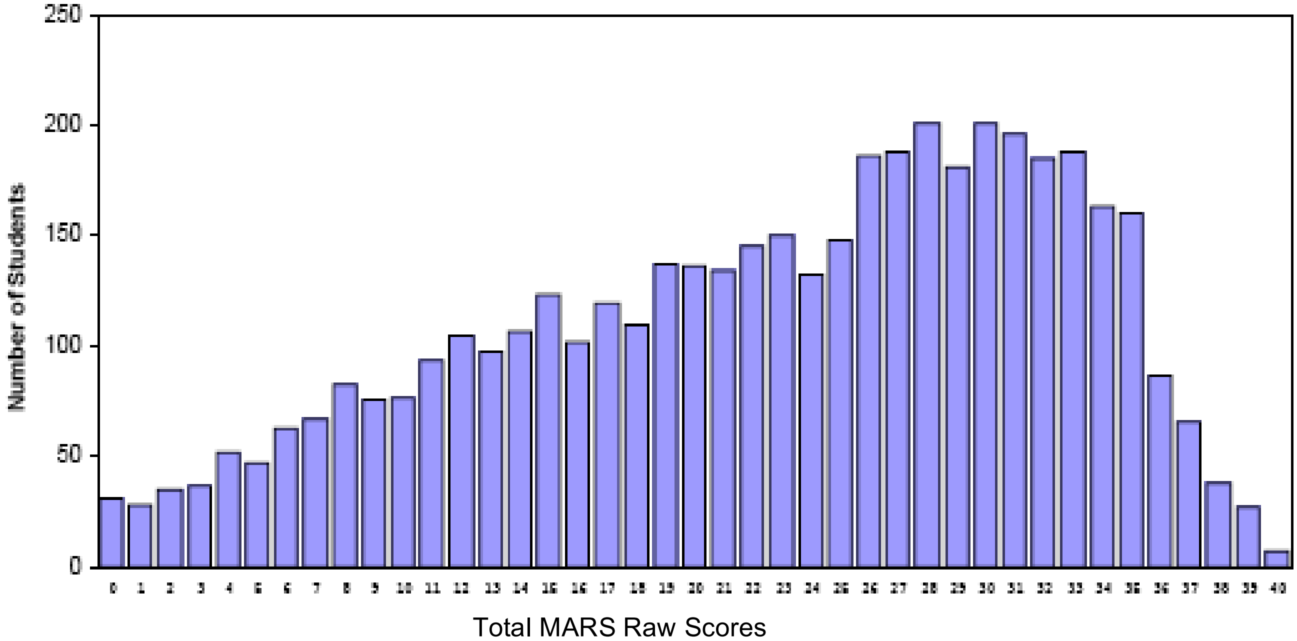
Overall Frequency Distribution by Total MARS Raw Scores, Course 1

Mean: 22.68 StdDev: 9.48



MARS Test Performance Level Frequency Distribution Table and Bar Graph

2005 – Number of Students tested in Course 1: 4512

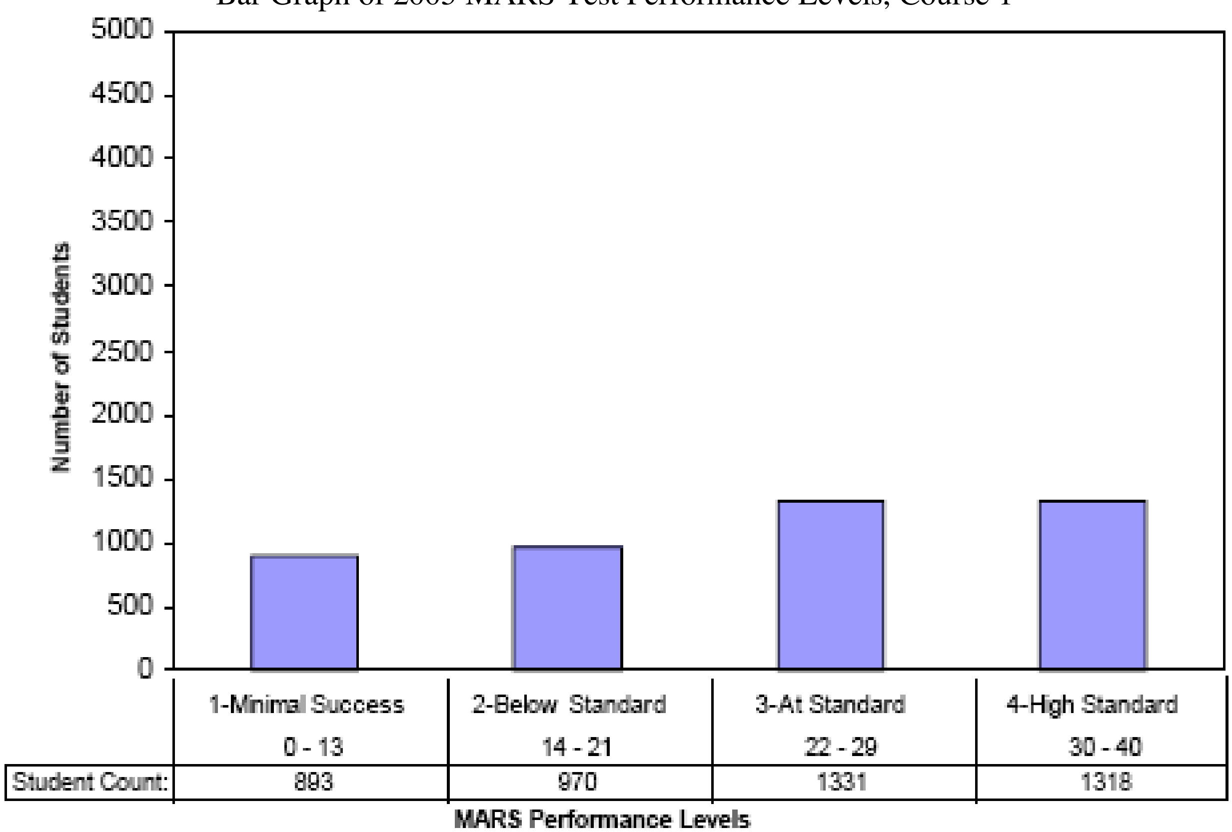
Frequency Distribution of MARS Test Performance Levels, Course 1

Year of Testing

	2000		2001		2002	
Perf. Level	% at	% at least	% at	% at least	% at	% at least
1	N/A	N/A	22%	100%	18%	100%
2	N/A	N/A	62%	78%	61%	82%
3	N/A	N/A	15%	16%	20%	22%
4	N/A	N/A	1%	1%	2%	2%

	2003		2004		2005	
Perf. Level	% at	% at least	% at	% at least	% at	% at least
1	66%	100%	39%	100%	20%	100%
2	29%	34%	39%	61%	21%	80%
3	5%	6%	19%	23%	29%	59%
4	1%	1%	3%	3%	29%	29%

Bar Graph of 2005 MARS Test Performance Levels, Course 1



*Total MARS raw scores is the summation of Tasks 1 through 5 on the MARS test.

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. Groups with Ns of less than 5 students are not reported.

Figure 8.1 Box and whisker plot of Total MARS Raw Scores by Ethnicity

Course: 1

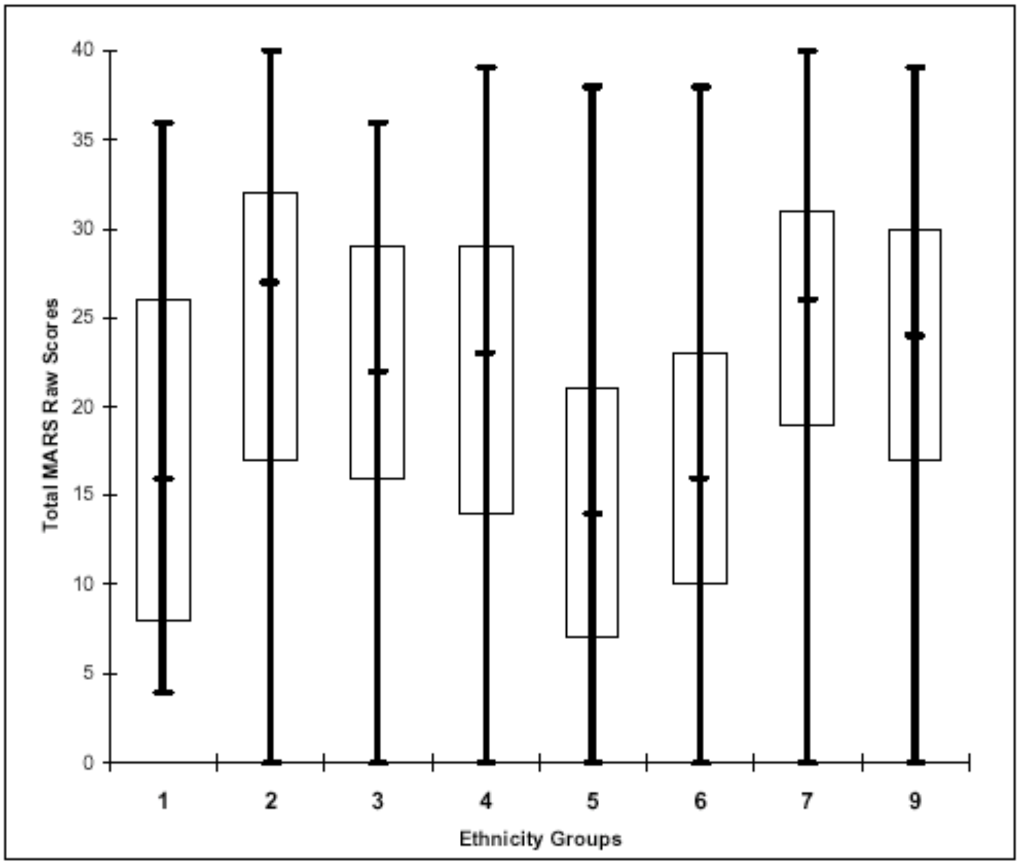


Table 8.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
1	American Indian	23
2	Asian/Asian American	1264
3	Pacific Islander	54
4	Filipino	346
5	Hispanic/Latino	1772
6	African American	333
7	White (Not Hispanic)	2060
9	Others/Unknown	123

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

Figure 19.1 Distribution of sampling means by Ethnicity

Course: 1

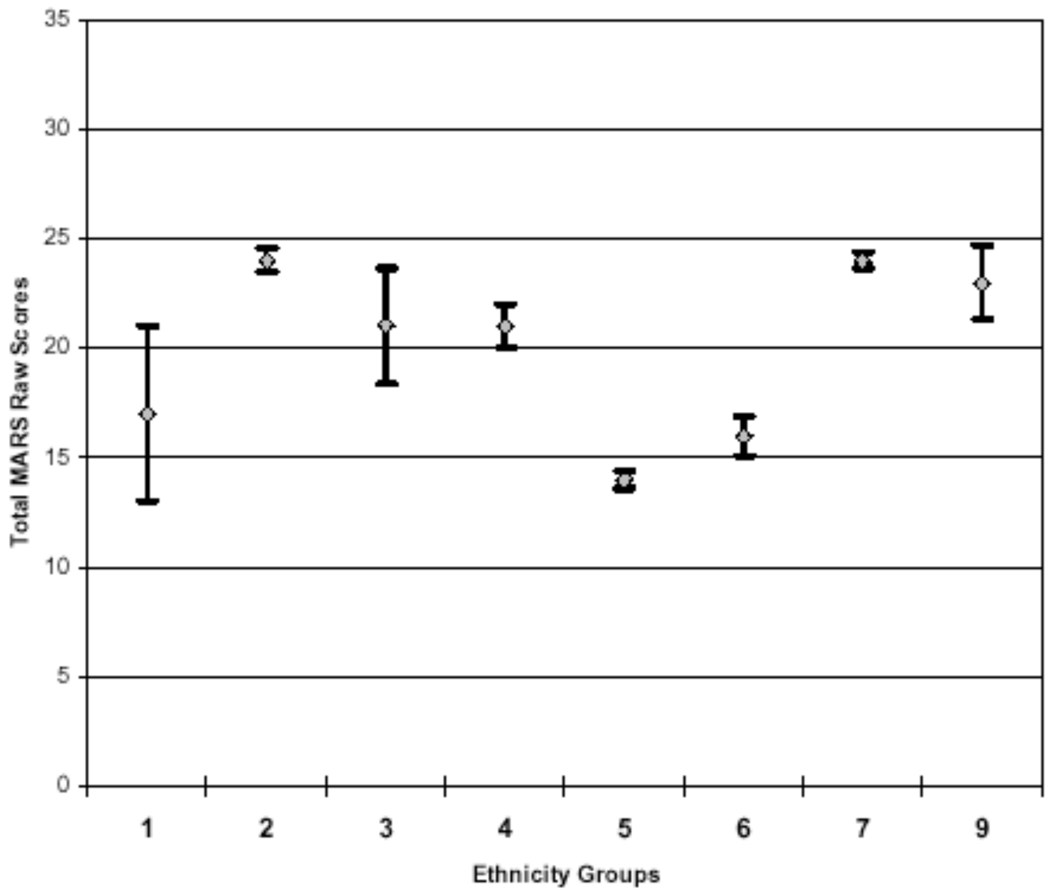


Table 19.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
1	American Indian	23
2	Asian/Asian American	1264
3	Pacific Islander	54
4	Filipino	346
5	Hispanic/Latino	1772
6	African American	333
7	White (Not Hispanic)	2060
9	Others/Unknown	123

Distribution of sampling means
Course 1 – All Students
Ethnicity

In this section, test scores are compared across different ethnic groups[36]. 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 Indian/Alaskan Native students are significantly lower than the scores of the Asian/Asian American and White students, and not significantly different from those of the students in any other ethnic group.

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

The scores of Pacific Islanders are significantly higher than the scores of Hispanic/Latino and African American students, but there are no significant differences between Pacific Islander students and any other group.

The scores of Filipino students are significantly lower than the scores of Asian/Asian American and White students, and higher than the scores of the Hispanic/Latino and African American students.

The scores of Hispanic/Latino students are significantly lower than the scores of students in any other ethnic group except Indian/Alaskan Native (not a significant difference).

The scores of African American students are significantly lower than the scores of students in all other ethnic groups except for Hispanic/Latino and Indian/Alaskan Native students. The scores of African American students are significantly higher than scores of Hispanic/Latino students, and not significantly different from those of the Indian/Alaskan Native students.

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

---

[36] 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.

Figure 8.2 Box and whisker plot of Total MARS Raw Scores by Home Language

Course: 1

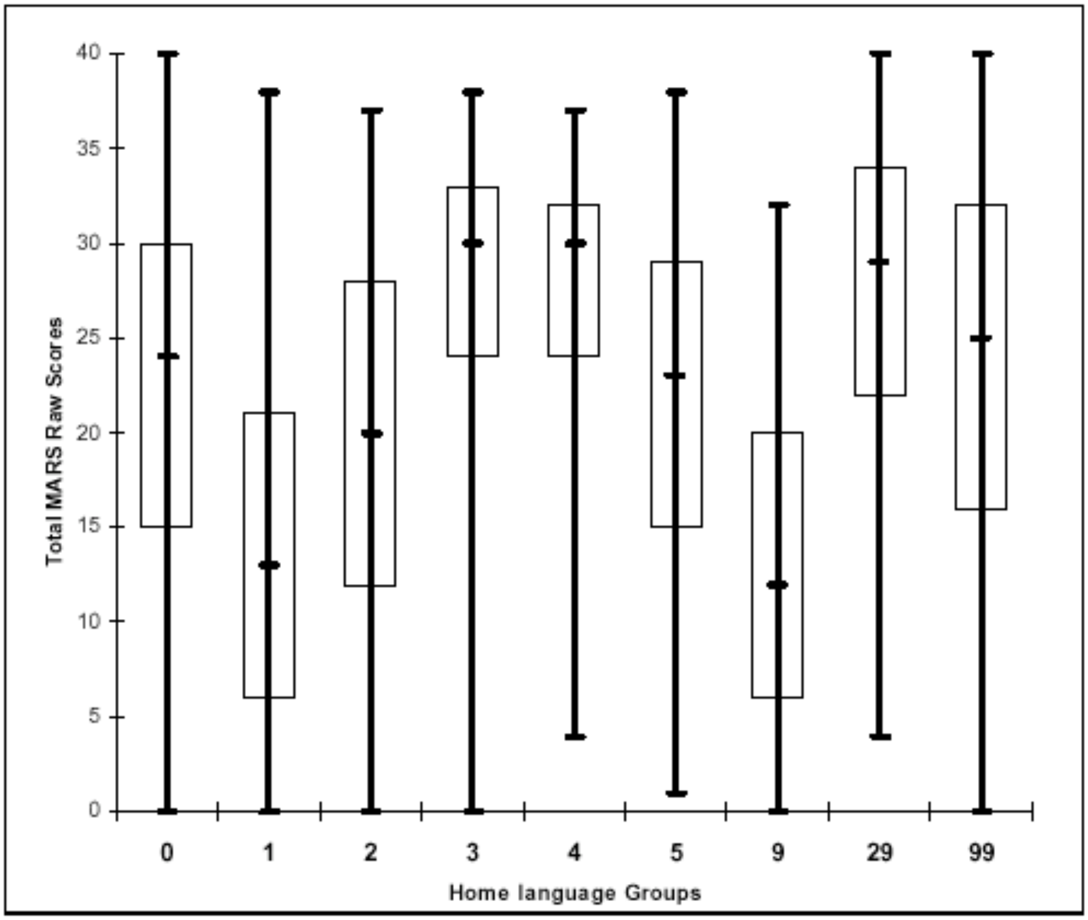


Table 8.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	3440
1	Spanish	1205
2	Vietnamese	317
3	Cantonese	112
4	Korean	56
5	Filipino	150
9	Cambodian	41
29	Russian	25
99	Others/Unknown	629

Figure 19.2 Distribution of sampling means by Home Language

Course: 1

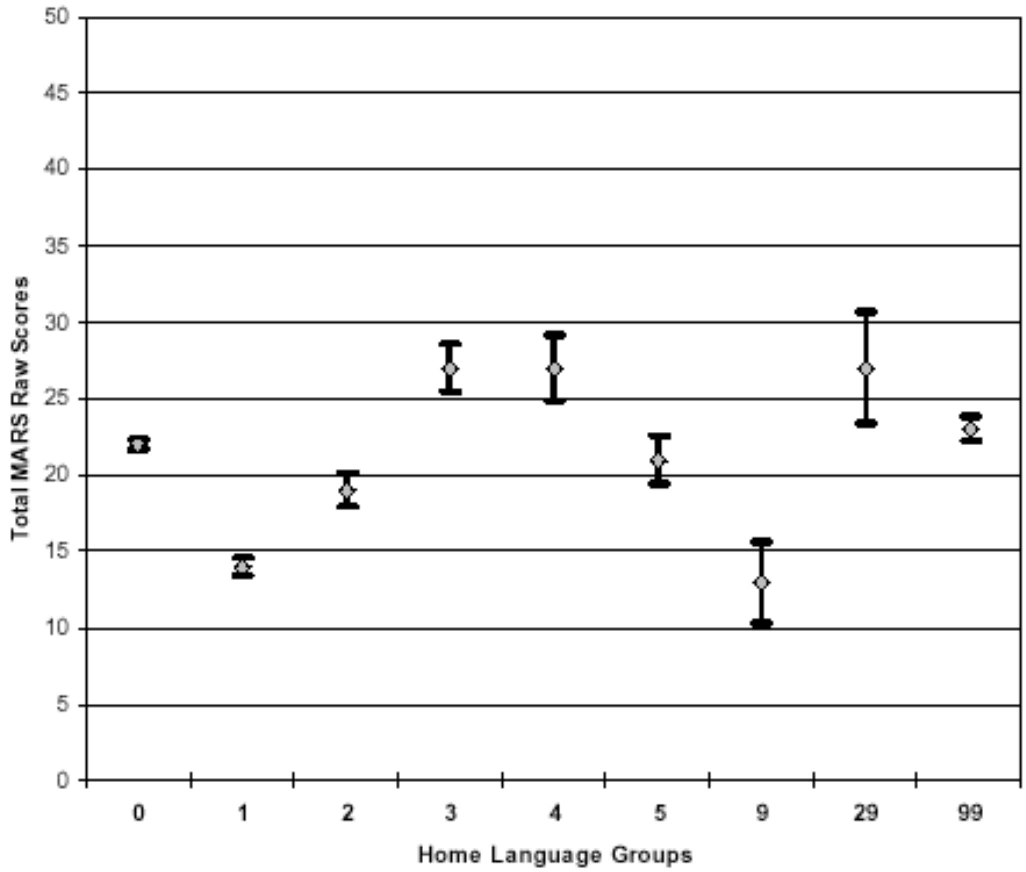


Table 19.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	3440
1	Spanish	1205
2	Vietnamese	317
3	Cantonese	112
4	Korean	56
5	Filipino	150
9	Cambodian	41
29	Russian	25
99	Others/Unknown	629

Distribution of sampling means
Course 1 – All Students
Home Language

In this section, test scores are compared across groups of students who speak different languages at home[37]. 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 higher than the scores of Spanish, Vietnamese and Khmer students, and lower than those of the Cantonese and Korean home language groups. There are no other significant differences between students with English as a home language and any other category.

The scores of students with Spanish as a home language are significantly lower than the scores of students in any other home language group except Khmer (not significantly different).

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 English, Cantonese, Korean, Russian or "Other" as a home language. There is no significant difference between students with Vietnamese as a home language and the Filipino home language group.

The scores of students with Cantonese as a home language are significantly higher than scores of students in any other home language group except Korean and Russian (not significantly different).

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

Students with Filipino as a home language scored significantly higher than Spanish and Khmer speaking students, and lower than students with Cantonese or Korean as a home language. There are no significant differences between students with Filipino as a home language and any other home language group.

The scores of students with Khmer as a home language are significantly higher than the scores of students in any other home language group except Spanish (not a significant difference).

---

[37] 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 Russian as a home language are significantly higher than the scores of students with Spanish, Vietnamese or Khmer as a home language. There are no significant differences between students with Russian as a home language and any other home language group.

