Test Date: June 1996

Harvard Quantitative Reasoning Test

As part of their graduation requirements, Harvard undergraduate students must pass an exit exam in quantitative reasoning. Norman Webb from the University of Wisconsin selected ten questions from the Harvard exam. Among the topics covered by the questions were statistical reasoning and interpretation of graphs.

A test consisting of these ten questions was administered in June 1996 to 150 students completing their junior year at Central High School. Of these, 91 were enrolled in IMP-3 and the other 59 in Algebra II. Eighth grade scores on national tests were used as baseline data.

Eighth Grade Scores of Central HS's 11th grade students:

	N	MEAN	STDEV
IMP	**91**	**94.53**	**4.18**
Algebra II	59	93.85	4.59

Questions taken from Harvard Exam given to Central HS's 11th grade students:

IMP	**91**	**5.09**	**0.23**
Algebra II	59	2.31	1.72

The difference was statistically significant ($p < .0001$).

COMPARISON ON A QUANTITATIVE REASONING TEST OF GRADE 11 INTERACTIVE MATHEMATICS PROGRAM (IMP) STUDENTS WITH ALGEBRA 2 STUDENTS AT ONE HIGH SCHOOL

April 24, 1997

Maritza Dowling
Norman L. Webb

Wisconsin Center for Education Research
School of Education
University of Wisconsin-Madison

The Interactive Mathematics Program Evaluation Project operates under the auspices of the Wisconsin Center for Education Research and is funded under a contract with the San Francisco State University Foundation, Inc. with resources that are provided by the National Science Foundation (award number ESI-9255262). The opinions, findings, and conclusions that are expressed in this paper do not necessarily reflect those of the supporting agencies.

Table of Contents

Page

Summary

Grade 11 students in four IMP Year 3 classes were compared to Grade 11 students in three algebra 2 classes at one high school in a large city located in an eastem state. Students were administered 10 multiple-choice items taken from a practice version of a quantitative reasoning test (QRT) developed by a prestigious university for first-year students. The test included items on data interpretation, probability, and statistics. IMP students significantly outperformed algebra 2 students as determined by tests of statistical significance. A matched-group analysis, which controlled for prior mathematics achievement, ethnicity and sex, produced consistent results, with IMP students significantly outperforming algebra 1 students.

COMPARISON ON A QUANTITATIVE REASONING TEST OF GRADE 11 INTERACTIVE MATHEMATICS PROGRAM (IMP) STUDENTS WITH ALGEBRA 2 STUDENTS AT ONE HIGH SCHOOL

Maritza Dowling and Norman L. Webb
Wisconsin Center for Education Research, School of Education,
University of Wisconsin-Madison

An increasing number of studies have compared the mathematical knowledge of students enrolled in the Interactive Mathematics Program (IMP) with that of students enrolled in the traditional algebra 1, geometry, algebra 2 sequence. Accumulating evidence shows that IMP students perform as well as, if not better than, students taking the traditional curriculum, as measured by standardized norm-referenced tests such as the Scholastic Assessment Test (SAT), Pre-Scholastic Assessment Test (PSAT), and the Comprehensive Test of Basic Skills (Interactive Mathematics Program, 1995; Schoen, 1993; S. Chew, personal communication, November, 1996; Webb & Dowling, 1995a, 1995b, 1995c, 1996). These traditional instruments measure students' knowledge of very general mathematics skills and reasoning in mathematics. Probability, statistics, quantitative reasoning, and problem solving are given little or no attention on these instruments however. It is increasingly important for students to have knowledge in these areas. This has been recognized by the NCTM Curriculum and Evaluation Standards for School Mathematics (1989) and by various state standards such as South Carolina Mathematics Framework (1993), Illinois Academic Standards Project (1996), and New Jersey Mathematics Curriculum Framework (Rosenstein, Caldwell, & Crown, 1996). Therefore, we set out to examine the performance of IMP students on these critical criteria, which have not been focused upon by past research.

A series of three studies was designed to produce information on how well IMP students perform on activities using probability, statistics, reasoning, and problem solving. The IMP students were compared to students who ,were enrolled in the traditional algebra I (Dowling & Webb, 1997a), geometry (Dowling & Webb, 1997b) and algebra 2 sequence. This is the report of the grade 11 study that compared IMP Year 3 students with algebra 2 students at one high school in a large eastern city. The school is a college-preparatory magnet public high school located in a large city in the eastern United States. In 1994, the school enrolled over 2,000 students – 35 percent black; 17 percent Asian, 4 percent Latino, and 44 percent white. A large number of students who spoke English as a second language were enrolled. Students had to meet minimum requirements to attend School 3. About one-third of the applicants were accepted. Nearly all, 98 percent, of the students continue their education at universities or colleges.

