. Dr. Becker serves on the M.A.C. Executive Committee and plays an important role in advising the director and membership on matters of mathematics education. In other roles, Dr. Becker helps the Collaborative keep focused on mathematics standards, provides mathematical expertise and helps to set performance level boundaries. The San Jose State University Foundation is the fiscal agent of M.A.C. Dr. Becker plays an essential role in overseeing the budget and expense payments. Dr. Becker assists with high school professional development and arranging the audit scoring sessions that employ San Jose State University students and are conducted in their Mathematics Department. In addition SCMVP and Dr. Becker provide ongoing professional development for the member districts through projects such as the Summer Lab Schools and the Summer Coaching Institutes. These grants and programs also provide support that allows participating teachers to attend the MAC professional development sessions.