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

Figure 8.3 Box and whisker plot of Total MARS Raw Scores by Parent Education

Course: 1

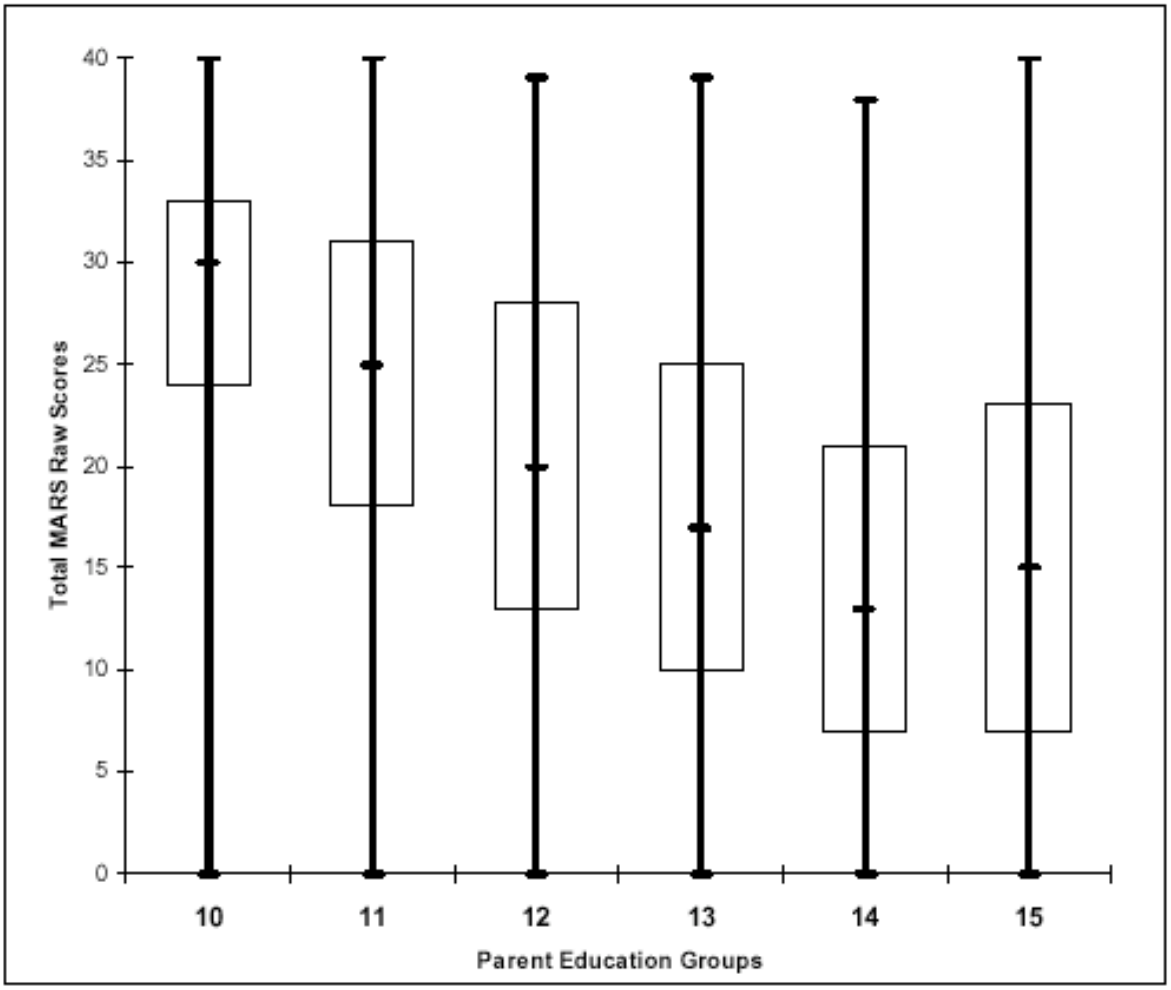


Table 8.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	1007
11	College graduate	1473
12	Some college	1077
13	High School graduate	1018
14	Not a high school graduate	605
15	Others/Unknown	795

Figure 19.3 Distribution of sampling means by Parent Education

Course: 1

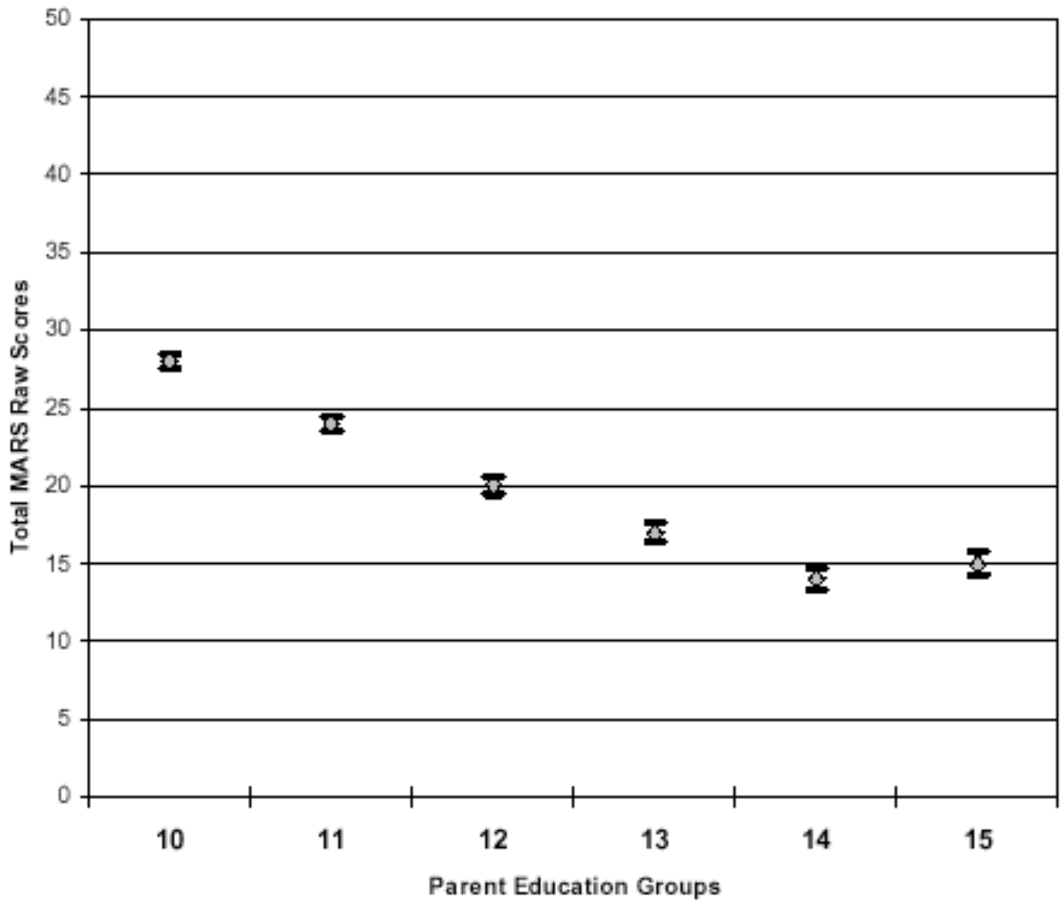


Table 19.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	1007
11	College graduate	1473
12	Some college	1077
13	High School graduate	1018
14	Not a high school graduate	605
15	Others/Unknown	795

Distribution of sampling means
Course 1 – All Students
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 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 in all other Parent Education categories.

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 scores of students whose parents are not high school graduates or whose education level is "Unknown," 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 except "Unknown" (not a significant difference).

The scores of students whose level of parent education is unknown are significantly lower than the scores of students whose parents are in any other category except "not high school graduates" (not a significant difference).

[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.

Figure 8.4 Box and whisker plot of Total MARS Raw Scores by Gender

Course: 1

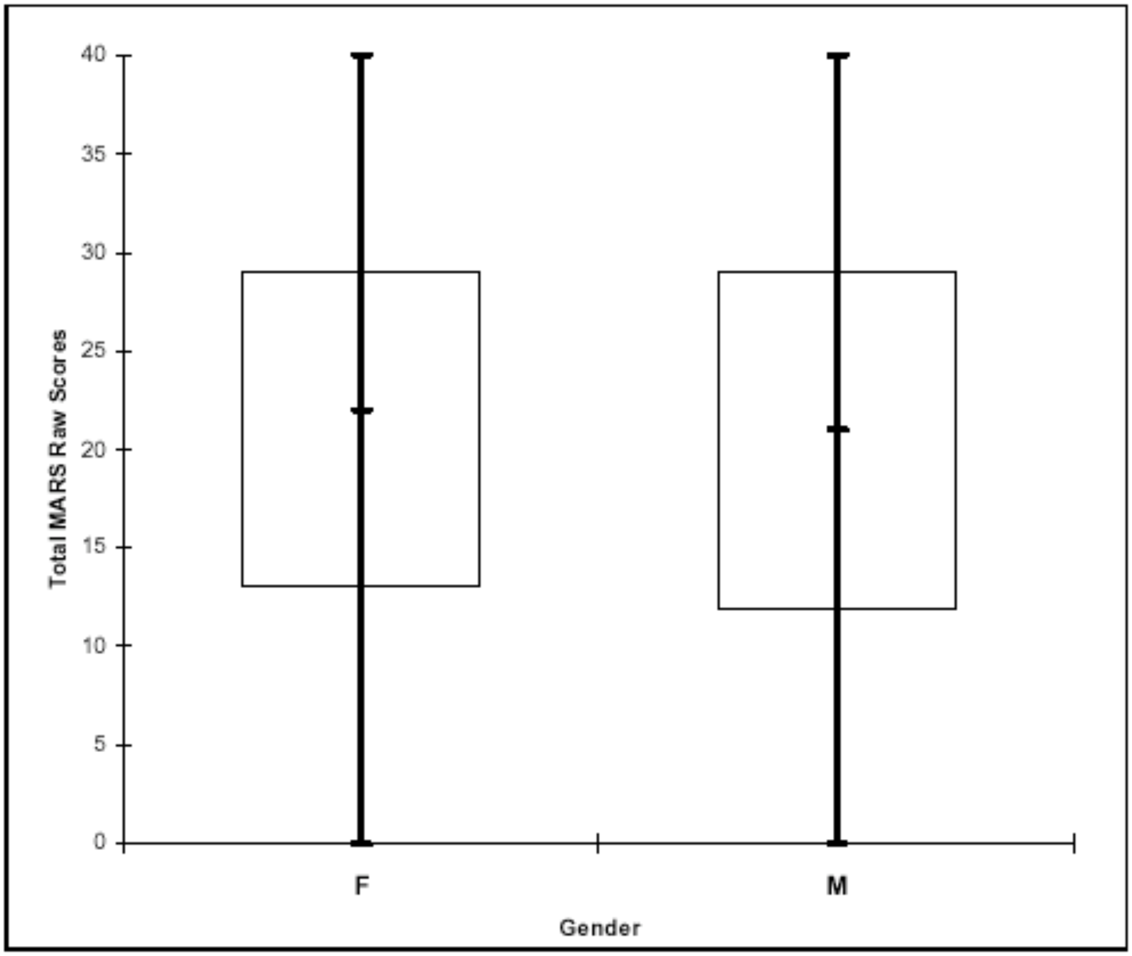


Table 8.4 Student count for Gender

Gender	Student Count
Female	2917
Male	3053

Figure 19.4 Distribution of sampling means by Gender

Course: 1

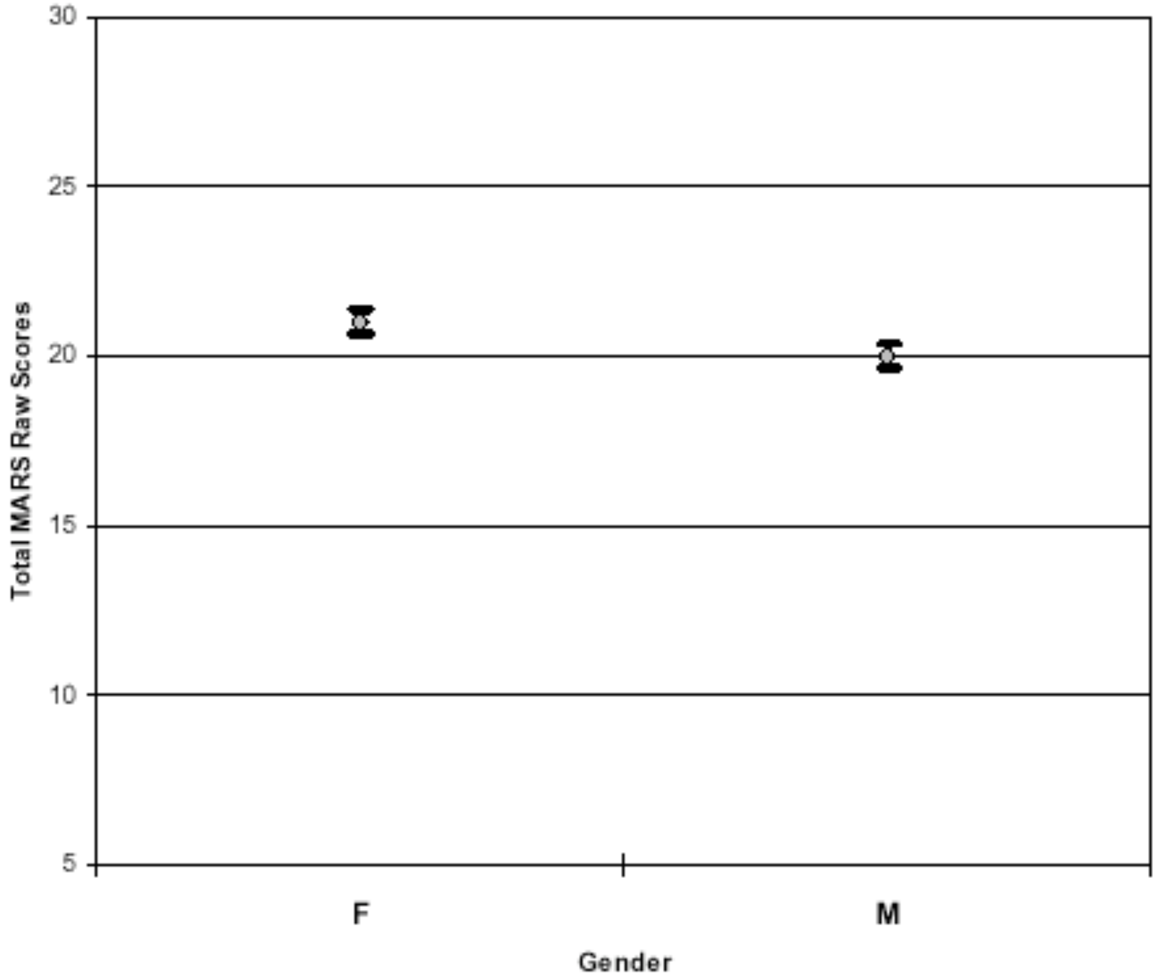


Table 19.4 Student count for Gender

Gender	Student Count
Female	2917
Male	3053

Distribution of sampling means
Course 1 – All Students
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.

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

Figure 8.5 Box and whisker plot of Total MARS Raw Scores by Language Fluency

Course: 1

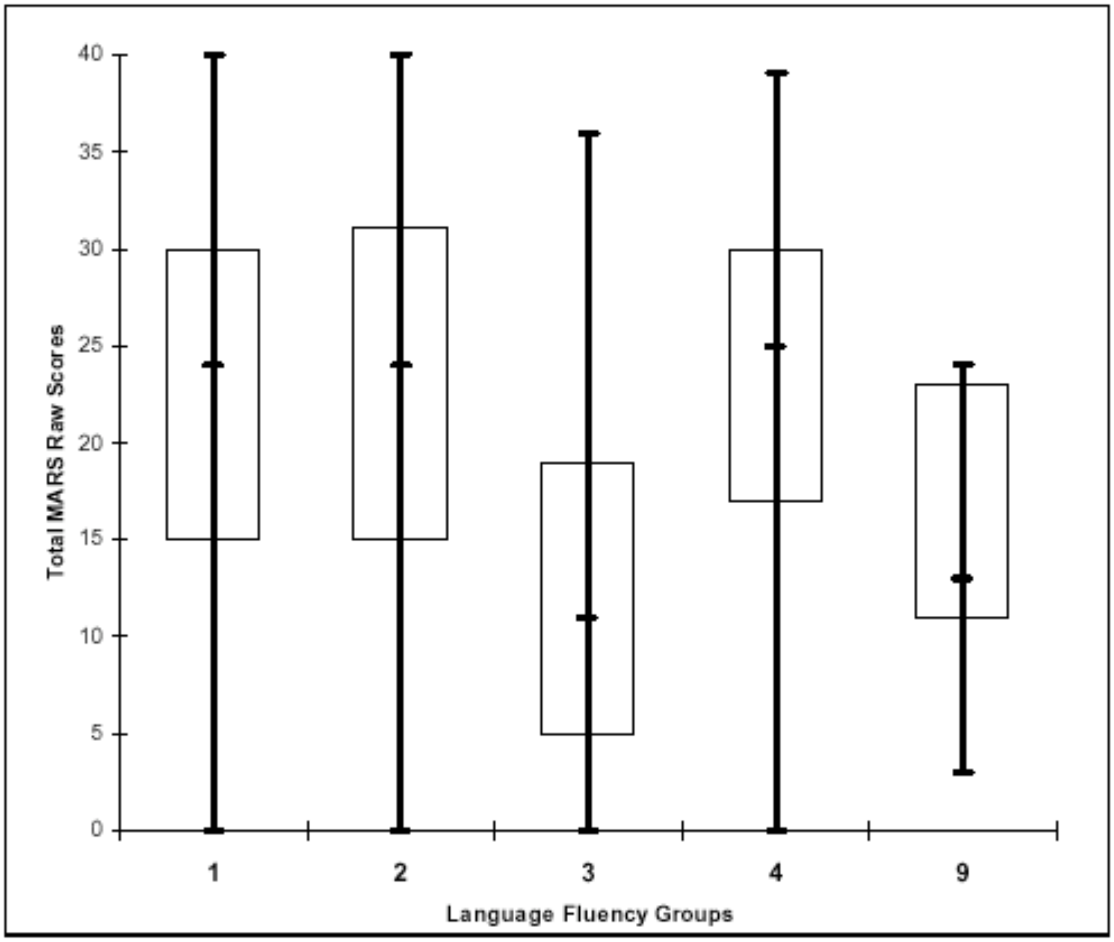


Table 8.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	3450
2	Initially Fluent (I-FEP)	776
3	English Learner	1082
4	ReDesignated (R_FEP)	662
9	Unknown	5

Figure 19.5 Distribution of sampling means by Language Fluency

Course: 1

Language Fluency

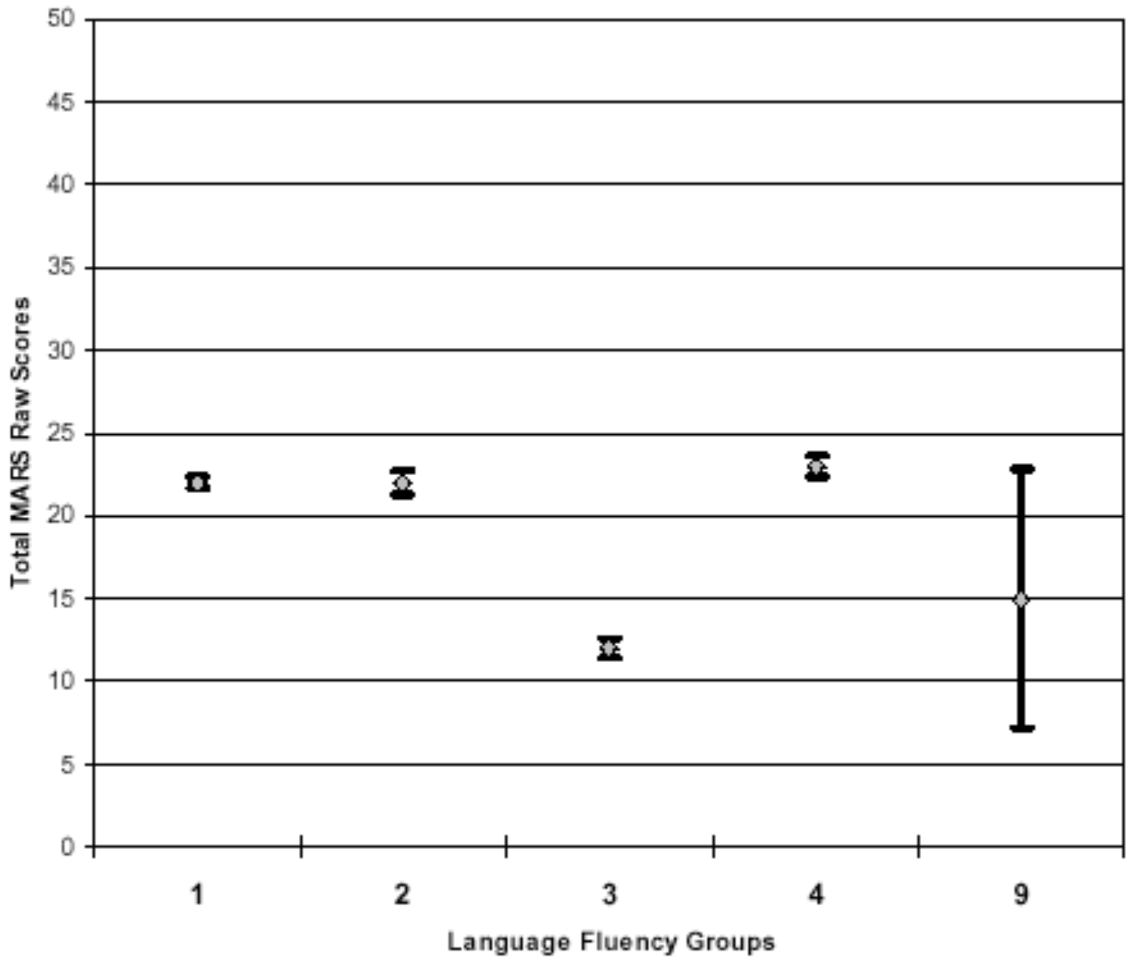


Table 19.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	3450
2	Initially Fluent (I-FEP)	776
3	English Learner	1082
4	ReDesignated (R_FEP)	662
9	Unknown	5

Distribution of sampling means
Course 1 – All Students
Language Fluency

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.

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 except "Unknown" (not a significant difference).

The scores of students in the Re-designated 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 Re-designated FEP category and any other group.

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. Groups with Ns of less than 5 students are not reported.