To make valid comparisons between IMP and the traditional mathematics course sequence, specific controls were imposed on the design of these three studies. Each participating school had to provide grade 8 standardized norm-referenced mathematics test scores on a significant number of IMP and traditional mathematics students who were tested. An adequate number of students, as close to 50 as possible or more in each program, had to be available to be tested. This meant that there had to be more than one IMP class at the target grade level. One or more teachers of the traditional mathematics course taken by students at the target grade level had to agree to participate in the study. The tests had to come from a source independent of IMP and had to be easily administered under the same conditions to classes of students in both course sequences. The inconvenience to teachers and students had to be kept to a minimum. Students' knowledge on as wide a range of content as possible was sought. Regional and school factors had to be reduced as much as possible. Effects had to be attributed to the curriculum rather than other factors such as teacher, school, or region.

Three studies were conducted. Each study was done in a different school at a different grade level with a different outcome measure. This design was chosen to meet the criteria above while maximizing the comparative information and minimizing the inference within any one school. Three different instruments were used. The study reported here compared the performance of grade 11 IMP Year 3 students with algebra 2 students on 10 multiple-choice items taken from a practice quantitative reasoning test (QRT) developed by a prestigious university for first-year students. The test focuses mainly on data interpretation. Its purpose is to gather evidence on using mathematics, probability, statistics, and Computation to solve problems. The test includes items on various topics: data and graph interpretation, probability, and basic understanding of statistics concepts, such as the standard deviation and the mean of a distribution. The test requires students to understand and apply arguments which are supported by numerical data.

Sample

In May 1996, the QRT test was administered to a total of 133 students. Students were allowed to use any form of calculator on the test. The test was administered to students enrolled in four IMP Year 3 classes and three algebra 2 classes. Only grade 11 students were included in the analysis. Fifty-nine students out of 126 (47 percent) had attended public schools before enrolling in high school; the other 67 (53 percent) had attended private schools. The type of middle school was not indicated for seven students (Table 1, p. 11). In addition to the students' responses to each item scores on standardized tests, and student demographic characteristics, such as ethnicity and sex, were obtained from the school. The grade 8 standardized test scores reported by the school were the California Achievement Test (CAT), Comprehensive Test of Basic Skills (CTBS), Stanford Achievement Test (SAT), and the Educational Record Bureau (ERB) test. All students who had attended a public middle school had taken the CAT test. Other students who bad attended private middle schools had taken a variety of standardized tests. Approximately 84 percent of the students took the CAT. The remaining 16 percent took one or more of the other tests.

A greater proportion of females than males took the exam; 72 out of 133 (54 percent of the test takers) were female (Table 2, p. 12). More than half of those tested (93 out of 133, or 70 percent) were IMP Year 3 students and the remaining 40 (30 percent) were students taking algebra 2. The ethnic composition of the total group w as predominantly white (60 percent), then black (23 percent), Asian (8 percent), other (5 percent), and Hispanic (3 percent) (Table 3, p. 13). Because of rounding, disaggregated percentages will not always total 100 percent. A total of seven classes or groups taught by four different teachers or pairs of teachers were administered the test (Table 4, p. 14). The three algebra 2 classes were taught by the same teacher during the 1995-96 school year. Four IMP classes were taught by three teachers—two taught by two different teachers, one taught by these two teachers as a team, and one taught by one of these teachers and a third teacher.

Results

Overall, IMP Year 3 students did significantly better on the test than the algebra 2 students (Table 5, p. 15). The item difficulty (the average of the proportion of students who correctly answered an item) was .50 for the IMP group and .24 for the algebra 2 group (Table 5, p. 15). The item difficulty levels ranged from .18 to .75 for the IMP group, and from .03 to .53 for the algebra 2 group (Table 12, p. 22). An "acceptable" item difficulty range for a five-options multiple-choice item is between .30 and .90. For the IMP group, 9 out of the 10 items were within this range. However, for the algebra 2 group only 3 out of the 10 items were within this "acceptable" range. Algebra 2 students had great difficulty with the test: their performance was essentially at the guessing level.

The items related to the application of the normal curve properties, especially the standard deviation, were relatively difficult for both groups (e.g., item 1; Table 6, p. 16). Other relatively difficult items, especially for the algebra 2 group, were those related to extrapolation using rates and percentages (items 4, 5, and 6). Item difficulties on these items ranged from .43 to .54 for the IMP group and from .10 to .23 for the algebra 2 group. IMP Year 3 students did considerably better on items related to probability concepts than algebra 2 students (items 7 and 8). The item difficulty values were .75 (item 7) and .67 (item 8) for the IW group and .53 and .38 respectively for the algebra 2 group.