Figure 10.1 Box and whisker plot of Total MARS Raw Scores by Ethnicity

Course 1 - Grade 8 and Lower grades

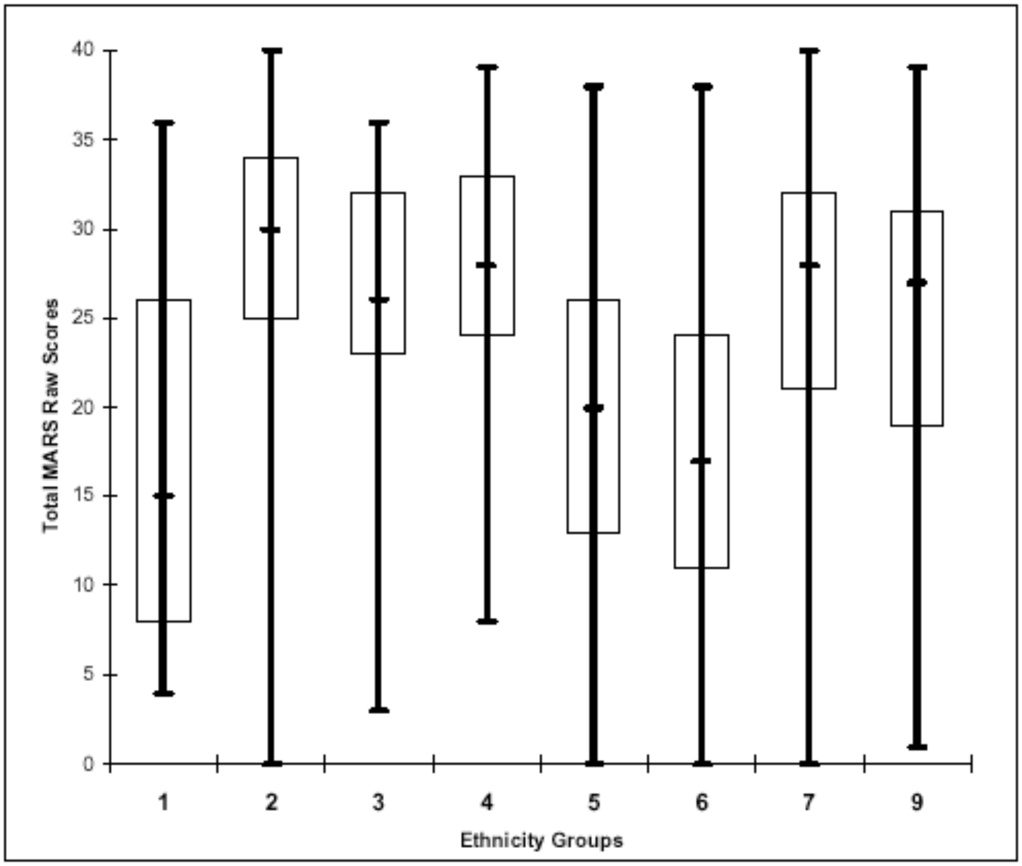


Table 10.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
1	American Indian	13
2	Asian/Asian American	777
3	Pacific Islander	29
4	Filipino	145
5	Hispanic/Latino	502
6	African American	225
7	White (Not Hispanic)	1600
9	Others/Unknown	105

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

Figure 21.1 Distribution of sampling means by Ethnicity

Course 1 - Grade 8 and Lower grades

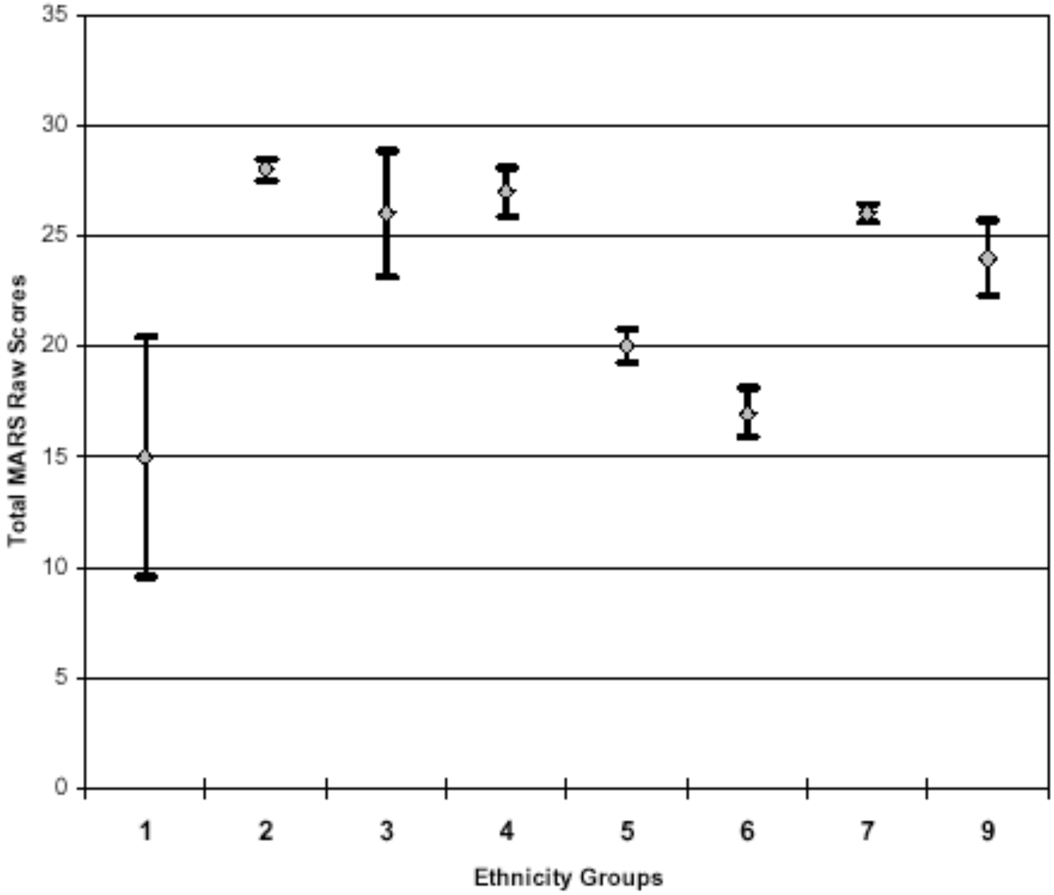


Table 21.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
1	American Indian	13
2	Asian/Asian American	777
3	Pacific Islander	29
4	Filipino	145
5	Hispanic/Latino	502
6	African American	225
7	White (Not Hispanic)	1600
9	Others/Unknown	105

Distribution of sampling means
Course 1 – Grade 8 and Lower
Ethnicity

In this section, test scores are compared across different ethnic groups[41]. 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 Indian/Alaskan Native students are significantly lower than the scores of students in all other groups except Hispanic/Latino and African American (not significant differences).

The scores of Asian/Asian American students are significantly higher than the scores of students in all other groups except Pacific Islander and Filipino (not significant differences).

The scores of Pacific Islanders are significantly higher than the scores of Indian/Alaskan Native, Hispanic/Latino and African American students, but there are no significant differences between Pacific Islander students and any other group.

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

The scores of Hispanic/Latino students are significantly lower than the scores of students in any other ethnic groups except Indian/Alaskan Native (not significantly different) and African American. The African American student scores were significantly lower than the scores of Hispanic/Latino students.

The scores of African American students are significantly lower than the scores of students in all other ethnic groups except for Indian/Alaskan Native students (not a significant difference).

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

[41] 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.

Figure 10.2 Box and whisker plot of Total MARS Raw Scores by Home Language

Course 1 - Grade 8 and Lower grades

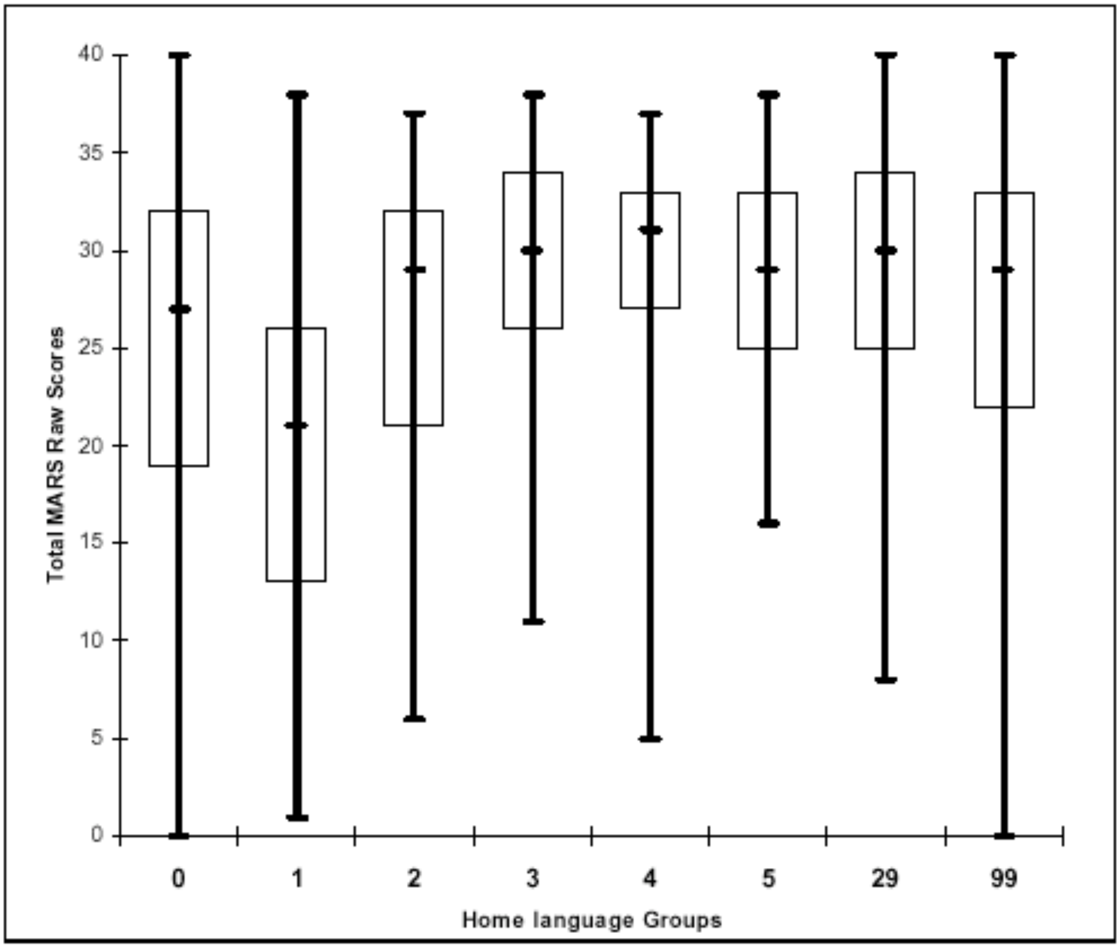


Table 10.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	2379
1	Spanish	294
2	Vietnamese	86
3	Cantonese	84
4	Korean	48
5	Filipino	64
29	Russian	22
99	Others/Unknown	419

Figure 21.2 Distribution of sampling means by Home Language
Course 1 - Grade 8 and Lower grades

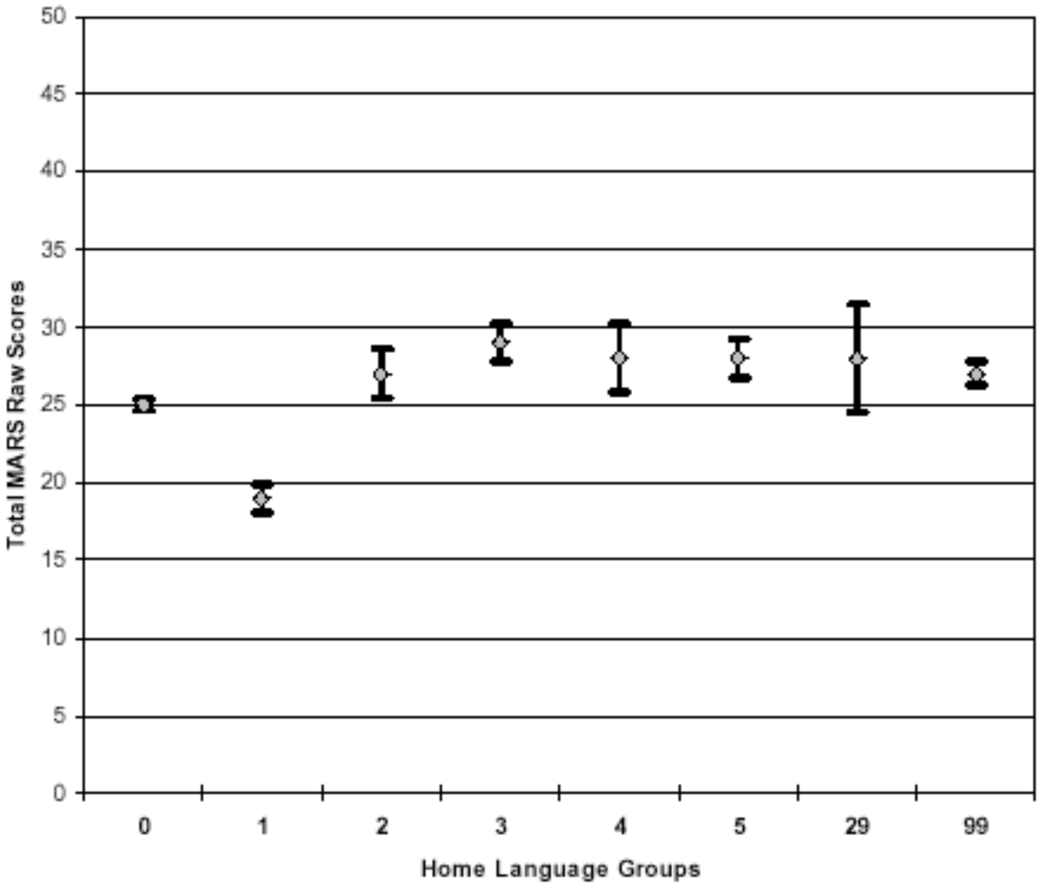


Table 21.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	2379
1	Spanish	294
2	Vietnamese	86
3	Cantonese	84
4	Korean	48
5	Filipino	64
29	Russian	22
99	Others/Unknown	419

Distribution of sampling means
Course 1 – Grade 8 and Lower
Home Language

In this section, test scores are compared across groups of students who speak different languages at home[42]. 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 the scores of students in the Cantonese, Filipino or "Other" home language groups, and significantly higher than those with a home language of Spanish. There are no other significant differences between students with English as a home language and any other category.

The scores of students with Spanish as a home language are significantly lower than the scores 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 as a home language. There is no significant difference between students with Vietnamese as a home language and any other home language group.

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

The students with Korean as a home language scored significantly higher than the students with Spanish as a home language. There are no significant differences between students with Korean as a home language and any other home language group.

Students with Filipino as a home language scored significantly higher than English or Spanish speaking students. There are no significant differences between students with Filipino as a home language and any other home language group.

The scores of students with Khmer as a home language were not significantly different from scores of any other home language group, but this is likely due to the small number of Khmer-speaking students in the sample (N=5).

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

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

Figure 10.3 Box and whisker plot of Total MARS Raw Scores by Parent Education

Course 1 - Grade 8 and Lower grades

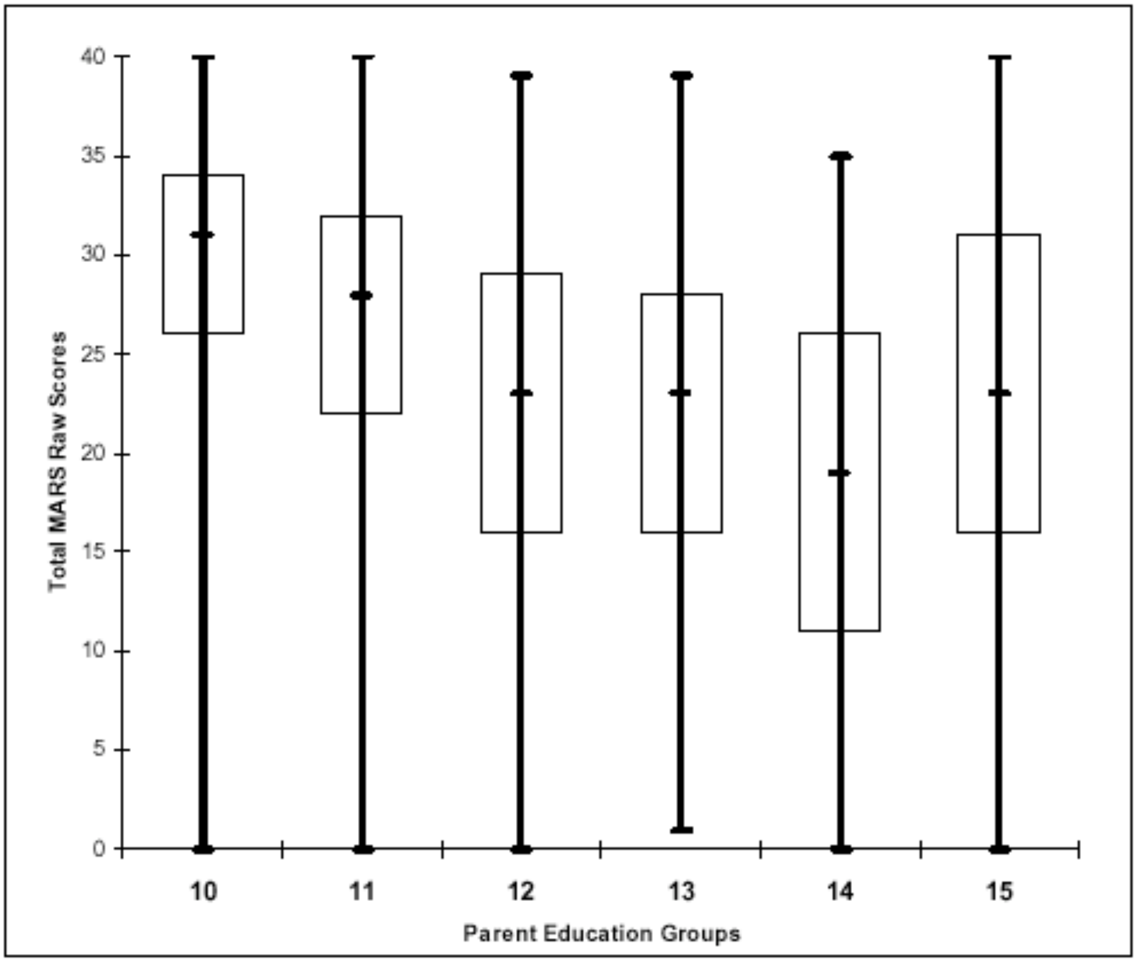


Table 10.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	890
11	College graduate	1005
12	Some college	693
13	High School graduate	445
14	Not a high school graduate	138
15	Others/Unknown	225

Figure 21.3 Distribution of sampling means by Parent Education

Course 1 - Grade 8 and Lower grades

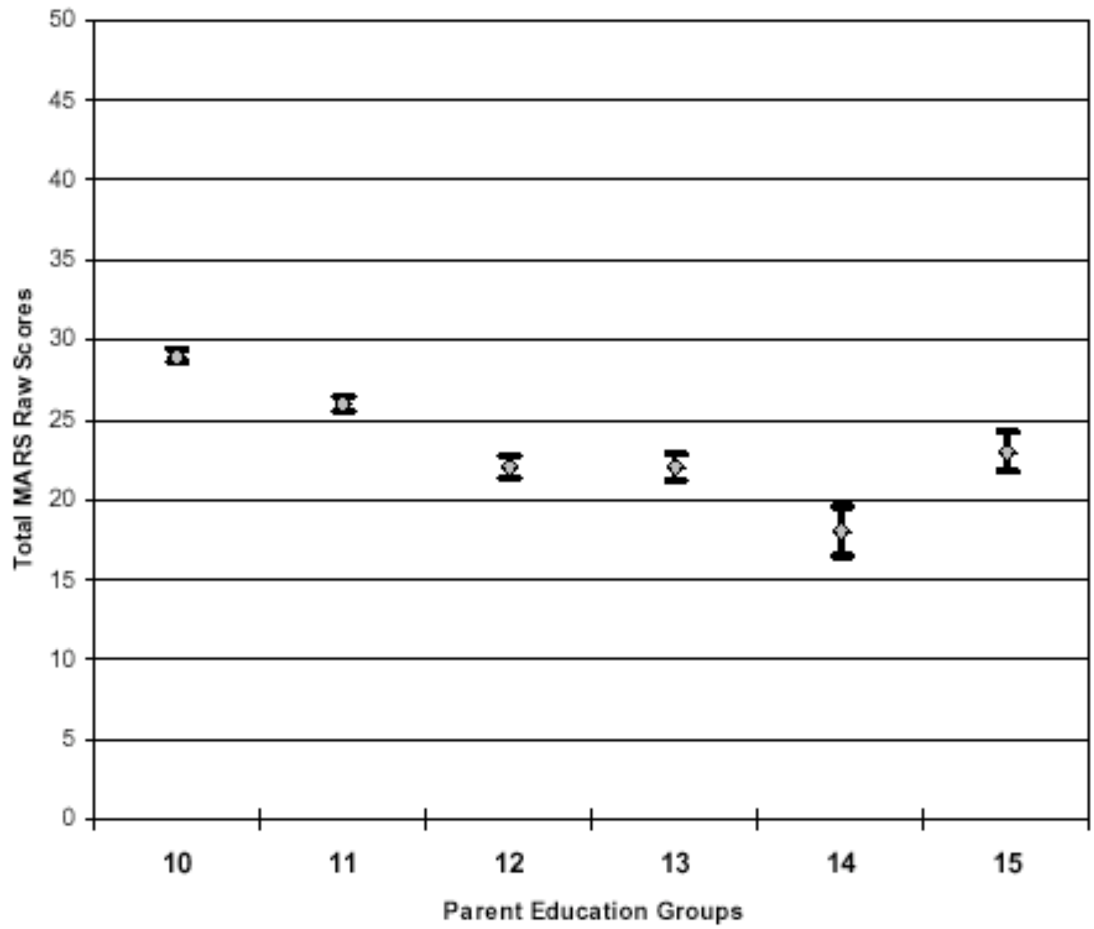


Table 21.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	890
11	College graduate	1005
12	Some college	693
13	High School graduate	445
14	Not a high school graduate	138
15	Others/Unknown	225

Distribution of sampling means
Course 1 – Grade 8 and Lower
Parent Education

In this section, test scores are compared across groups of different levels of parent education[43]. 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 in all other Parent Education categories.

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 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 scores of students whose parents are not high school graduates, and significantly lower than students with parents in the college graduate or graduate school 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 parent education level is unknown 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.

[43] 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.

Figure 10.4 Box and whisker plot of Total MARS Raw Scores by Gender

Course 1 - Grade 8 and Lower grades

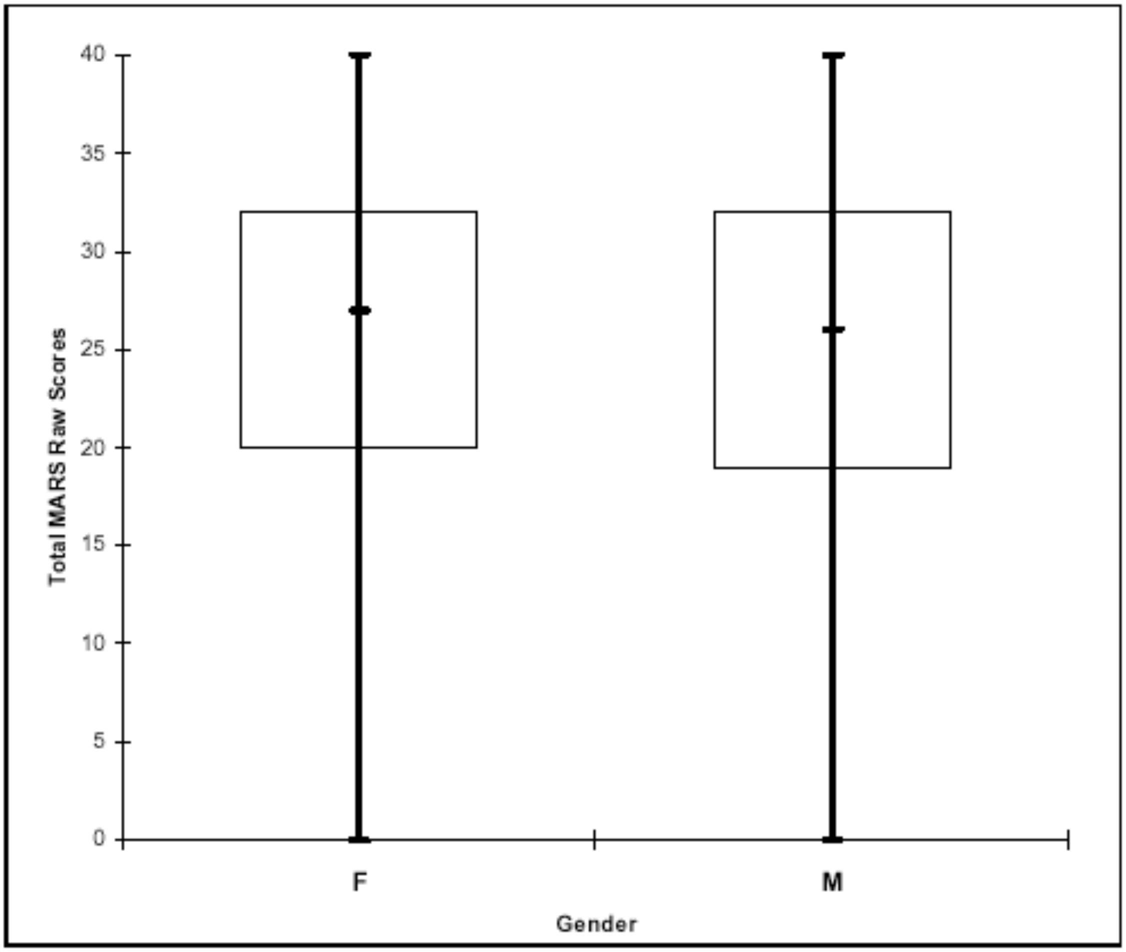


Table 10.4 Student count for Gender

Gender	Student Count
Female	1693
Male	1703

Figure 21.4 Distribution of sampling means by Gender
Course 1 - Grade 8 and Lower grades

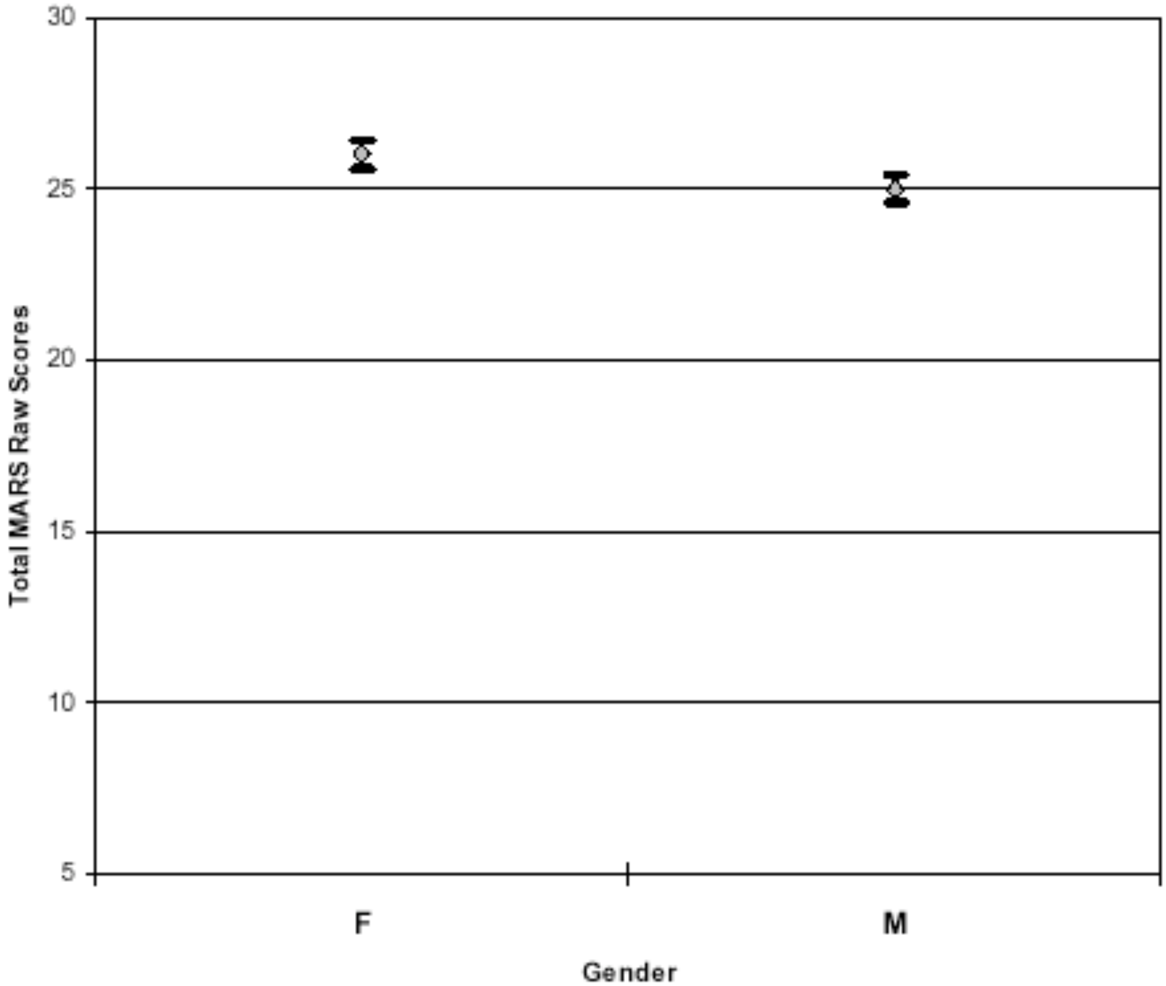


Table 21.4 Student count for Gender

Gender	Student Count
Female	1693
Male	1703

Distribution of sampling means
Course 1 – Grade 8 and Lower
Gender

In this section, test scores are compared across gender[44]. 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 females are significantly higher than the scores of males.

Figure 10.5 Box and whisker plot of Total MARS Raw Scores by Language Fluency Course 1 - Grade 8 and Lower grades

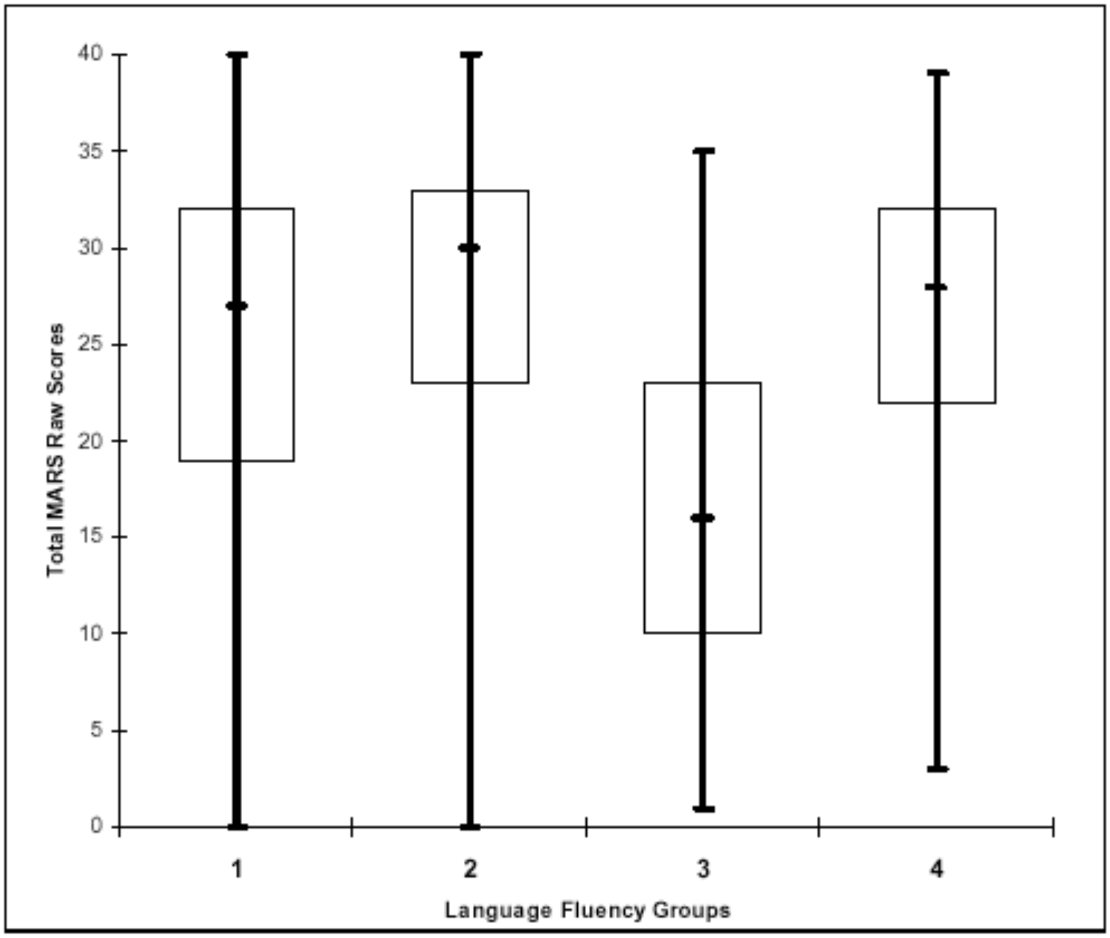


Table 10.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	2369
2	Initially Fluent (I-FEP)	417
3	English Learner	204
4	ReDesignated (R_FEP)	406

Figure 21.5 Distribution of sampling means by Language Fluency

Course 1 - Grade 8 and Lower grades

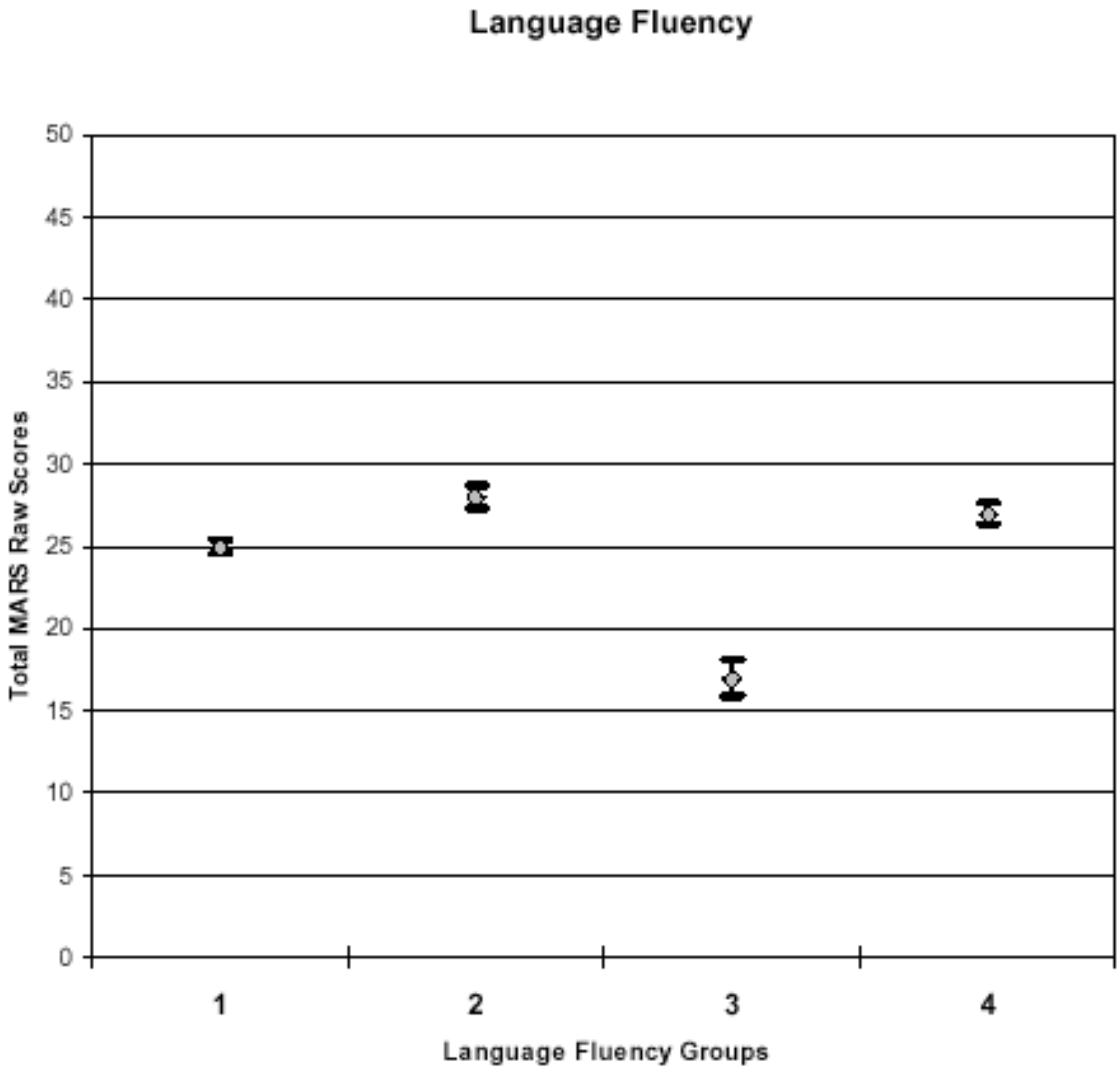


Table 21.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	2369
2	Initially Fluent (I-FEP)	417
3	English Learner	204
4	ReDesignated (R_FEP)	406

Distribution of sampling means
Course 1 – Grade 8 and Lower
Language Fluency

In this section, test scores are compared across groups of different language fluency[45]. 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 the scores 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 the scores of students in the English Only and English Learner categories. The scores of students in the FEP category are not significantly different from the scores of students in the re-designated FEP group.

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

The scores of students in the Re-designated FEP category are significantly higher than the scores of students described as English Only or English Learner.

[45] 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 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. Groups with Ns of less than 5 students are not reported.

Figure 11.1 Box and whisker plot of Total MARS Raw Scores by Ethnicity

Course 1- Grade 9 and Higher grades

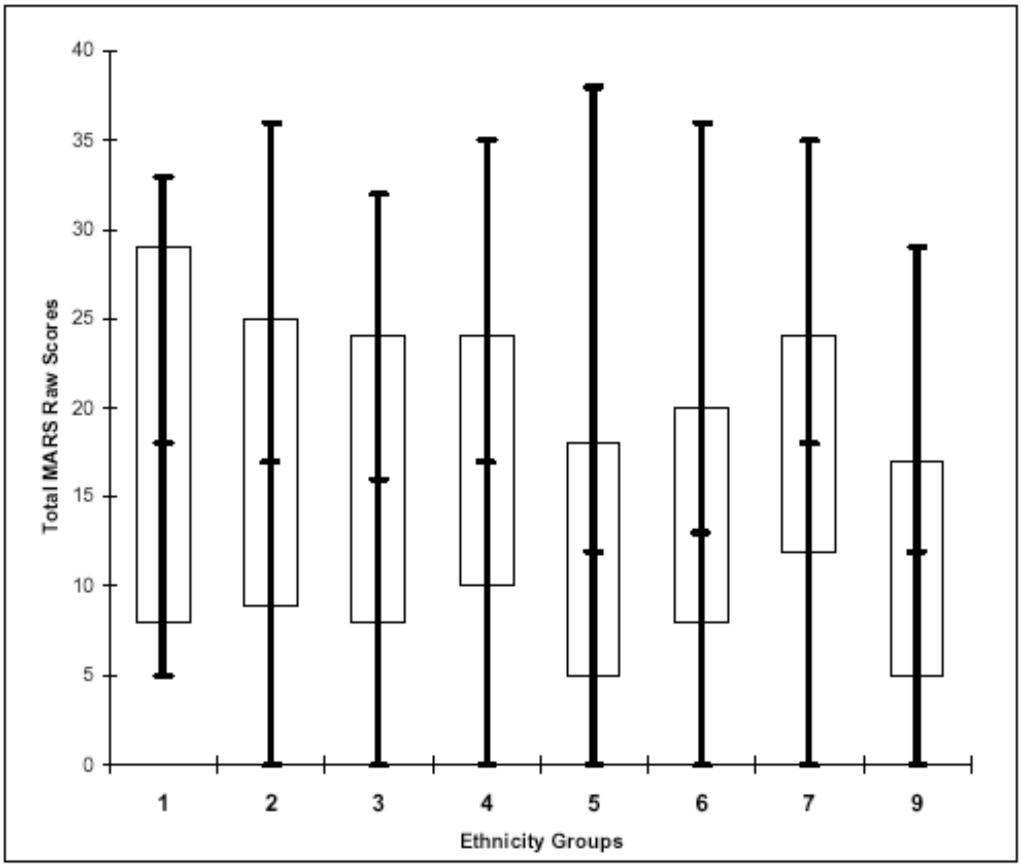


Table 11.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
1	American Indian	10
2	Asian/Asian American	487
3	Pacific Islander	25
4	Filipino	201
5	Hispanic/Latino	1270
6	African American	108
7	White (Not Hispanic)	460
9	Others/Unknown	13

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

Figure 22.1 Distribution of sampling means by Ethnicity

Course 1- Grade 9 and Higher grades

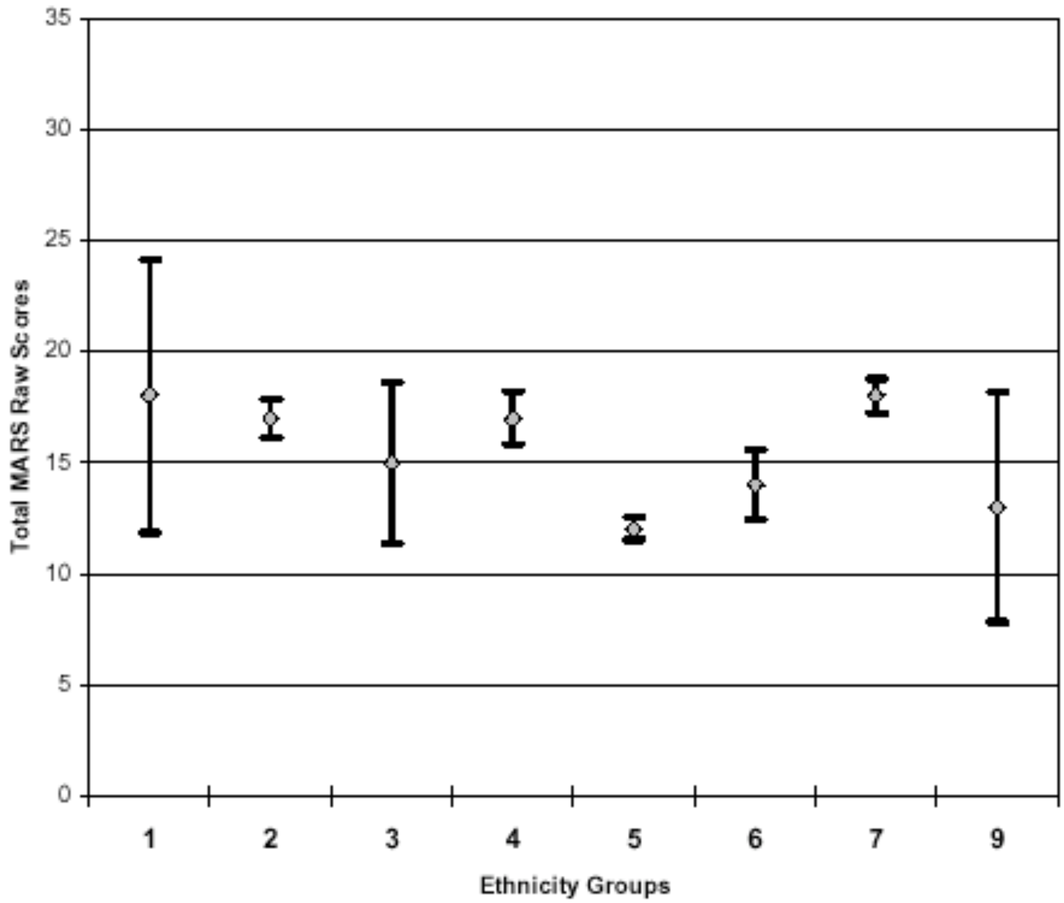


Table 22.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
1	American Indian	10
2	Asian/Asian American	487
3	Pacific Islander	25
4	Filipino	201
5	Hispanic/Latino	1270
6	African American	108
7	White (Not Hispanic)	460
9	Others/Unknown	13

Distribution of sampling means
Course 1 – Grade 9 and Higher
Ethnicity

In this section, test scores are compared across different ethnic groups[46]. 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 were few statistically significant differences in scores betweenamong the various ethnic groups who took the Course 1 test and who were in grade 9 or above. AsianAsian American and White students scored higher than HispanicLatino and African American students, and Filipino students also scored higher than HispanicLatino students.

[46] 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.