The alpha reliability coefficient for a test provides an estimate of the accuracy of the test scores as indicated by the internal consistency, or degree to which a of the items in the test are consistent in rank ordering the students. The obtained reliability estimate for the total group (n = 133) was .65. This is considered a reasonably high estimate, given the fact that the test was relatively short (only 10 items) and it covered a variety of topics. The standard error of measurement for the test was 1.36 raw score units (Table 5, p. 15). As expected, given the obtained range of item difficulty values for each group, the reliability estimate was higher for the IMP Year 3 sample (.56) than for the algebra 2 sample (.33).

Averaging the national percentile scores across students, discounting that not all students had taken the same test, the mean score on grade 8 mathematics test was significantly different in favor of IMP students (Table 7, p. 17). National percentile scores determined by different instruments are not equivalent. The mean scores computed here have to be interpreted very broadly. In a matched-group analysis, reported in the next section, only students with a score from the same grade 8 mathematics test were included in the analysis.

When using grade 8 CAT national percentile scores for the “public middle school” sample and grade 8 national percentile scores for the other standardized tests (i.e., a combination of CAT, CTBS, SAT, and ERB) for the “private” sample as covariates, the difference in QRT test scores between the IMP Year 3 and algebra 2 students was found to be highly significant ($p < .0001$; $p < .0001$) (Tables 8 and 9, pp. 18-19). When using the score obtained in the grade 8 standardized test in mathematics as a covariate, for the total group, the result was consistent: that is, the IMP group performed significantly better than the algebra 2 group ($p < .0001$) (Table 10, p. 20).

The highest attainable score on the QRT test was 10. Students obtained scores ranging from 0 to 10. Two IMP students obtained a perfect score on the test. None of the algebra 2 students did this. Three students obtained a score of 0: one IMP student and two algebra 2 students. Additionally, males outperformed females on the test ($p < .016$) (Table 11, p. 21; Table 13, p. 23). No significant differences in performance by ethnicity were found on mean test scores (Tables 12 and 13, pp. 22-23).

An analysis of variance (anova) was used to test for differences among mean scores across class periods. As mentioned before, a total of 7 classes (four IMP Year 3 classes and three algebra 2 classes) taught by four different teachers were administered the quantitative reasoning test (Table 14, p. 24). There were no statistically significant differences in mean total test score by class periods within any group, IMP Year 3 or algebra 2 (Tables 15 and 16, pp. 25-26). A series of anovas were performed with the purpose of testing for mean differences between class periods across treatment (IMP vs. algebra). Almost all the mean differences between all 12 possible pairwise combinations between IMP and algebra 2 classes were statistically significant **except** for one: teacher C-period 3 (IMP) with a mean of 4.44, and teacher A-period 7 (algebra 2) with a mean of 3.13.

Matched-Group Analysis

Several variables—such as prior mathematical experience, sex, and ethnicity—could have influenced the outcome of the study. Upon entering high school, the group of IMP students who participated in the study had a higher average mathematics achievement than the group of algebra 2 students (see Table 17, p. 27). One way to control for this initial difference was to perform an analysis of covariance. The results of this analysis were discussed in the previous section of this paper. Another approach used to establish some statistical control over suspected nuisance variables possibly affecting students’performance on the QRT test was the large-sample, rnatched-group Wilcoxon test. For this test, students from both groups were matched according to prior achievement. Only students who had taken the CAT test in grade 8 were included in this analysis. Students in both groups, a total of 110 students who had taken the CAT in grade 8, were rank-ordered-77 IMP Year 3 students, and 33 algebra 2 students. After ranking students on CAT scores, the two groups (IMP and algebra) were matched, with the two highest and comparable-scoring students paired, the next two comparable-scoring students paired, and so on, with the two lowest and comparable- scoring students forming the last pair. A total of 31 pairs were formed. Additionally, the groups were formed taking into account ethnicity, sex, and the middle school attended. The main purpose of the pairing process was to form two groups (IMP Year 3 and algebra 2) that matched as closely as possible on mean CAT scores and by sex and ethnicity (Tables 17, 18 and 19, pp. 27-29).

The results of the matched-group analysis were consistent with those obtained in the analysis of the complete sample: The DOP Year 3 students outperformed the algebra 2 students. The large-sample Wilcoxon test comparing the group distribution was highly significant (Table 20, p. 30). The overall differences in mean test scores (5.42 for the IMP group and 2.39 for the algebra 2 group) and item difficulties (.54 for the IMP group and .24 for the algebra group) were larger than those obtained in the complete sample analysis (See Table 5, p. 15 and Table 21, p. 31). Also consistent with the results obtained in the complete sample analysis, males performed significantly better than females, although the difference in mean test scores and item difficulties were not as dramatic as those obtained in the prior analysis (see Table II, p. 21 and Table 22, p. 32).

The obtained item difficulty values of the IMP group in the matched-group analysis were slightly higher than those obtained in the complete sample analysis. On the other hand, the item difficulty values obtained by the algebra 2 group in the matched-group test were almost identical to those obtained before (Table 6, p. 16 and Table 23, p. 33). The item difficulty levels ranged from .23 to .84 for the IMP Year 3 group and from .03 to .52 for the algebra 2 group.