Figure 11.2 Box and whisker plot of Total MARS Raw Scores by Home Language Course 1- Grade 9 and Higher grades

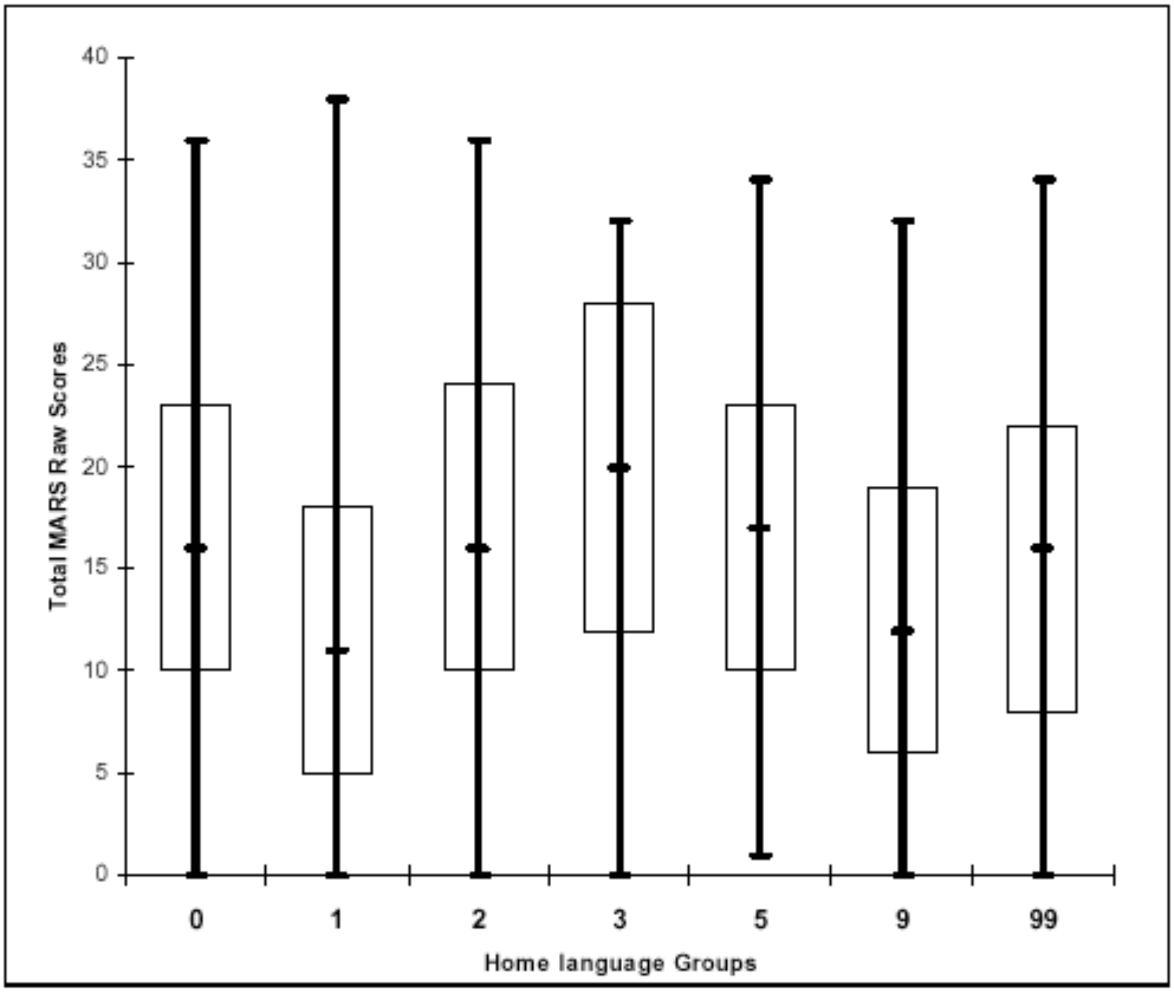


Table 11.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	1062
1	Spanish	911
2	Vietnamese	231
3	Cantonese	28
5	Filipino	86
9	Cambodian	36
99	Others/Unknown	220

Figure 22.2 Distribution of sampling means by Home Language
Course 1- Grade 9 and Higher grades

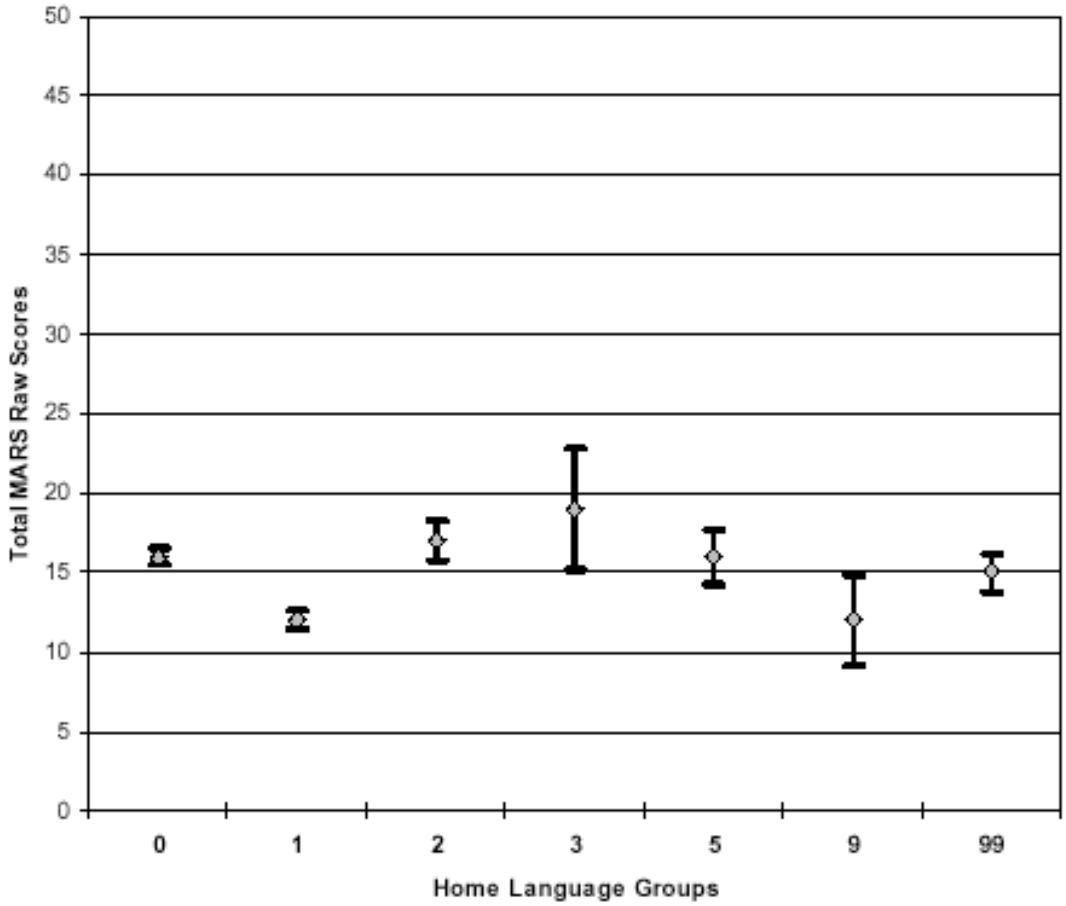


Table 22.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	1062
1	Spanish	911
2	Vietnamese	231
3	Cantonese	28
5	Filipino	86
9	Cambodian	36
99	Others/Unknown	220

Distribution of sampling means
Course 1 – Grade 9 and Higher
Home Language

In this section, test scores are compared across groups of students who speak different languages at home[47]. 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 higher than the scores of students with Spanish as a home language. There are no significant differences between students with English as a home language and any other category.

The scores of students with Spanish as a home language are significantly lower than the scores of students in any other home language group except Khmer (not significantly different).

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 home language groups.

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

There are no significant differences in scores between students with Korean as a home language and any other home-language group.

Students with Filipino as a home language scored significantly higher than students whose home language is Spanish. There are no significant differences between students with Filipino as a home language and any other home language group.

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

[47] 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 "Other" as a home language are significantly higher than the scores of students with Spanish as a home language. There are no significant differences between students with "Other" as a home language and any other home language group.

Figure 11.3 Box and whisker plot of Total MARS Raw Scores by Parent Education Course 1- Grade 9 and Higher grades

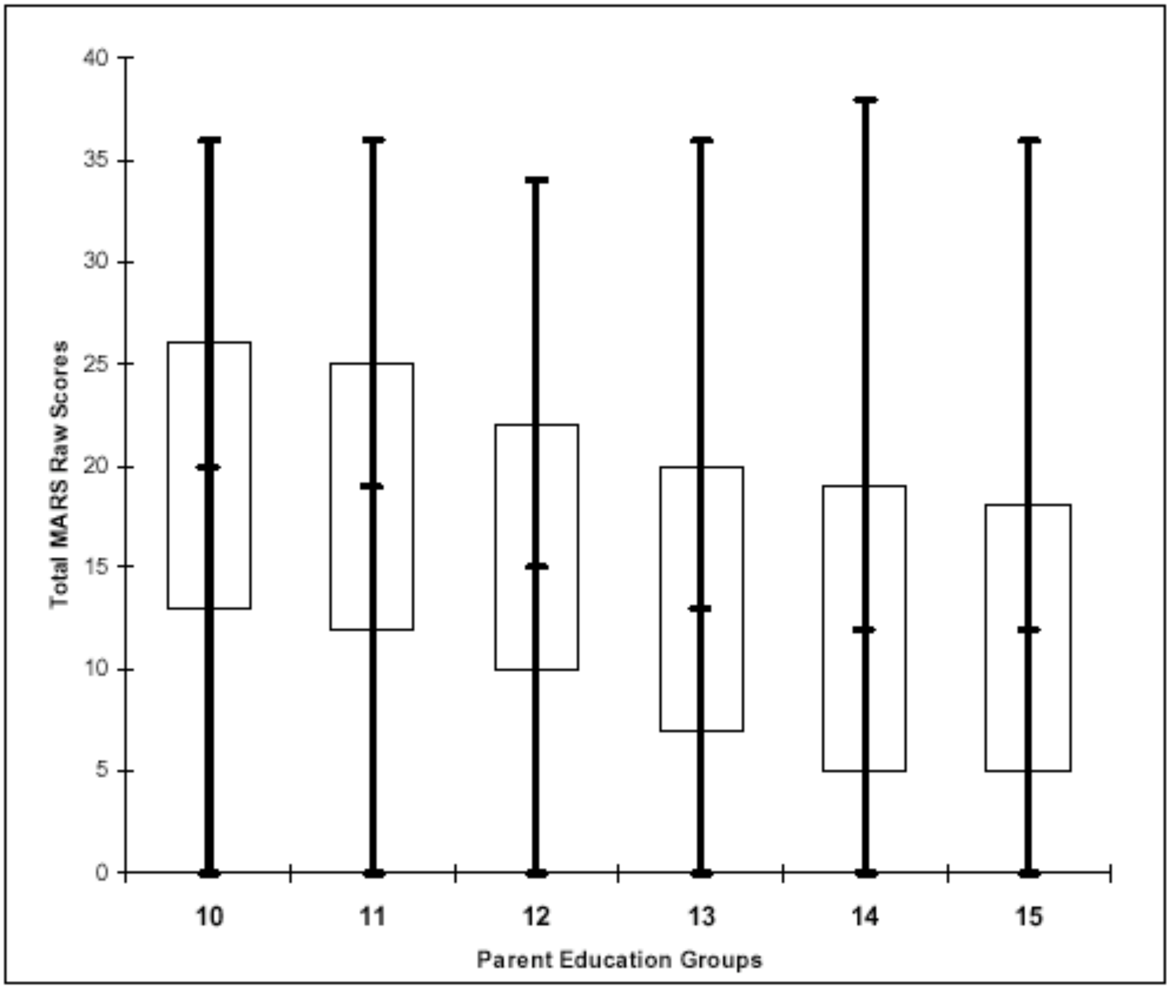


Table 11.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	117
11	College graduate	468
12	Some college	384
13	High School graduate	573
14	Not a high school graduate	467
15	Others/Unknown	565

Figure 22.3 Distribution of sampling means by Parent Education

Course 1- Grade 9 and Higher grades

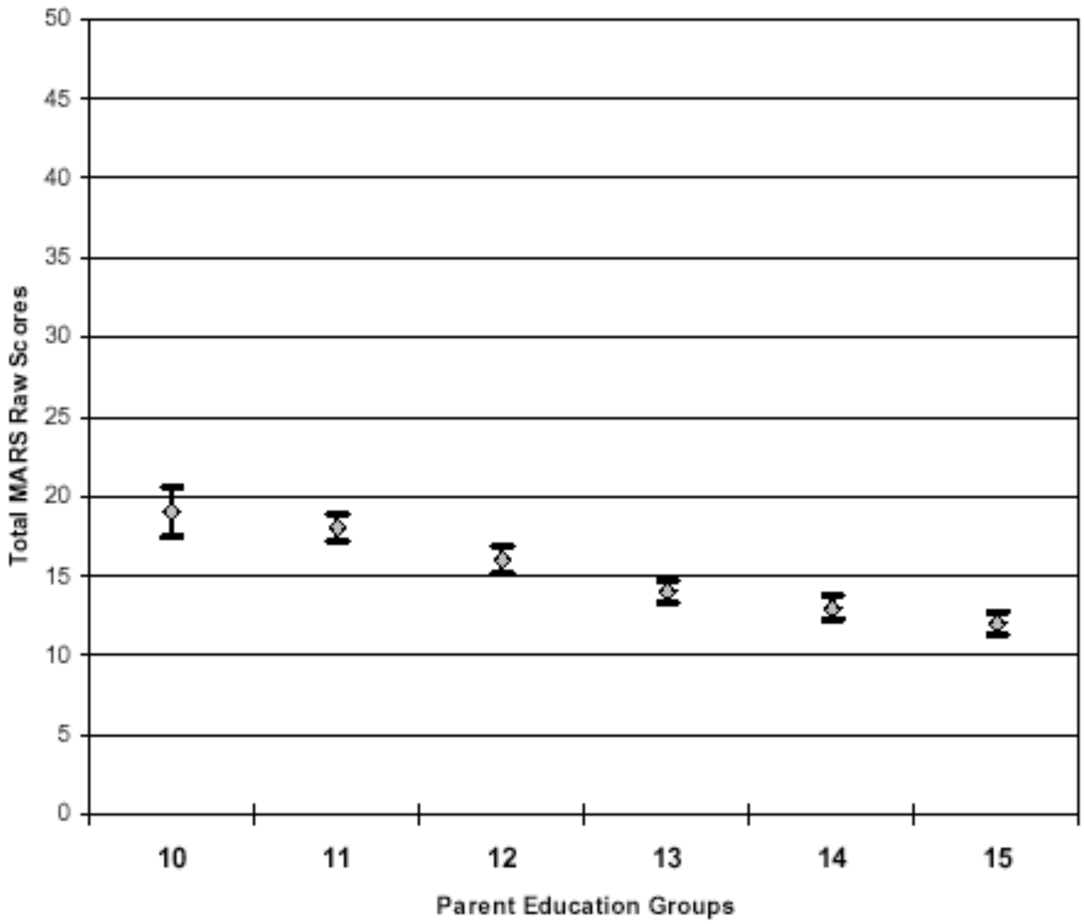


Table 22.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	117
11	College graduate	468
12	Some college	384
13	High School graduate	573
14	Not a high school graduate	467
15	Others/Unknown	565

Distribution of sampling means
Course 1 – Grade 9 and Higher
Parent Education

In this section, test scores are compared across groups of different levels of parent education[48]. 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 in the H.S. Graduate and not H.S. Graduate categories of parent education levels are not significantly different from each other, or from those whose parent education is unknown. The scores of all three of these groups, however, are significantly lower than the scores of students whose parents have some college, a college education, or a graduate education. There is no statistically significant difference in the scores of students whose parents are college graduates or have a graduate education.

[48] 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.

Figure 11.4 Box and whisker plot of Total MARS Raw Scores by Gender
Course 1- Grade 9 and Higher grades

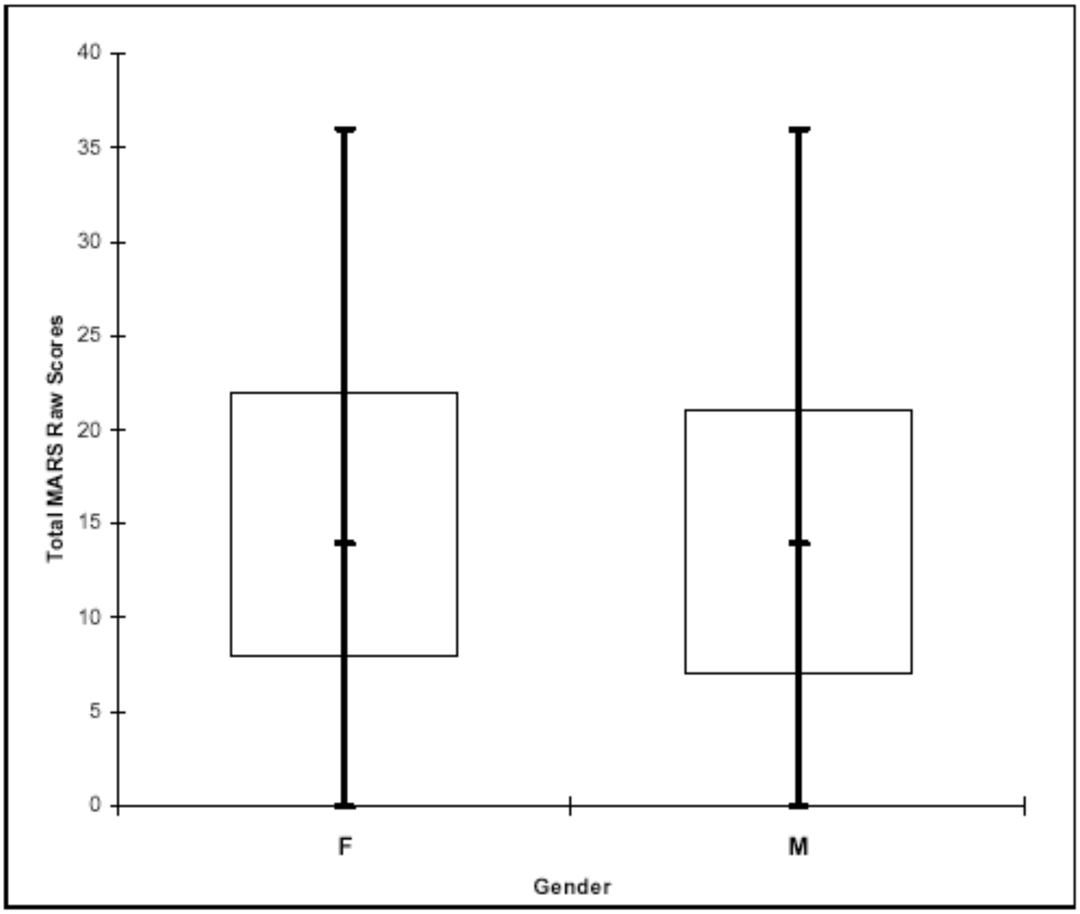


Table 11.4 Student count for Gender

Gender	Student Count
Female	1224
Male	1349

Figure 22.4 Distribution of sampling means by Gender

Course 1- Grade 9 and Higher grades

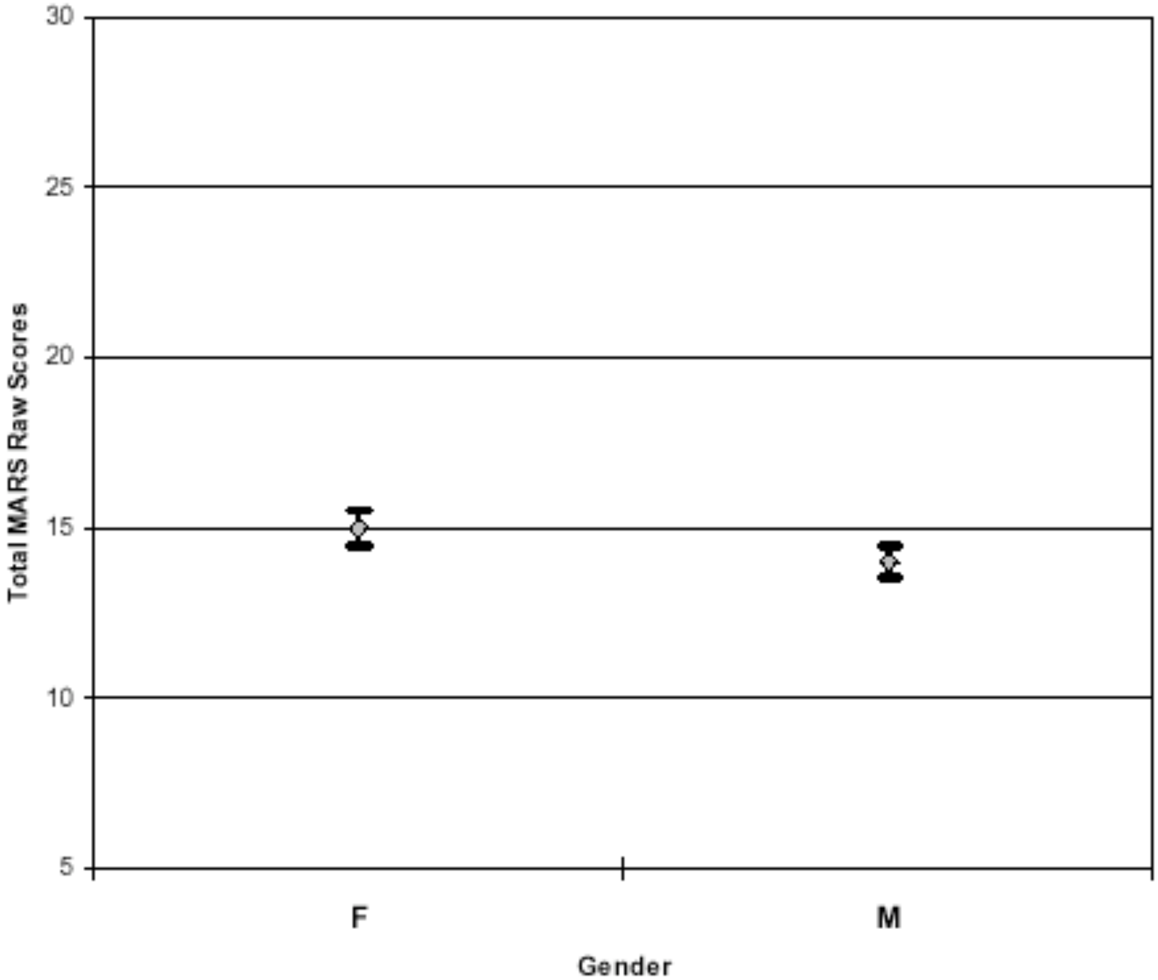


Table 22.4 Student count for Gender

Gender	Student Count
Female	1224
Male	1349

Distribution of sampling means
Course 1 – Grade 9 and Higher
Gender

In this section, test scores are compared across gender[49]. 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 females are significantly higher than the scores of males.

[49] 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.