Differences in performance by item between IMP and algebra students were larger in the matched-group analysis compared to the complete sample analysis. For example, on items related to statistical concepts (specifically the application of the normal curve properties, items 1 to 3 on Table 23, p. 33), the difference between IMP and algebra 2 students was even greater for the matched-group analysis than it was for the complete sample analysis. The item difficulty values ranged from .23 to .68 for IMP students and from .03 to .26 for algebra students. Overall, IMP Year 3 students did significantly better than algebra 2 students on items related to graph interpretation and analysis using the slope concept (items 9 and 10) and extrapolation using rates and percentages (items 4, 5, and 6). Also consistent to prior findings, IMP students performed significantly better than algebra students on items related to probability concepts (items 7 and 8). The item difficulty values were .68 (item 7) and .71 (item 8) for the IMP Year 3 group and .52 and .36 respectively for the algebra 2 group.

The obtained reliability estimates were higher than those obtained in the complete sample analysis. This may be explained by the fact that the item difficulty values were slightly higher, especially for the IMP Year 3 group, compared to those obtained in the complete-sample analysis. Higher reliability estimates also suggest that the ability range in the subsample was broader than the complete group, or the subsample was more heterogeneous in ability. The obtained reliability estimate for the total matched-group (n = 62) was .71. The reliability estimate was .60 for the IMP Year 3 group and .41 for the algebra 2 group.

Conclusions

IMP Year 3 students scored significantly higher than algebra 2 students on 10 multiple-choice items taken from a practice version of a quantitative reasoning test developed by a prestigious university. Two analyses were applied to compare the performance of the groups: an analysis of covariance using grade 8 national percentile scores on standardized tests as a covariate and a “matched-group” large-sample Wilcoxon test. After establishing statistical controls over possible extraneous variables such as grade 8 CAT scores, sex, and ethnicity, the matched-group analysis confirmed the results obtained in the analysis of covariance. The overall differences in mean test scores and mean item difficulties obtained in the matched-group analysis were even larger than those obtained in the prior analyses.

On both analyses, males performed significantly better than females, although in the matched-group analysis the difference in mean test scores and item difficulties by sex were smaller than those obtained in the analysis of the full group. On both analyses, IMP students performed significantly better than algebra students on all items, including those requiring the application of statistical concepts such as the mean, the standard deviation, and normal curve properties.

The findings from this study strongly suggest that the IMP curriculum is successful in emphasizing reform-inspired topics such as probability and statistics. Overall, IMP Year 3 students demonstrated greater facility in applying probability and statistics concepts, interpreting data and analyzing graphs, and using rates and percentages to extrapolate than algebra 2 students.

References

Crosswhite, F. J., Dossey, J. A., Swafford, J. O., McKnight, C. C., Cooney, T. J., Downs, F. L., Grouws, D. A., & Weinzweig, A. 1. (1986). *Second International Mathematics Study, detailed report for the United States*. Champaign, IL: Stipes.

Dowling, M. & Webb, N. L. (1997a). *Comparison on statistics items of grade. 9 Interactive Mathematics Program (IMP) students with algebra students at one high school*. Project Report 97-2 from the Interactive Mathematics Program Evaluation Project. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Dowling, M. & Webb, N. L. (1997b). *Comparison on problem solving and reasoning of grade 10 Interactive Mathematics Program (IMP) students with geometry students at one high school*. Project Report 97-3 from the Interactive Mathematics Program Evaluation Project. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Illinois State Board of Education. (1996). *Preliminary draft - Illinois academic standards. English language arts, mathematics* Volume One: State goals 1-10. Springfield, IL: Illinois Academic Standards Project.

Interactive Mathematics Program. (1995). Standardized tests: HigMights from current studies of IMP student performance. *IMP Evaluation Update* [Newsletter], 1, Spring, p. 2.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Rosenstein, J. G., Caldwell, J. H., & Crown, W. D. (1996). *New Jersey Mathematics Curriculum Framework*. A collaborative effort of the New Jersey Mathematics Coalition and the New Jersey Department of Education. New Brunswick, NJ: The New Jersey Mathematics Coalition, Rutgers, The State University of New Jersey.

Schoen, H. L. (1993). Report to the National Science Foundation on the impact of the Interactive Mathematics Project (IMP). In N. L. Webb, H. Schoen, & S. D. Whitehurst (Eds.), *Dissemination of nine precollege mathematics instructional materials projects funded by the National Science Foundation, 1981-91*. Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Education Research.

South Carolina State Department of Education. (1993). *South Carolina Mathematics Framework*. Columbia, SC: Author.