Figure 11.5 Box and whisker plot of Total MARS Raw Scores by Language Fluency Course 1- Grade 9 and Higher grades

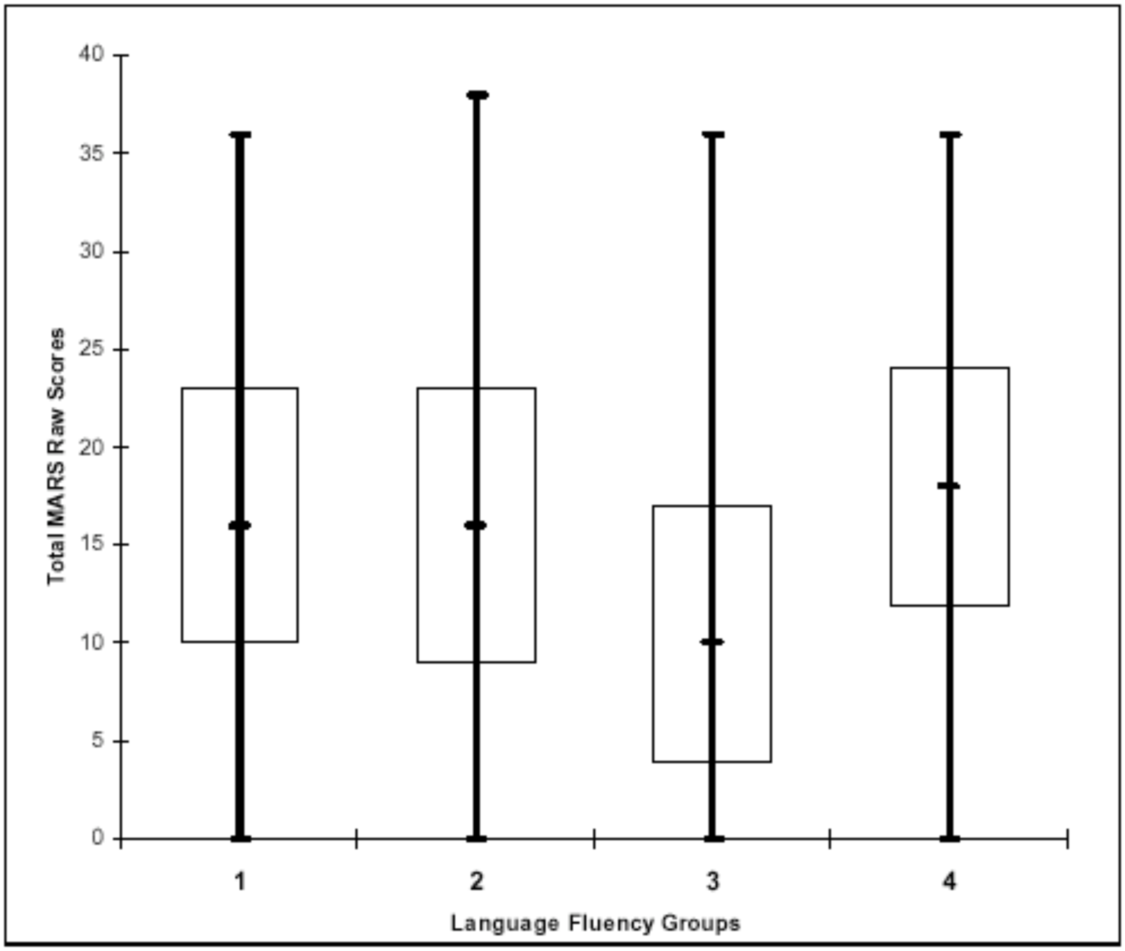


Table 11.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	1081
2	Initially Fluent (I-FEP)	359
3	English Learner	878
4	ReDesignated (R_FEP)	256

Course 1- Grade 9 and Higher grades

Language Fluency

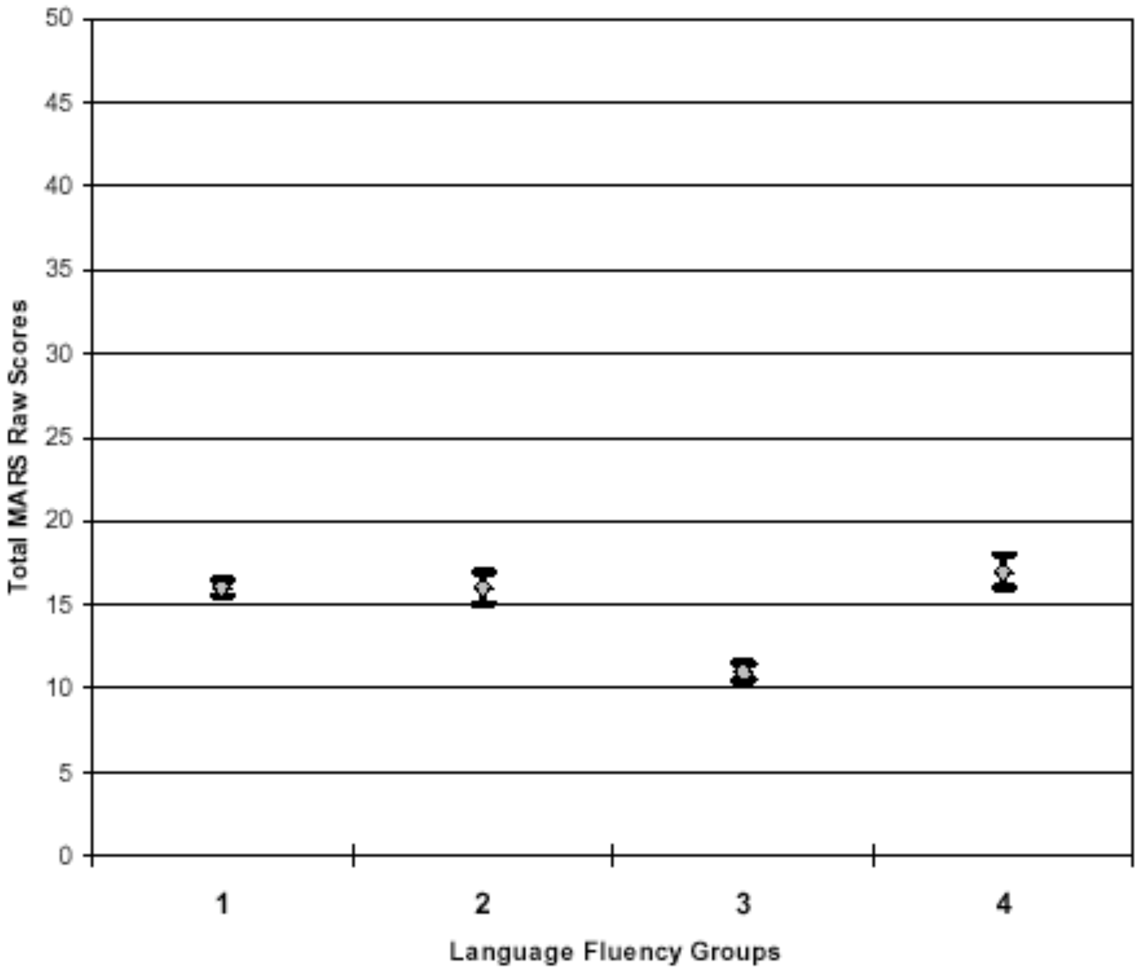


Table 22.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	1081
2	Initially Fluent (I-FEP)	359
3	English Learner	878
4	ReDesignated (R_FEP)	256

Distribution of sampling means
Course 1 – Grade 9 and Higher
Language Fluency

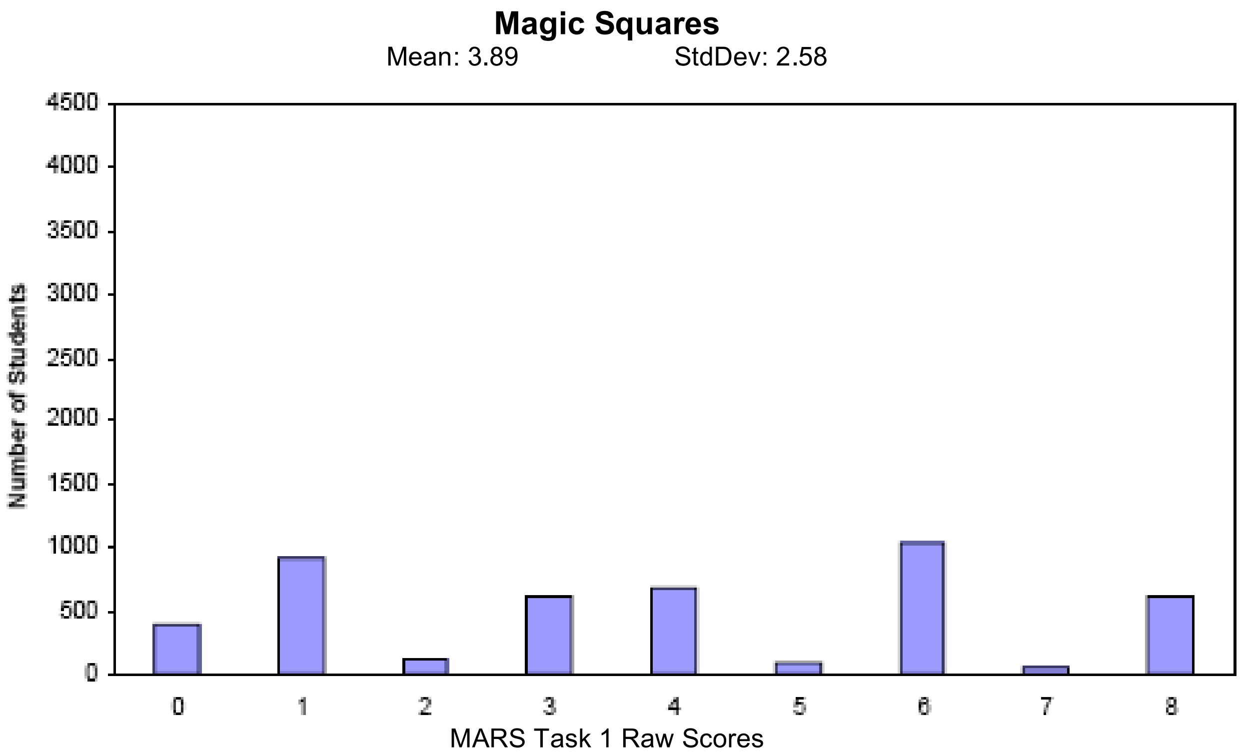
In this section, test scores are compared across groups of different language fluency[50]. One way to look at the group differences is to study the graphs. If scores from group A are above and do not overlap the scores from group B then group A is significantly higher than group B. Conversely, if scores 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.

When students are grouped according to language fluency classification there are few significant differences. Specifically, students described English Learners have scores significantly lower than the scores of students in all other categories.

Course 1 Task 1 Magic Squares

Student Task	Use symbolic algebraic notation to calculate values in "magic" squares where each row, column and diagonal adds to the same number.
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, including deductive and inductive reasoning, making and testing conjectures and using counterexamples and indirect proof.** • Show mathematical reasoning in a variety of ways, including words, numbers, symbols, pictures, charts, graphs, tables, diagrams, and models

Frequency Distribution for Task 1 – Course 1 – Magic Squares



Score:	0	1	2	3	4	5	6	7	8
Student Count	391	925	124	618	686	77	1037	48	606
% < =	8.7%	29.2%	31.9%	45.6%	60.8%	62.5%	85.5%	86.6%	100.0%
% > =	100.0%	91.3%	70.8%	68.1%	54.4%	39.2%	37.5%	14.5%	13.4%

The maximum score available for this task is 8 points.
The minimum score for a level 3 response, meeting standards, is 5 points.

Most students, about 91%, total the numerical values of the rows and columns in the first magic square. Many students, 68%, could total the values of rows and columns in the magic squares both numerically and using symbolic notation. About half the students, could add the number values for the first magic square, fill in numerical values for the third magic square, and give the total for the rows and columns in the third magic square. Some students, 39%, could find the totals for all the magic squares using numbers and symbolic expressions and fill in most of the missing values in the third magic square. 13% of the students could meet all the demands of the task including using algebra to find the missing values for the last magic square. Almost 9% of the students scored no points on this task. 76% of those students attempted the task.

Based on teacher observation, this is what algebra students know and are able to do:

- Add numerical data
- Use guess and check to find missing values in a magic square

Areas of difficulty for algebra students:

- Combining like terms
- Understanding the difference between x + x + x and x^3
- Using algebra to find a solution in a problem-solving setting
- Setting up an equation or system of equations to match a problem situation
- Using as few variables as possible/ trying to define one term in terms related to other ideas or parts of the problem instead of inserting new variables
- Understanding that the sums of the diagonals needed to be the same as the sums for the rows and columns in a magic square

Strategies used by successful students:

- Finding the value of x, y, and z, then using substitution to find missing variable
- Setting up simultaneous equations with 2 unknowns
- Looking for expressions with similar terms that with one value would cancel out
- Using the diagonal total to set up equation with 1 unknown
- Comparing symbolic and numeric values for the 4 corners

Implications for Instruction:

Students need contexts when learning to combine and manipulate algebraic notation. If students merely spend time simplifying expressions in isolation of a context, then they will never develop a reason of why, when or how to use algebraic notation in a problem situation. Students must have learning experiences around developing algebraic models in problem situations These must be non-routine situations and not a series of building algebraic models following a formalized process. Most students did not use algebra in this problem. That indicates a lack of familiarity with experiences where students need to develop and use algebraic models.

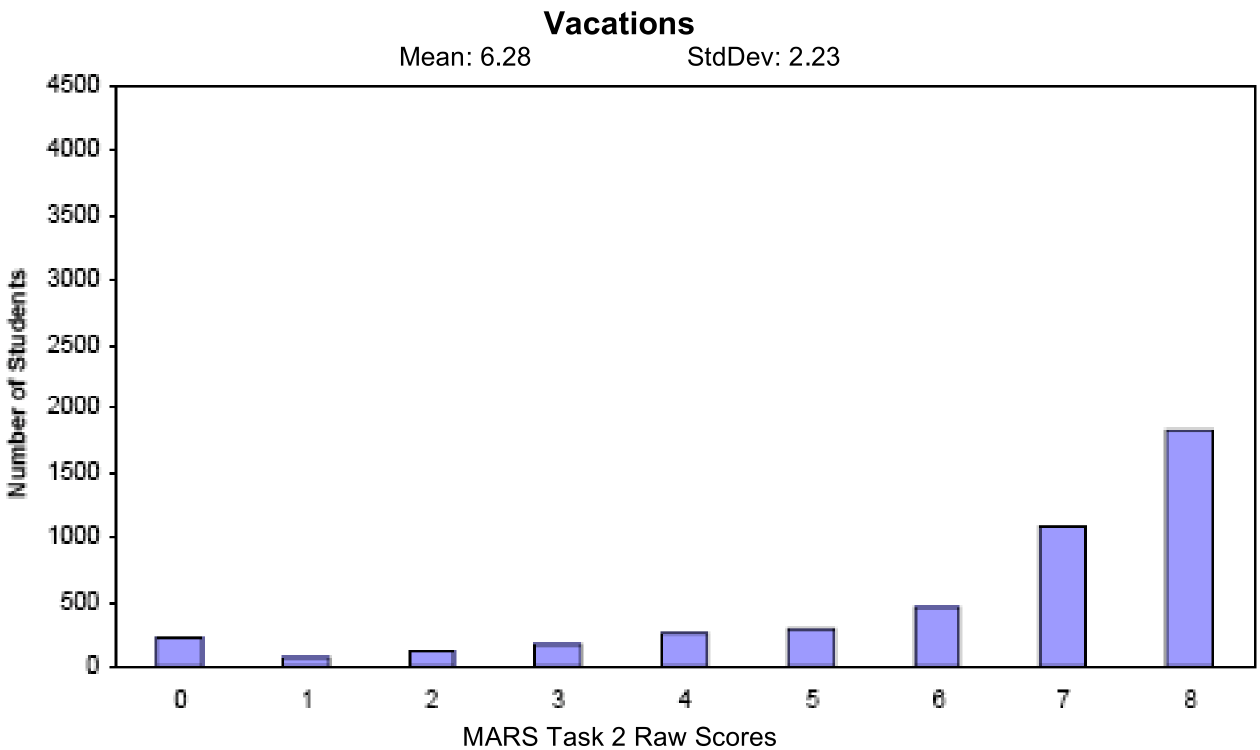
After a task like this has been given to students for informal assessment, the teacher might consider ways to design a lesson around the variety of strategies possible for solving the problem. The lesson might start by posing a hypothetical debate between two students. For example, "Cindy thinks that the problem can be solving by using the algebraic sum of the rows. Sam disagrees because the numerical value of the row is not given. Could Cindy be correct?" Then students could work in groups to see if they can use Cindy's idea. A different start might be, "Fred says that some of the information includes variables that cancel each other out. Would this be helpful? How would this help solve the problem?" Again students could work in groups and discuss ways to make this work. Consider, " Laura makes all the missing values into new variables. Kim says that the problem can be solved using only one variable. Juan says that his class has been working on solving two equations with two unknowns. Who is right? Could more than one person be right?" Having students examine a variety of strategies, helps them start to develop a generalization about classes of problems rather than solving each problem as a unique event. Just as we want students to develop generalizations for a function pattern, they need to be looking for generalizations about classes of problems. The discussion and debate process helps them make this transition. After students have found a variety of strategies, the teacher might ask further probing questions: "What did all these solutions have in common? How were they different? What did we have to know to use each one of the strategies? How were different pieces of information needed and unique to each solution type?" These types of questions help students reflect beyond the answer to making generalizations about what is needed to set up equations for solving problems.

This lesson structure is good for a variety of problems. Ideally, enough variety of solutions will come from the students in the class to provide the prompts for the lesson. However, the teacher should anticipate a variety of solution paths and be able to pull in some other strategies to push students' thinking if the students do not come up with them on their own. The prompts try not to lead or scaffold the student thinking, but give enough hints for students to do some thinking on their own and lead to debate. In the debating process, students are then confronted with ideas about justification. What makes a convincing argument? What does not make a convincing argument? In this process, students are also developing their logic as well as their problem-solving skills. Students are applying algebraic techniques and practicing procedures, but they are also learning the mathematics and big ideas behind the procedures and symbolic manipulation.

Course 1 Task 2 Vacations

Student Task	Match graphic displays to the written descriptions of how some students are paying for their summer vacations. Write a formula that describes each of the matched relationships and then write a possible description for a new vacation saving formula.
Core Idea 1 Functions and Relations	**Understand patterns, relations, and functions.** • Generalize patterns using explicitly defined functions • Understand relations and functions and select, convert flexibly among, and use various representations for them • Analyze functions of one variable by investigating local and global behavior, including slopes as rates of change, intercepts and zeros
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 • Use symbolic algebra to represent and explain mathematical relationships • Use symbolic expressions to represent relationships arising from various contexts • Approximate and interpret rates of change from graphic and numeric data
Core Idea 2 Mathematical Reasoning	**Employ forms of mathematical reasoning and proof appropriate to the solution of the problem, including deductive and inductive reasoning, making and testing conjectures and using counterexamples and indirect proof.** • Use induction to make conjectures and use deductive reasoning to prove conclusions • Draw reasonable conclusions about a situation being modeled

Frequency Distribution for Task 2 – Course 1 – Vacations



Score:	0	1	2	3	4	5	6	7	8
Student Count	216	74	122	167	260	290	462	1094	1827
% < =	4.8%	6.4%	9.1%	12.8%	18.6%	25.0%	35.3%	59.5%	100.0%
% > =	100.0%	95.2%	93.6%	90.9%	87.2%	81.4%	75.0%	64.7%	40.5%

The maximum score available for this task is 8 points.
The minimum score for a level 3 response, meeting standards, is 5 points.

Most students, about 93%, could match the equations to the verbal descriptions. Many students, 81%, could match the equations to verbal descriptions, write an equation for Carla, and match at least two of the verbal descriptions to the appropriate graphs and justify the choice based on slope or y-intercept. About half the students could match the verbal descriptions to the graphs with justification, match the equations to the verbal descriptions, and either write an equation for Carla or write a verbal description, fitting the context of the problem, to match the given equation in part 3. More than 40% of the students could meet all the demands of the task. 4.8% of the students scored no points on the task. None of the middle school students in the sample scored 0 points. 66% of the high school students with a score of 0 did not attempt the task.

Based on teacher observation, this is what algebra students knew and were able to do:

- Recognize positive and negative slopes
- Recognize the y-intercept as the initial amount of money
- Match graph to descriptions
- Match equations to descriptions

Areas of difficulty for algebra students:

- In A= 50n +150, students did not see the 150 as the initial amount of money
- Did not provide enough information for choosing the graph or did not focus their reasons on what made the description or graph unique
- Understanding or qualifying the differences between the graphs for Ben and Carla
- Writing an equation for Carla

Implications for Instructions on Vacations:

Learning a concept or idea happens over time and deepens with opportunities to reflect on what is being studied. Working problems in context helps students to connect meaning to symbolic notation and graphic representations being studied and gives them opportunities to deal with significant features in ways that are not possible with problems out of context.

Having students tackle making a graph to fit a situation helps them to think about the importance of y-intercept, slope, and the relationship between variables. Students might be given a description of soda in a soda machine and how it changes based on lunch, being filled, and time of day. Students need to think about rate of change, relative values for the dependent variable and scale. Context gives students reason to think why time is better for the x-axis, or why one-variable is dependent upon another. These messy, but interesting ideas are not confronted when working exercises without context. Students might also be given a context where the graph is the inverse of the actual action in the narrative, for example speed on a roller coaster increases as the roller coaster goes down hill. This situation gets students to think about many important issues. Students might compare graphs for the roller coaster speed and time to graphs of height and time.

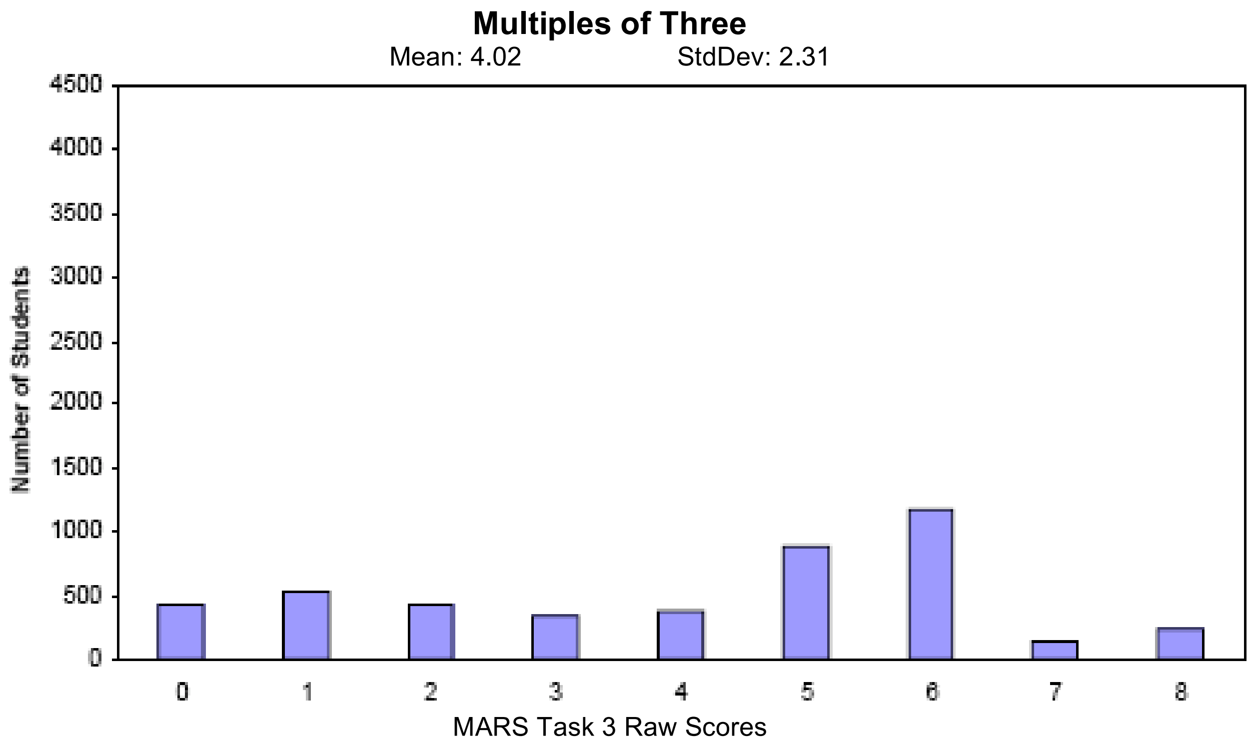
Another good technique for getting students to think about graphs is to provide them with an unlabeled graph and write a description that might fit the ideas on the graph. For example students might write a narrative to describe who won a race or what happened to a runner during a race for different types of lines. Also students might be asked to think about a blank graph and given situation and try to give realistic values to fit the graph.