Webb, N. L. & Dowling, M. (1995a). *Impact of the Interactive Mathematics Program on the retention of underrepresented students: Class of 1993 transcript report for school 1*: Brooks High School. Project Report 95-3 from the Interactive Mathematics Program Evaluation

Project. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Webb, N. L. & Dowling, M. (1995b). *Impact of the Interactive Mathematics Program on the retention of underrepresented students: Class of 1993 transcript report for school 2: Hill High School*. Project Report 95-4 from the Interactive Mathematics Program Evaluation Project. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Webb, N. L. & Dowling, M. (1995c). *Impact of the Interactive Mathematics Program on the retention of underrepresented students: Class of 1993 transcript report for school 3: Valley High School*. Project Report 95-5 from the Interactive Mathematics Program Evaluation Project. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Webb, N. L. & Dowling, M. (1996). *Impact of the Interactive Mathematics Program on the retention of underrepresented students: Cross-school analysis of transcripts for the class of 1993 for three high schools*. Project Report 96-2. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

List of Tables

List of Tables (Continued)

Page

Table 1

Distribution of Students by School and Course

Middle Schools	Total		IMP Year 3		Algebra 2	
	N	%	N	%	N	%
Public	59	46.8	38	41.8	21	60.0
Private	67	53.2	53	58.2	14	40.0
Total	126[a]	100.0	91[a]	100.0	35[a]	100.0

[a] Type of middle school was not reported for two IMP Year 3 and five Algebra 2 students.

Table 2

Distribution of Students by Sex, Course, and Teacher

Group	Total				IMP Year 3				Algebra 2			
	M	%[a]	F	%	M	%	F	%	M	%	F	%
Teacher A	14	35.0	26	65.0	-	-	-		14	35.0	26	65.0
Teacher B	11	45.8	13	54.2	11	45.8	13	54.2	-	-	-	-
Teacher C	13	52.0	12	48.0	13	52.0	12	48.0	-	-	-	-
Teacher B/ Teacher C	10	50.0	10	50.0	10	50.0	10	50.0	-	-	-	-
Teacher B/ Teacher D	13	54.2	11	45.8	13	54.2	11	45.8	-	-	-	-
Total	61	45.9	72	54.1	47	50.5	46	49.5	14	35.0	26	65.0

[a] Horizontal percentage.

Table 3

Distribution of Students by Ethnicity and Course

Ethnicity	Total		IMP Year 3		Algebra 2	
	N	%	N	%	N	%
Asian or Pacific Islander	11	8.3	6	6.5	5	12.5
Black – not of Hispanic Origin	31	23.3	19	20.4	12	30.0
Hispanic	4	3.0	3	3.2	1	2.5
White – not of Hispanic Origin	80	60.1	59	63.4	21	52.5
Other	7	5.3	6	6.5	1	2.5
Total	**133**	**100.0**	**93**	**100.0**	**40**	**100.0**

Table 4

Distribution of Students by Period, Course, and Teacher

Teacher	Period 1	Period 2	Period 3	Period 4	Period 5	Period 7	Total
	N	N	N	N	N	N	N
Teacher A (Algebra 2)	-	-	18	-	14	8	40
Teacher B (IMP Year 3)	-	-	-	24	-	-	24
Teacher C (IMP Year 3)	-	-	25	-	-	-	25
Teacher B/ Teacher C (IMP Year 3)	-	20	-	-	-	-	20
Teacher B/ Teacher D (IMP Year 3	24	-	-	-	-	-	24
Total	**24**	**20**	**43**	**24**	**14**	**8**	**133**

Table 5

Summary, Statistics on the Test by Course

	Total	IMP Year 3	Algebra 2
Number of students	133	93	40
Mean score	4.25	5.04	2.40
Standard deviation	2.30	2.12	1.52
Reliability (Alpha)	0 (3)	0 (1)	0 (2)
Standard error of measurement	1.36	1.40	1.25
Mean item difficulty	.42	.50	.24
Maximum attained score [a]	10 (2)	10 (2)	8 (1)
Minimum attained score [a]	0 (3)	0 (1)	0 (2)

[a] The number of students is indicated in parenthesis.

Table 6

Item Difficulty on the Test by Course

Item	Total	IMP Year 3	Algebra 2
1	.14	.18	.03
2	.46	.54	.28
3	.23	.29	.10
4	.43	.52	.23
5	.41	.54	.10
6	.33	.43	.10
7	.68	.75	.53
8	.58	67	.38
9	.36	.42	.23
10	.36	.71	.45
Total no. of Students	**133**	**93**	**40**

Table 7

Mean and Standard Deviation on Grade 8 Standardized Tests Scores by Course and Middle School

Middle School	Total			IMP Year 3			Algebra 2		
	M	S.D.	N	M	S.D.	N	M	S.D.	N
Public Middle School									
Mathematics [a]	92.31	4.46	59	92.95	4.36	38	91.14	4.51	21
English [a]	91.14	7.82	59	92.87[c]	5.80	38	88.00	9.97	21
Private Middle School									
Mathematics [b]	95.54	3.73	67	95.66	3.70	53	95.07	3.97	14
English	95.24	4.24	67	95.26	4.33	53	95.14	4.04	14
Total									
Mathematics	94.03	4.38	126	94.53[d]	4.18	91	92.71	4.67	35
English	93.32	6.49	126	94.26[e]	5.10	91	90.86	8.79	35

[a] California Achievement Test (CAT) in mathematics taken in grade 8.

[b] Students who attended a private middle school took a variety of standardized tests. Approximately 84% of the students took California Achievement Test (CAT). The remaining 16% took one or more of the following tests: Comprehensive Test of Basic Skills (CTBS), the Stanford Achievement Test (SAT), the Iowa Test of Basic Skills (IOWA), and the Educational Record Bureau (ERB) test.

[c] Statistically significant differences between IMP and algebra 2 using an analysis of variance ($p < .021$).

[d] Statistically significant differences between IMP and algebra 2 using an analysis of variance ($p < .037$).

[e] Statistically significant differences between IMP and algebra 2 using an analysis of variance ($p < .008$).

Table 8

Analysis of Variance: Total Test Score by Grade 11 Course-Public Middle School

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Sig of F
Covariate: Grade 8 Scores in Math	16.793	1	16.793	4.135	.0470
Main Effects Course	121.456	1	121.456	29.908	.0001
Explained	138.249	2	69.124	17.022	.0001
Residual	227.412	56	4.061		
Total	365.661	58	6.305		

Table 9

Analysis of Variance: Total Test Score by Grade 11 Course-Private Middle School

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Sig of F
Covariate: Grade 8 Scores in Math	43.354	1	43.354	13.056	.0010
Main Effects Course	55.801	1	55.801	16.804	.0001
Explained	99.154	2	49.577	14.930	.0001
Residual	212.518	64	3.321		
Total	311.672	66	4.722		

Table 10

Analysis of Variance: Total Test Score by Grade 11 Course-Total Group

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**F**	**Sig of F**
Covariate: Grade 8 Scores in Math	50.780	1	50.780	13.413	.0001
Main Effects Course	161.210	1	161.210	45.581	.0001
Explained	211.990	2	105.995	27.997	.0001
Residual	465.669	123	3.786		
Total	677.659	125	5.421		

Table 11

Summary Statistics on the Test by Sex

	Male	**Female**
Number of Students	61	72
Mean Score	4.93[a]	3.67[a]
Standard Deviation	2.25	2.19
Reliability (Alpha)	.61	.65
Standard error of measurement	1.41	1.30
Mean item difficulty	.49	.37
Maximum attained [b]	10 (2)	9 (1)
Minimum attained [b]	1 (4)	0 (3)

[a] Statistically significant differences were found using an analysis of variance ($p < .016$)

[b] The number of students is indicated in parenthesis.

Table 12

Mean and Standard Deviation by Ethnicity and Course

	Total			IMP Year 3			Algebra 2		
	N	**Mean**	**S.D.**	**N**	**Mean**	**S.D.**	**N**	**Mean**	**S.D.**
Asian or Pacific Islander	11	4.00	2.37	6	5.67	1.63	5	2.00	1.22
Black – not of Hispanic Origin	31	3.32	2.04	19	4.16	2.03	12	2.00	1.21
Hispanic	4	4.25	2.63	3	4.67	3.06	1	3.00	-
White – not of Hispanic Origin	80	4.59	2.27	59	5.27	2.06	21	2.67	1.77
Other	7	4.86	2.67	6	5.17	2.79	1	3.00	-
Total	133	4.25	2.30	93	5.04	2.12	40	2.40	1.52

Table 13

Analysis of Variance: Total Test Score by Grade 11 Course, Sex, and Ethnicity

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Sig of F
Main Effects	237.567	6	39.595	10.884	.0001
Course	158.370	1	158.370	43.535	.0001
Sex	21.731	1	21.731	5.974	.016
Ethnicity	13.654	4	3.413	.938	.444
2-Way Interaction Course by Ethnicity	5.527	4	1.382	.380	.823
Course by Sex	.009	1	.009	.002	.960
Sex by Ethnicity	34.102	4	8.526	2.344	.059
3-Way Interaction Course by Sex by Ethnicity	3.921	2	1.960	.539	.585
Residual	418.346	115	3.638		
Total	696.812	132	5.279		

Table 14

Total Test Score Means and Standard Deviations by Period, Course, and Teacher [a]

Teacher	Period 1		Period 2		Period 3		Period 4		Period 5		Period 7	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Teacher A (Algebra 2)	-	-	-	-	2.22	1.17	-	-	2.21	1.25	3.13	2.42
Teacher B (IMP Year 3)	-	-	-	-	-	-	5.04	2.03	-	-	-	-
Teacher C (IMP Year 3)	-	-	-	-	4.44	2.49	-	-	-	-	-	-
Teacher B/ Teacher C (IMP Year 3)	-	-	5.25	1.41	-	-	-	-	-	-	-	-
Teacher B/ Teacher D (IMP Year 3)	5.50	2.25	-	-	-	-	-	-	-	-	-	-

[a] all pairwise comparisons of means between IMP classes of algebra 2 classes were statistically significant except for teacher C-period 3 (IMP) and teacher A-period 7 (algebra 2) using an analysis of variance.