The task itself provides a good model for worthwhile learning activities. After students have worked the problem as an assessment activity, students should be given the opportunity to discuss big ideas, like when does Ernie receive the $150? The problem might be posed, " Sara thinks that Ernie saves $50 a month and then her dad gives her $150 for a birthday present. Harry disagrees, saying that it doesn't matter when Ernie gets the $150. It could be in February or June but the answer will be the same. Walt thinks they are both wrong! What do you think? Can you make a convincing argument?" This use of cognitive conflict gets students to really analyze the situation and match the mathematics to the context and allows them to verify their ideas with numbers that will make sense. Numbers and letters stripped of context do not allow students to verify or check the reasonableness of their assumptions.

Course 1 Task 3 Multiples of Three

Student Task	Given a statement regarding multiples of three, test it to see if it is true, find examples that match the statement and explain and justify conclusions.
Core Idea 3 Algebraic Properties and Representations	**Represent and analyze mathematical situations and structures using algebraic symbols.** • Compare and contrast the properties of numbers and number systems including real numbers • Use symbolic algebra to represent and explain mathematical relationships • Use symbolic expressions to represent relationships arising from various contexts
Core Idea 2 Mathematical Reasoning	**Employ forms of mathematical reasoning and proof appropriate to the solution of the problem, including deductive and inductive reasoning, making and testing conjectures and using counterexamples and indirect proof.** • Explain the logic inherent in a solution process • Use induction to make conjectures and use deductive reasoning to prove conclusions • Draw reasonable conclusions about a situation being modeled

Frequency Distribution for Task 3 – Course 1 – Multiples of Three



Score:	0	1	2	3	4	5	6	7	8
Student Count	425	529	429	335	373	889	1166	131	235
% < =	9.4%	21.1%	30.7%	38.1%	46.3%	66.0%	91.9%	94.8%	100.0%
% > =	100.0%	90.6%	78.9%	69.3%	61.9%	53.7%	34.0%	8.1%	5.2%

The maximum score available for this task is 8 points.
The minimum score for a level 3 response, meeting standards, is 5 points.

Most students, about 90%, could state that 4721 was not a multiple of three. Many students knew that 4721 was not a multiple of 3 and could provide a counterexample for why adding two multiples of 3 does not always equal a multiple of 6. About half the students, 54%, could use the divisibility rule given to prove why 4721 was not a multiple of 3 and to create a 5-digit number that was a multiple of 3. 34% could also give a counter example to disprove the conjecture in part 4. 5% of the students could meet all the demands of the task including developing a convincing argument about distributive property of groups of 3's to show why adding two different multiples of 3 would always make a multiple of 3. 9.4% of the students scored no points on this examine. All the middle school students attempted the task. Only 37.5% of the high school students with this score attempted the task.

Based on teacher observation, this is what algebra students know and are able to do:
- Could interpret and apply the divisibility rule for multiples of 3
- Could prove a number was a multiple of three using division
- Could provide examples to illustrate a conjecture

Areas of difficulty for algebra students:
- Make a justification based on properties of the number rather than by giving examples
- Find counter examples to disprove a statement
- Being specific, giving examples to support their statements

Implications for Instruction:

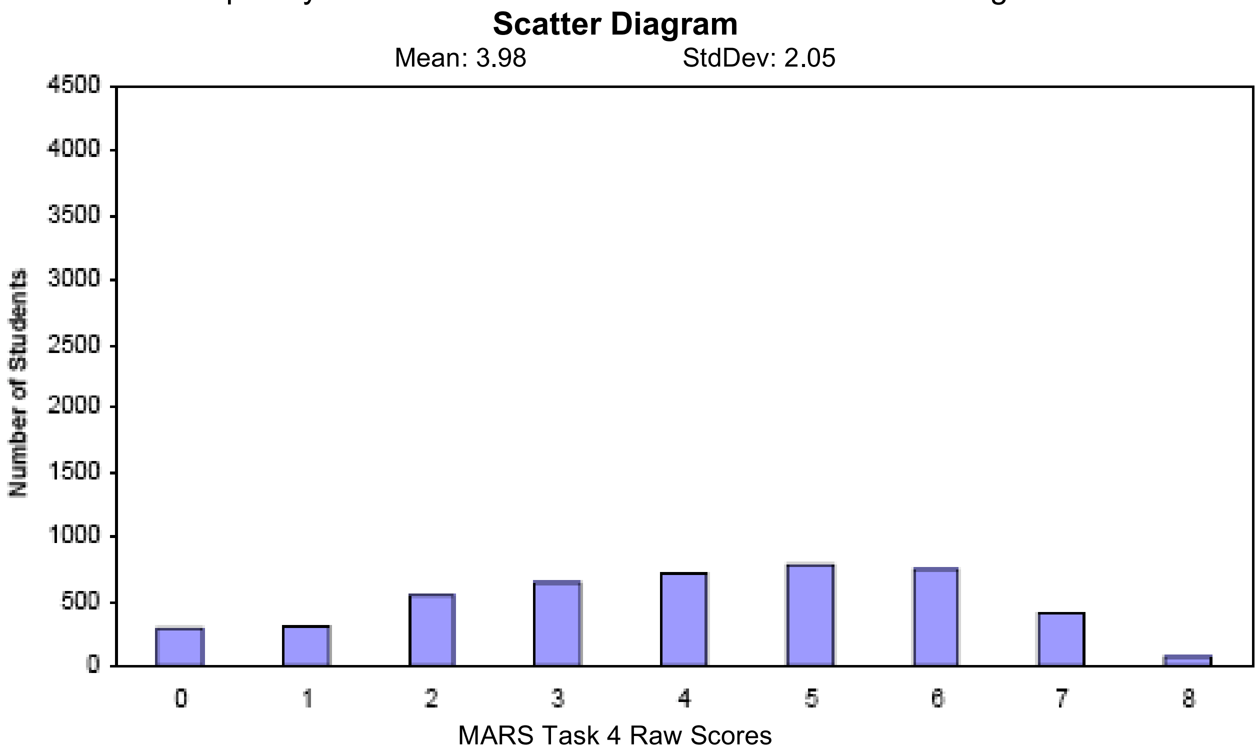
Justification and proof are the cornerstones of mathematics in a college prep curriculum. This requires students to develop the higher level thinking skills of analysis and synthesis. Students need opportunities to explore number theory problems then pose and justify their conjectures. Students need opportunities to use real mathematical theorems by interpreting their meanings and applying them to other situations. Students may or may not come to class knowing rules of divisibility such as the Multiple of Three problem. Teachers should provide the class with the opportunity to investigate other rules of divisibility and justify why they work. This will provide opportunities for students to develop their skills at proving mathematical ideas.

Mark Driscoll's book, Fostering Algebraic Thinking Book, has a myriad of problems that allow students to build their justification skills. Chapter 3 looks at justification around Number Sense and Number Theory, with problems about the product of 3 consecutive numbers and the differences of perfect squares. Books on problem solving, like Creative Problem Solving in School Mathematics by George Lenchner, often have whole chapters devoted to number theory and divisibility rules.

Course 1 Task 4 Scatter Diagram

Student Task	Explain the information presented in a scatter plot of students' scores on two tests. Evaluate statements made about the relationships found from the data and revise the statements if necessary.
Core Idea 5 Data Analysis	**Select and use appropriate statistical methods to analyze data and understand and apply basic concepts of probability.** • Understand the relationship between two sets of data (bivariate) and describe trends and shape of the plot including correlations (positive, negative, or no) and lines of best fit • Make inferences based on the data and evaluate the validity of conclusions drawn
Core Idea 2 Mathematical Reasoning	**Employ forms of mathematical reasoning and proof appropriate to the solution of the problem, including deductive and inductive reasoning, making and testing conjectures and using counterexamples and indirect proof.** • Use induction to make conjectures and use deductive reasoning to prove conclusions • Draw reasonable conclusions about a situation being modeled

Frequency Distribution for Task 4 – Course 1 - Scatter Diagram



Score:	0	1	2	3	4	5	6	7	8
Student Count	296	307	547	649	712	783	740	408	70
% < =	6.6%	13.4%	25.5%	39.9%	55.7%	73.0%	89.4%	98.4%	100.0%
% > =	100.0%	93.4%	86.6%	74.5%	60.1%	44.3%	27.0%	10.6%	1.6%

The maximum score available for this task is 8 points.
The minimum score needed for a level 3 response, meeting standards, is 4 points.

Most students, about 93%, were able to see that the scatterplot showed a positive correlation between scores on the two tests. Many students, about 74%, were able to plot the point and recognize that statements 1 and 3 in the table were true. More than half the students were able to plot the point, recognize that statements 1 and 3 were correct, and either draw a line of best fit or rewrite statement 2 or 4. About 10% of the students could meet most of the demands of the task including plotting a point, drawing a line of best fit, comparing lowest scores on A and B, finding range for test B, comparing highest scores on A and B, and finding the biggest difference. Only about 2% of the students could explain what the line of best fit represents in terms of the context of the problem. 7% of the students scored no points on this task. Only 13% of those students attempted the task.

Based on teacher observation, this is what algebra students knew and were able to do:

- Plot a point
- Draw a line of best fit
- Identify statements about the graph which were correct
- Identify statements about the graph which were incorrect

Areas of difficulty for algebra students:

- Rewriting incorrect statements to make true statements about the graph, particularly the two lowest scores, the highest scores, and the largest difference in scores
- Understanding the purpose of the line of best fit

Implications for Instruction : Students at this grade level should be familiar with a variety of ways to display and interpret data. When given a set of data, they should be comfortable picking the display that best represents that data or that suits the purpose for answering the questions for which the data was collected. Giving students opportunities to pose questions and collect their own data helps make the choice of data display selection seem more relevant and clarifies the purpose of each choice.

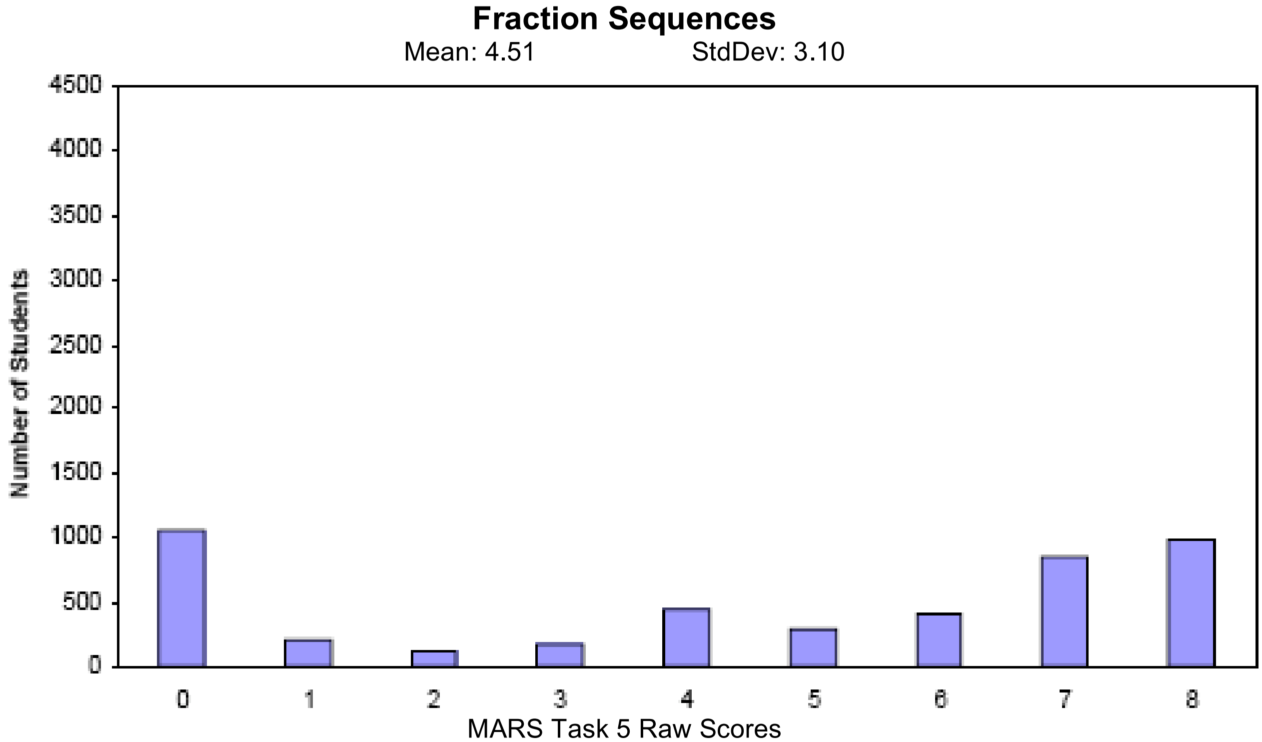
Students should see that scatter plots help to answer questions about the relationship between two items, is there a cause and effect? Does one seem to help predict the other? Is there a relationship between shoe size and weight? Shoe size and speed running a 50 yd. dash? The amount of homework completed and test scores?

Knowing students ideas about average as the most common misinterpretation of tests scores, the teacher might plan a lesson on returning the papers to help students confront this misconception. The teacher might ask them to talk about what the average is for each test, what the average might mean for one particular student on the graph, say point (24,10) or (19,26). Where would you plot a point to represent each average? If you started to graph all the points, what would you notice about the plots of the averages? Now have students find the number of points and try to design a line dividing the points into equal groups above and below the line? How does this line look different? Which line is the best predictor of someone's score? Test out several cases. Then have the students examine and think about these same questions using some other scatter plots. Students should be exposed to several that have different correlations, positive, negative, no correlation, tight correlation, and loose correlation. Then students should be able to answer why this is a useful tool for researchers. Students should also be given situations that appear to have a correlation, but don't have a cause and effect relationship, where the results are just coincidental. It is in unpacking these types of discrepancies that the mathematics becomes engaging for students and the ideas start to take on meaning.

Course 1 Task 5 Fraction Sequences

Student Task	Extend a sequence of fractions and compare the values. Make conjectures about the patterns in the values of the terms as well as their equivalent decimal values.
Core Idea 1 Functions and Relations	**Understand patterns, relations, and functions.** • Generalize patterns using explicitly defined functions • Understand relations and functions and select, convert flexibly among, and use various representations for them
Core Idea 2 Mathematical Reasoning	**Employ forms of mathematical reasoning and proof appropriate to the solution of the problem, including deductive and inductive reasoning, making and testing conjectures and using counterexamples and indirect proof.** • Use induction to make conjectures and use deductive reasoning to prove conclusions • Draw reasonable conclusions about a situation being modeled

Frequency Distribution for Task 5 – Course 1 – Fraction Sequences



Score:	0	1	2	3	4	5	6	7	8
Student Count	1052	209	119	176	437	294	403	841	981
% < =	23.3%	27.9%	30.6%	34.5%	44.2%	50.7%	59.6%	78.3%	100.0%
% > =	100.0%	76.7%	72.1%	69.4%	65.5%	55.8%	49.3%	40.4%	21.7%

The maximum score available on this task is 8 points.
The minimum score for a level 3 response, meeting standards, is 4 points.

Many students could find some of the fractional values in either sequence 1 or 3. More than half the students, 65%, could find all the fractions and their equivalent decimals for one of the two sequences (usually sequence 3). Almost half the students, 49%, could use substitution to continue the fraction sequence and find decimal equivalents for both patterns. Several students, about 40%, could also describe the pattern for sequence 3. 22% of the students could meet all the demands of the task including recognizing the pattern for the first sequence, noting that the numbers were all about 7 tenths or that the decimal values increased and decreased. 23% of the students scored no points on this task. 64% of the middle school students with this score attempted the task. *40% of the high school students with this score attempted the task.*

Based on teacher observation, this is what algebra students knew and were able to do:

- Substitute values into algebraic expressions
- Convert fractions to decimals
- Calculate the value of a complex expression

Areas of difficulty for algebra students:

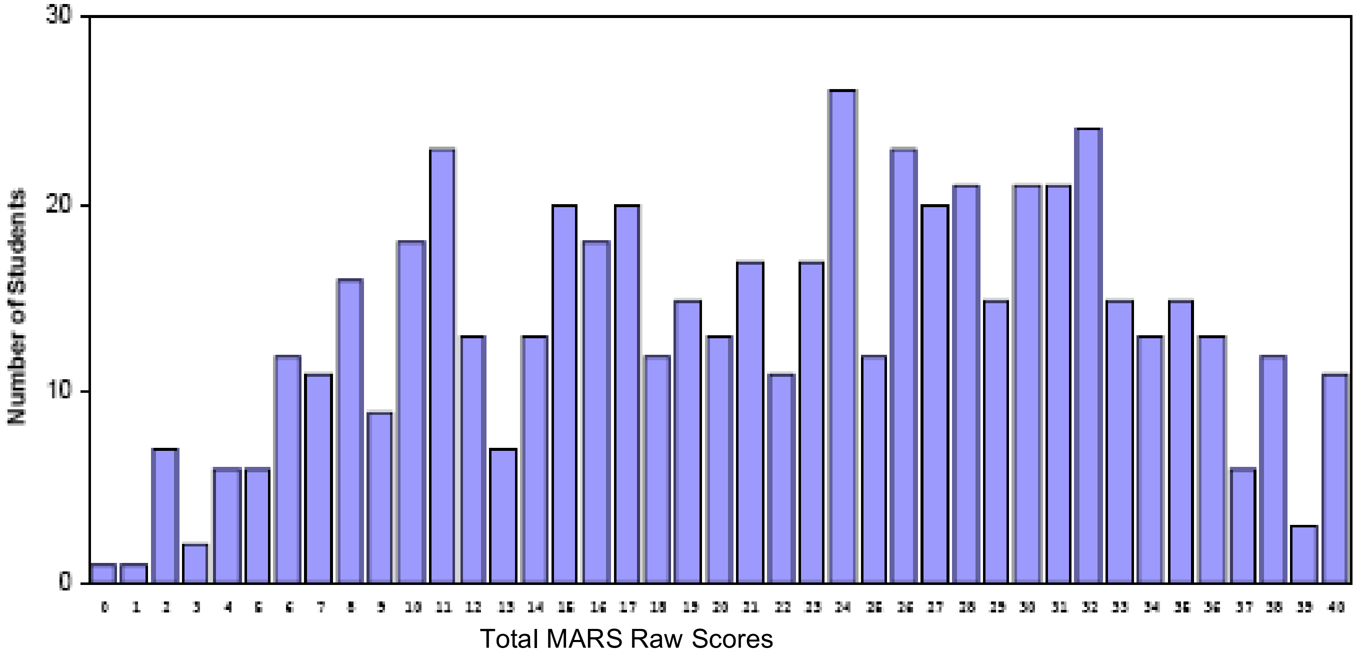
- Using rounding to see that values are almost all the same
- Comparing fractions and decimals
- Decimal place value
- Describing a pattern

Implications for Instruction:

Using a recursive rule is challenging for many students, yet it is a simple task for any computer. This process requires organization and an understanding of variable. For each new step (iteration) the variable changes (varies with a new value). Students have many experiences with substituting single values into equations. This task requires that students use those result values to then re-substitute those new values into the equation, thus generating a new sequence of terms. Students are also usually familiar with growing or decaying sequences mostly linear but also non-linear. Students are less familiar with sequences that converge or oscillate. Providing students with experience around all these different types of patterns is important in algebra.

Overall Frequency Distribution by Total MARS Raw Scores, Course 2

Mean: 21.96 StdDev: 9.89



MARS Test Performance Level Frequency Distribution Table and Bar Graph

2005 – Number of Students tested in Course 2: 559

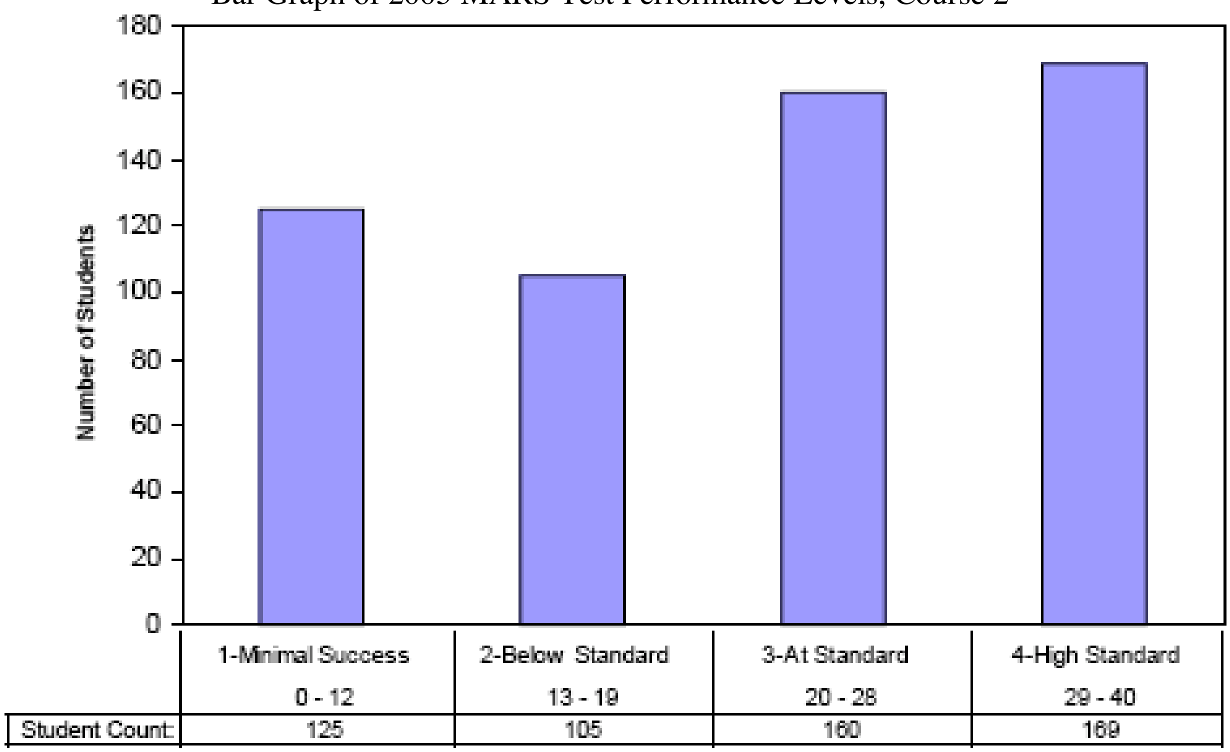
Frequency Distribution of MARS Test Performance Levels, Course 2

Year of Testing

	2000		2001		2002	
Perf. Level	% at	% at least	% at	% at least	% at	% at least
1	33%	100%	32%	100%	44%	100%
2	45%	67%	52%	68%	40%	56%
3	18%	22%	14%	16%	9%	16%
4	4%	4%	2%	2%	6%	6%

	2003		2004		2005	
Perf. Level	% at	% at least	% at	% at least	% at	% at least
1	25%	100%	6%	100%	22%	100%
2	39%	75%	27%	94%	19%	78%
3	23%	37%	29%	67%	29%	59%
4	14%	14%	38%	38%	30%	30%

Bar Graph of 2005 MARS Test Performance Levels, Course 2



*Total MARS raw scores is the summation of Tasks 1 through 5 on the MARS test.