Table 15

Analysis of Variance: Total Test Score by IMP Year 3 Class Periods

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Sig of F
Main Effects Course	14.960	3	4.987	1.118	.346
Explained	14.760	3	4.987	1.118	.346
Residual	396.868	89	4.459		
Total	411.828	92	4.476		

Table 16

Analysis of Variance: Total Test Score by Algebra 2 Class Periods

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Sig of F
Main Effects Course	5.257	2	2.628	1.153	.327
Explained	5.257	2	2.628	1.153	.327
Residual	84.343	37	2.280		
Total	89.600	39	2.297		

Table 17

Distribution of Students by School and Course (Matched-Group Analysis)

Middle School	Total		IMP Year 3		Algebra 2	
	N	%	N	%	N	%
Public	31	50.0	14	45.2	17	54.8
Private	31	50.0	17	54.8	14	45.2
Total	62	100.0	31	100.0	31	100.0

Table 18

Distribution of Students by Sex and Course (Matched-Group Analysis)

Sex	Total		IMP Year 3		Algebra 2	
	N	%	N	%	N	%
Female	40	64.5	19	61.3	21	67.7
Male	22	35.5	12	38.7	10	32.3
Total	62	100.0	31	100.0	31	100.0

Table 19

Distribution of Students by Ethnicity and Course (Matched-Group Analysis)

Ehnicity	Total		IMP Year 3		Algebra 2	
	N	%	N	%	N	%
Asian or Pacific Islander	4	6.5	1	3.2	3	9.7
Black-not of Hispanic Origin	18	29.0	8	25.8	10	32.3
Hispanic	2	3.2	1	3.2	1	3.2
White-not of Hispanic Origin	33	53.2	17	54.9	16	51.6
Other	5	8.1	4	12.9	1	3.2
Total	62	100.0	31	100.0	31	100.0

Table 20

Large-Sample Matched-Group Wilcoxon Test Results [1]

Treatment: IMP Year 3 Vs. Algebra 2
Observed or Outcome Variable: Total Test Score

RESULTS	
E (T)	248.00
VARIANCE (T)	2586.50
STANDARD DEVIATION	50.86
$P_{.025}$	147.81
$P_{.975}$	348.19
T	481.40

$H_0: M_d = 0$
$T > T_u$

∴ *therefore*

Reject H_0

Z-Statistic Approximation, continuity corrected

Z = 4.5814 2-Tail p value = 0.0001** (Highly Significant)

Limits for the two-tailed Confidence Interval (95%)

$2 \leq M_d \leq 4$

[1] The statistical package NONPAR developed by Dr. Ronald Serlin from the University of Wisconsin-Madison was used to perform the Wilcoxon test.

Table 21

Summary Statistics on the Test by Course
(Matched-Group Analysis)

	Total	IMP Year 3	Algebra 2
Number of students	62	31	31
Mean score	3.90	5.42	2.39
Standard deviation	2.45	2.20	1.61
Reliability (Alpha)	.71	.60	.41
Standard error of measurement	1.32	1.39	1.24
Mean item of difficulty	.39	.54	.24
Maximum attained score [a]	10 (1)	10 (1)	8 (1)
Minimum attained score [a]	0 (2)	0 (1)	0 (1)
Grade 8: *California Achievement Test* (CAT)[b]	93.69 (3.73)	93.71 (3.75)	93.68 (3.77)

[a] The number of students is indicated in parenthesis.

[b] The standard deviation is indicated in parenthesis.

Table 22

Summary Statistics on the Test by Sex (Matched-Group Analysis)

	Male	Female
Number of students	22	40
Mean score	4.41	3.63
Standard deviation	2.56	2.37
Reliability (Alpha)	.70	.71
Standard error of measurement	1.40	1.28
Mean item of difficulty	.44	.36
Maximum attained score [a]	10 (1)	9 (1)
Minimum attained score [a]	1 (4)	0 (2)

[a] The number of students is indicated in parenthesis.

Table 23

Item Difficulty on the Test by Course (Matched-Group Analysis)

Item	Total	IMP Year 3	Algebra 2
1	.13	.23	.03
2	.47	.68	.26
3	.26	.39	.13
4	.36	.45	.26
5	.34	.58	.10
6	.23	.39	.07
7	.60	.68	.52
8	.53	.71	.36
9	.36	.48	.23
10	.65	.84	.45
Total no. of Students	62	31	31

Test

To the student:

This test consists of 10 multiple-choice questions. Calculators are allowed. If it is necessary, the test administrator will provide you will additional paper to work the problems.
Before beginning the test please complete all of the information requested on the cover sheet.

Please circle your choice to each test question in the booklet. Try to answer every questions, but do not spend too much time on any one question. If you have time at the end of the test, go back and answer any questions that you have skipped. If you are not sure of an answer, circle the answer you think is right.

If you want to change an answer, erase your first mark completely. Then circle the choice you think is correct. Circle only one answer for each question.

You may start answering the questions as soon as you complete your name and the other information requested on the cover sheet. Please write your name on the top of each page.

May 1996

Cover Sheet

Student's Name: ____

Grade: ____

Mathematics Course Title: ____________________________________

Class Period: ____

Teacher: ____

Gender: (circle one) F M

(optional)

Ethnicity:
(circle one)

(1) American Indian or Alaskan

(2) Asian or Pacific Islander

(3) Black (not of Hispanic origin)

(4) Hispanic

(5) White (not of Hispanic origin)

(6) Other

May 1996

Student Name: ______________________

The oil consumption of Town A in the year 1985 was 60,000 barrels. The consumption increases at a rate of 6% per year. (Use this information to answer questions 4 to 6.)

4. By what percent did Town A's oil consumption increase from 1985 to 1988?

   a. 6%
   b. 12%
   c. 18%
   d. 19%
   e. 26%

5. Assuming that oil consumption continues to increase by 6% per year, approximately how many years will it take for the town's annual oil consumption to reach 240,000 barrels?

   a. 12 years
   b. 24 years
   c. 36 years
   d. 46 years
   e. 50 years

6. The oil consumption for a nearby metropolitan area, City B, increases at a rate of 10,000 barrels each year. City B used 200,000 barrels in 1985. Assume that oil consumption continues to increase indefinitely for both Town A and City B at the given rates. Which of the following statements is true?

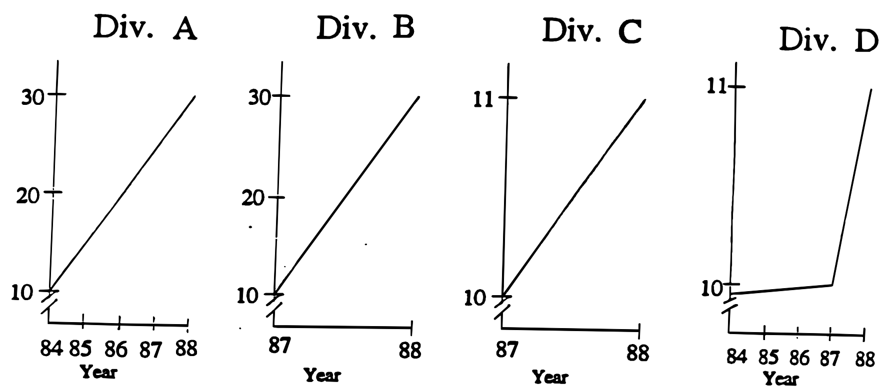
   a. Town A will consume more oil per year that City B by the year 2005.
   b. Town A will eventually consume more oil per year than City B, but after the year 2005.
   c. Town A's annual oil consumption will always remain under half of City B's.
   d. Town A's annual oil consumption will never exceed City B's, but will come within 10,000 barrels of City B's consumption.
   e. None of the above is true.

Student Name: ____________________

9. The head of Division A in a large corporation was pleased with the revenue growth in his division over the past several years. In a meeting with his supervisor, he presented her a graph of the revenue in his division for the years 1984 to 1988, inclusive, pointing out that the 1984 revenue was $10 million and that the revenue had increased linearly until it reached $30 million in 1988. (See the graph for Division A below.) Which of the following is the best interpretation of the slope of his graph?

   a. The average yearly revenue for the period from 1984 to 1988.
   b. The total revenue for the period from 1984 to 1988.
   c. The increase for the period from 1984 to 1988.
   d. The yearly increase in revenue.
   e. The yearly percent increase in revenue.

10. The supervisor wanted to see how the other three divisions under her control compared to the division she had just seen. She asked each of the other three division leaders to submit a graph of the revenue in his or her division. She now has the follow graphs.



(all revenue given in millions of dollars)

Which of the following statements is true?

a. During the period 1987-1988, the revenues in Division D increased at a rate greater than the rate of increase in Division C.
b. During the period 1987-1988, the revenues in Division B increased at a rate less than the rate of increase in Division A.
c. During the period 1987-1988, the rate of increase of revenue was the same in Division B and Division C.
d. (a), (b), and (c) are all true.
e. None of (a), (b), and (c) are true.