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. Groups with Ns of less than 5 students are not reported.

Figure 9.1 Box and whisker plot of Total MARS Raw Scores by Ethnicity

Course: 2

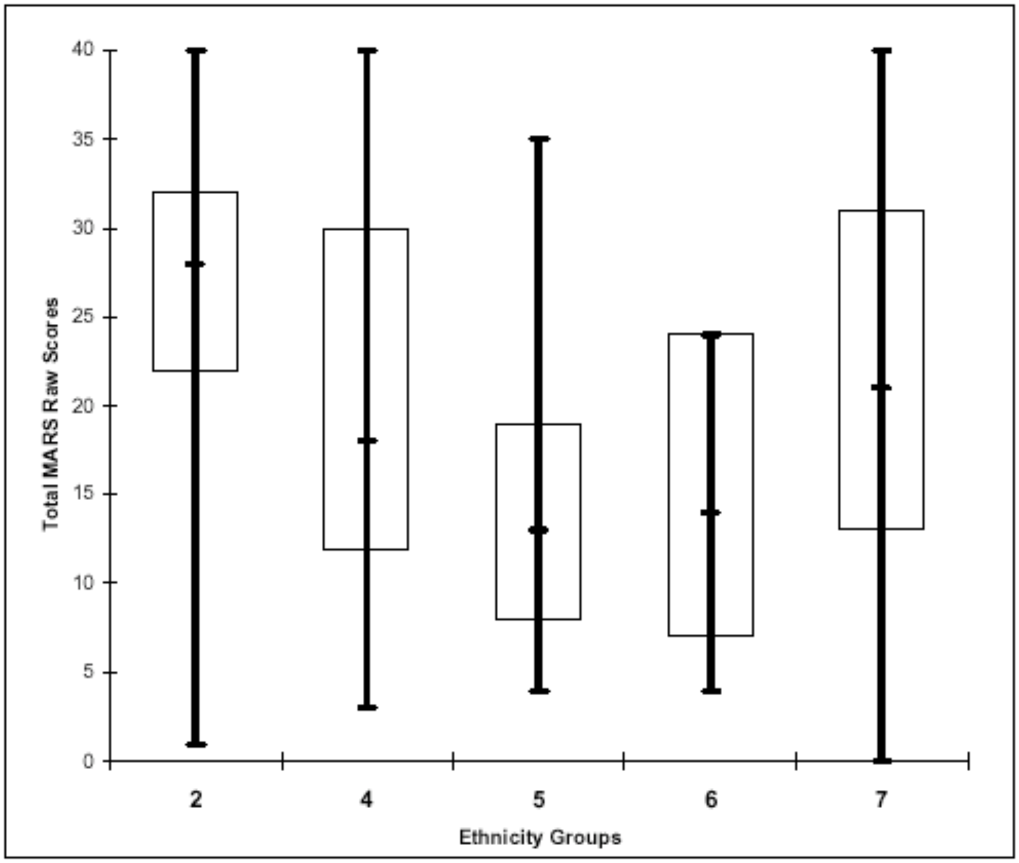


Table 9.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
2	Asian/Asian American	267
4	Filipino	16
5	Hispanic/Latino	26
6	African American	9
7	White (Not Hispanic)	156

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

Figure 20.1 Distribution of sampling means by Ethnicity

Course: 2

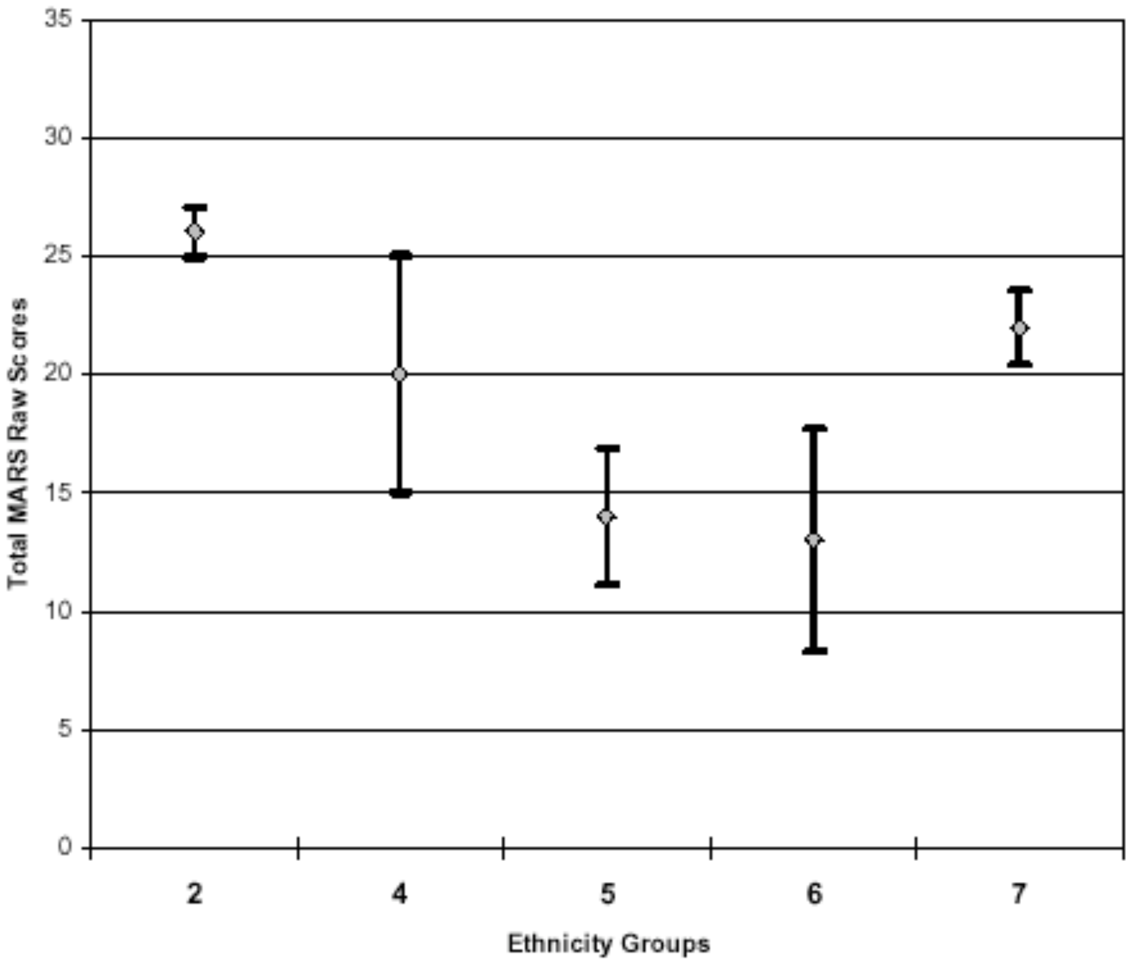


Table 20.1 Student count for Ethnicity

Ethnicity Groups	Ethnicity	Student Count
2	Asian/Asian American	267
4	Filipino	16
5	Hispanic/Latino	26
6	African American	9
7	White (Not Hispanic)	156

Distribution of Sampling Means
Grade 10
Ethnicity

In this section, test scores are compared across different ethnic groups[51]. 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 Filipino students are significantly lower than the scores of Asian/Asian American students, but not significantly different from the scores of students in any other ethnic group.

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

Students in the African American ethnic group scored significantly lower than Asian/Asian American and White students, but their scores were not significantly different from those of students in any other category.

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

[51] 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.

Figure 9.2 Box and whisker plot of Total MARS Raw Scores by Home Language

Course: 2

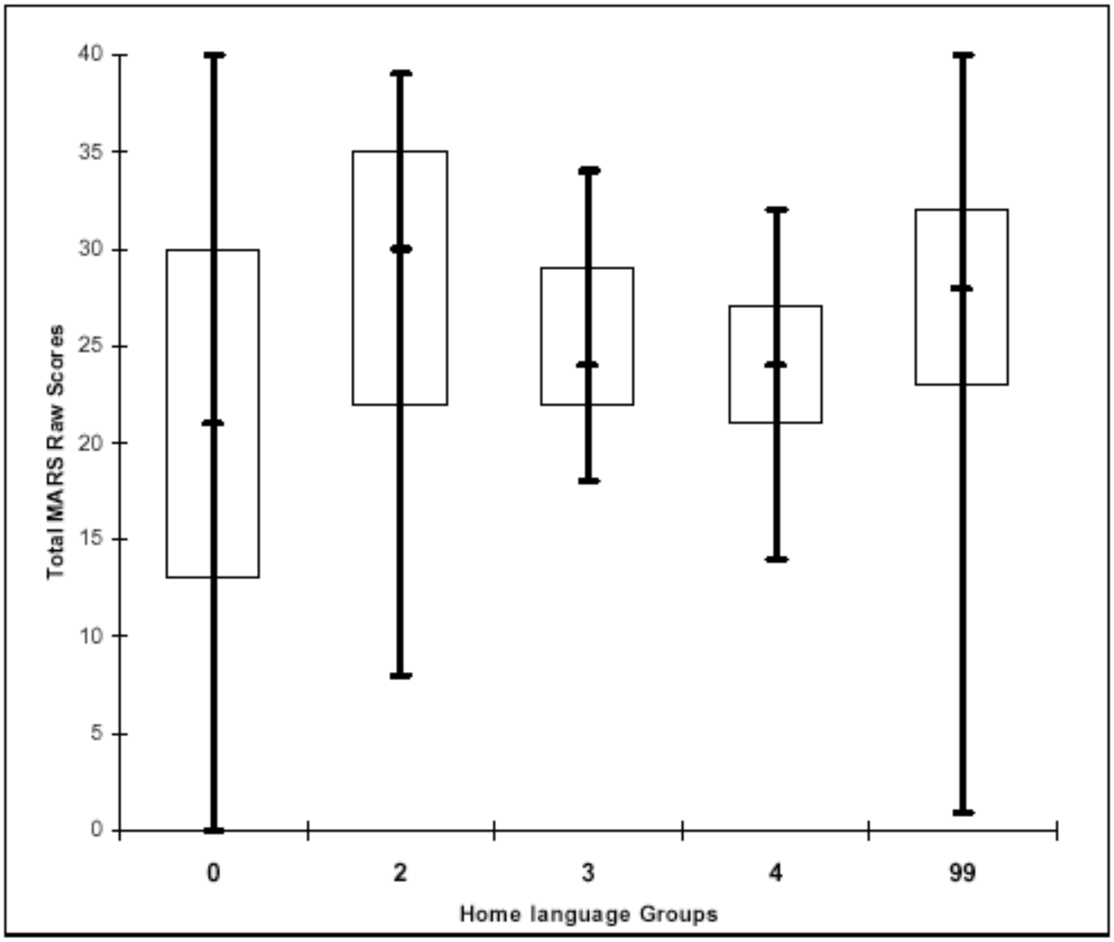


Table 9.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	267
2	Vietnamese	16
3	Cantonese	19
4	Korean	13
99	Others/Unknown	161

Figure 20.2 Distribution of sampling means by Home Language

Course: 2

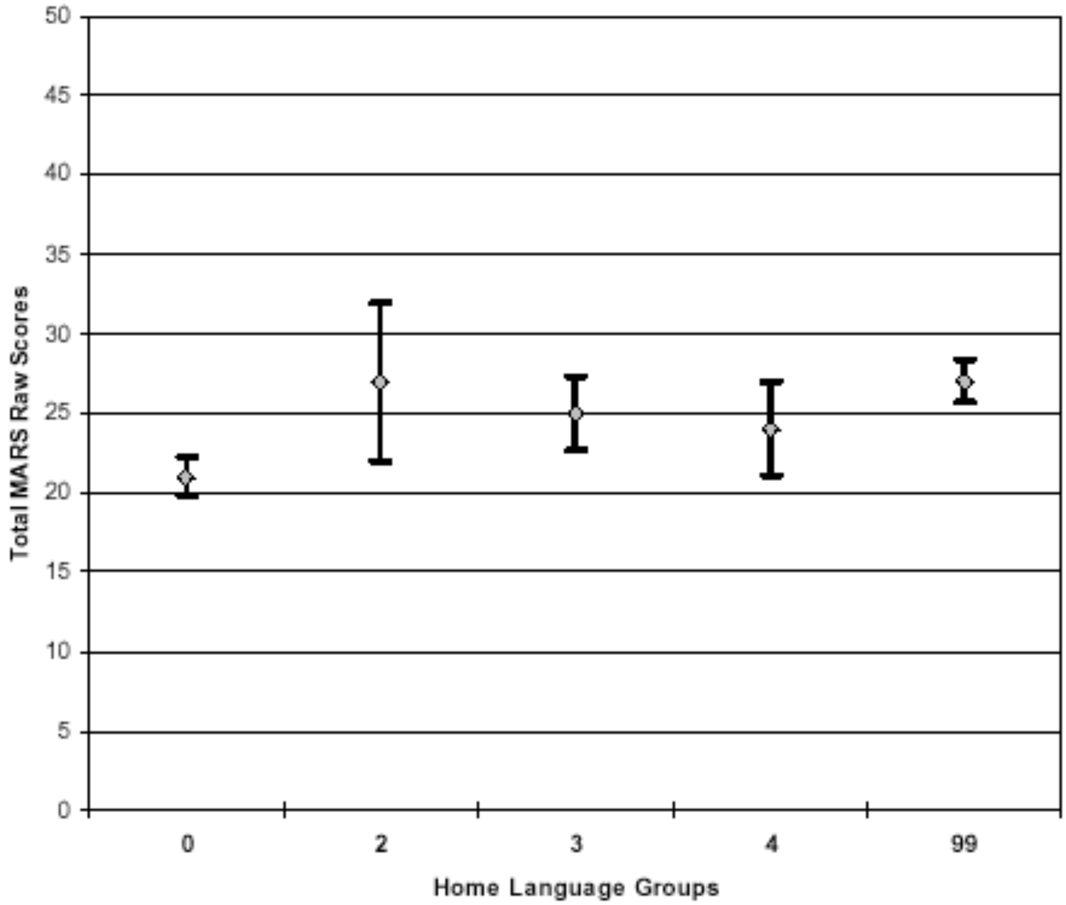


Table 20.2 Student count for Home Language

Home Language Group	Home Language	Student Count
0	English	267
2	Vietnamese	16
3	Cantonese	19
4	Korean	13
99	Others/Unknown	161

Distribution of sampling means
Grade 10
Home Language

In this section, test scores are compared across groups of students who speak different languages at home[52]. 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.

Because of the relatively small numbers of students in the different home language groups, there are few significant differences for this grade level. Specifically, students who speak English as a home language scored lower than those who are in the "Other" category of home language. No other differences are statistically significant.

[52] 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.

Figure 9.3 Box and whisker plot of Total MARS Raw Scores by Parent Education

Course: 2

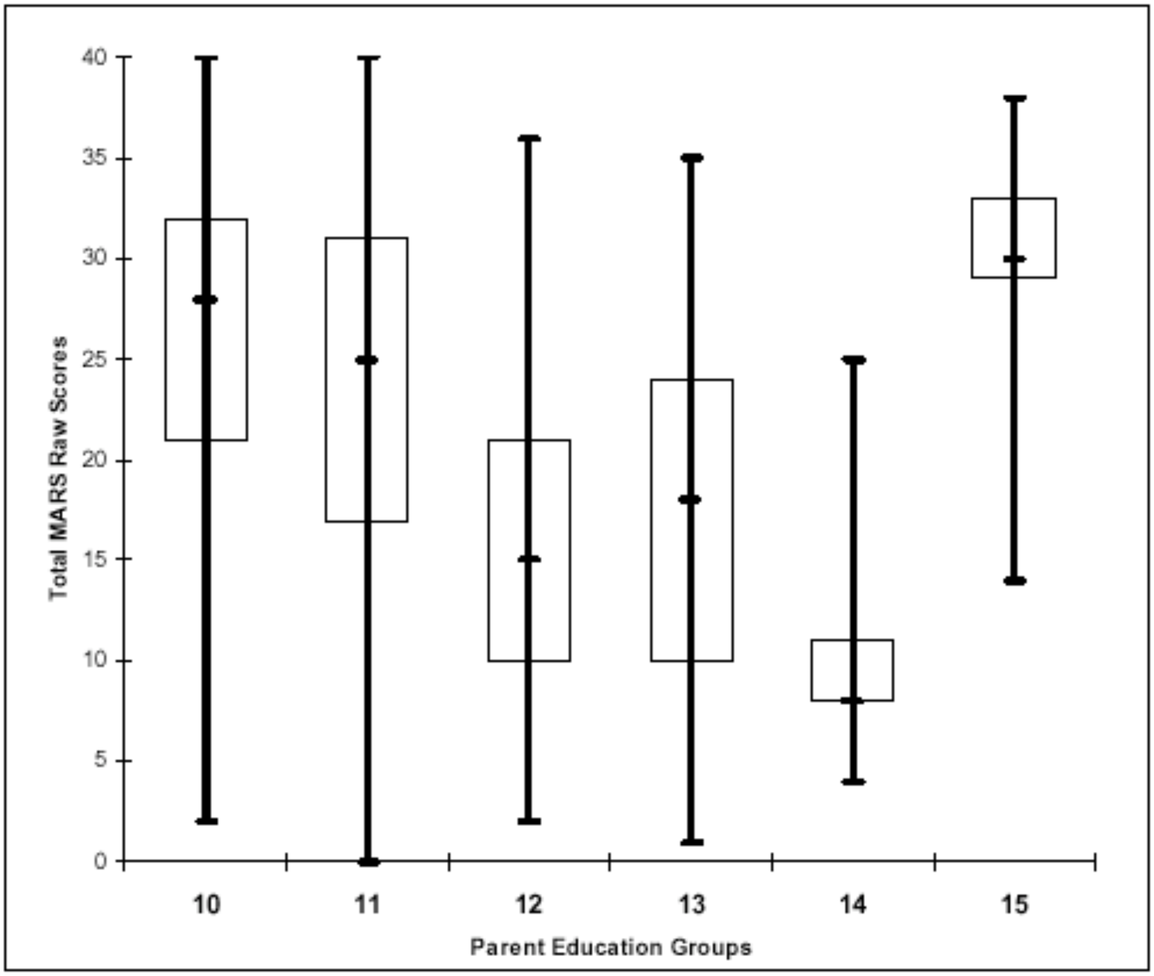


Table 9.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	213
11	College graduate	165
12	Some college	59
13	High School graduate	21
14	Not a high school graduate	5
15	Others/Unknown	13

Figure 20.3 Distribution of sampling means by Parent Education

Course: 2

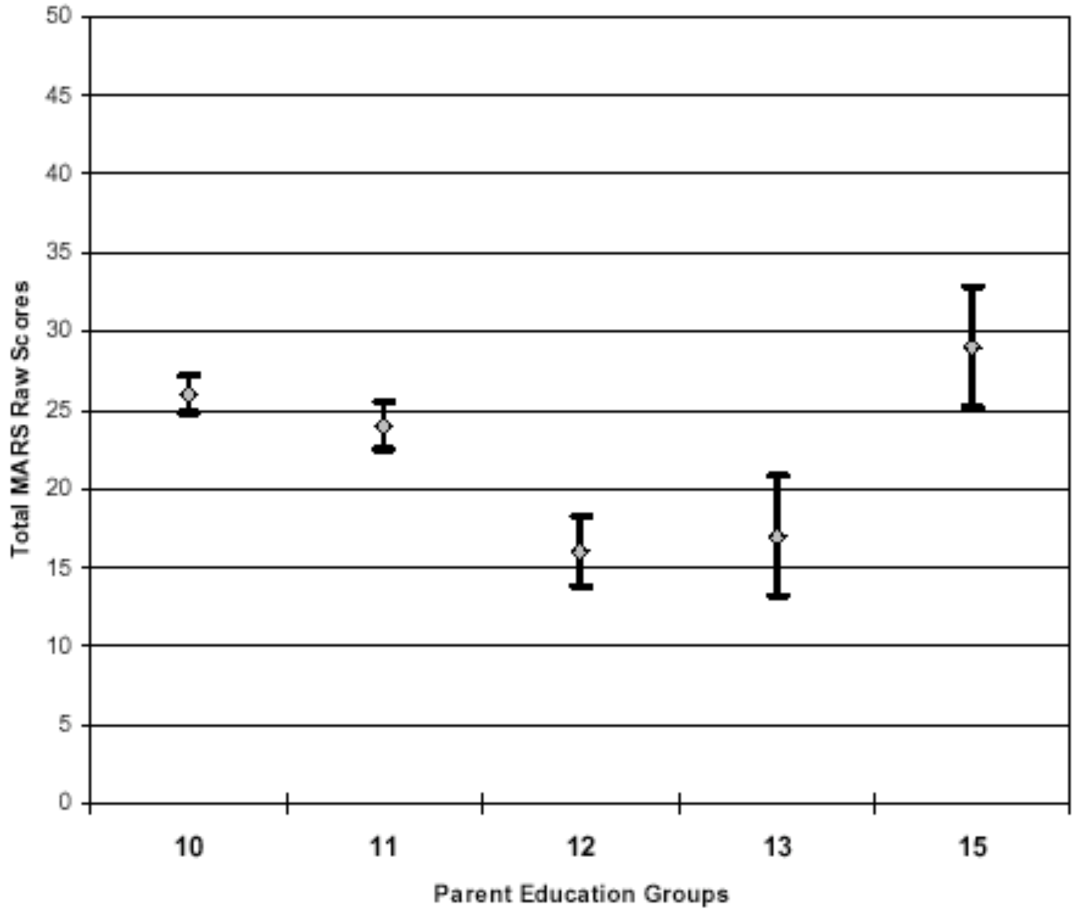


Table 20.3 Student count for Parent Education

Parent Education Group	Parent Education	Student Count
10	Graduate School	213
11	College graduate	165
12	Some college	59
13	High School graduate	21
14	Not a high school graduate	5
15	Others/Unknown	13

Distribution of Sampling Means
Grade 10
Parent Education

In this section, test scores are compared across groups of different levels of parent education[53]. 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 breakdown of statistically significant differences for the scores of students with parents in different education categories follows a pattern in which higher levels of parent education are associated with higher student scores. However, the difference in scores between students whose parents had a graduate education and the scores of students whose parents were college graduates is not significantly different. And, the scores of students with parents in the “some college” category are not significantly different from those whose parents are or are not H.S. graduates. Again, the lack of significant differences is most likely attributable to the small numbers of students in some of the parent education categories.

Figure 9.4 Box and whisker plot of Total MARS Raw Scores by Gender

Course: 2

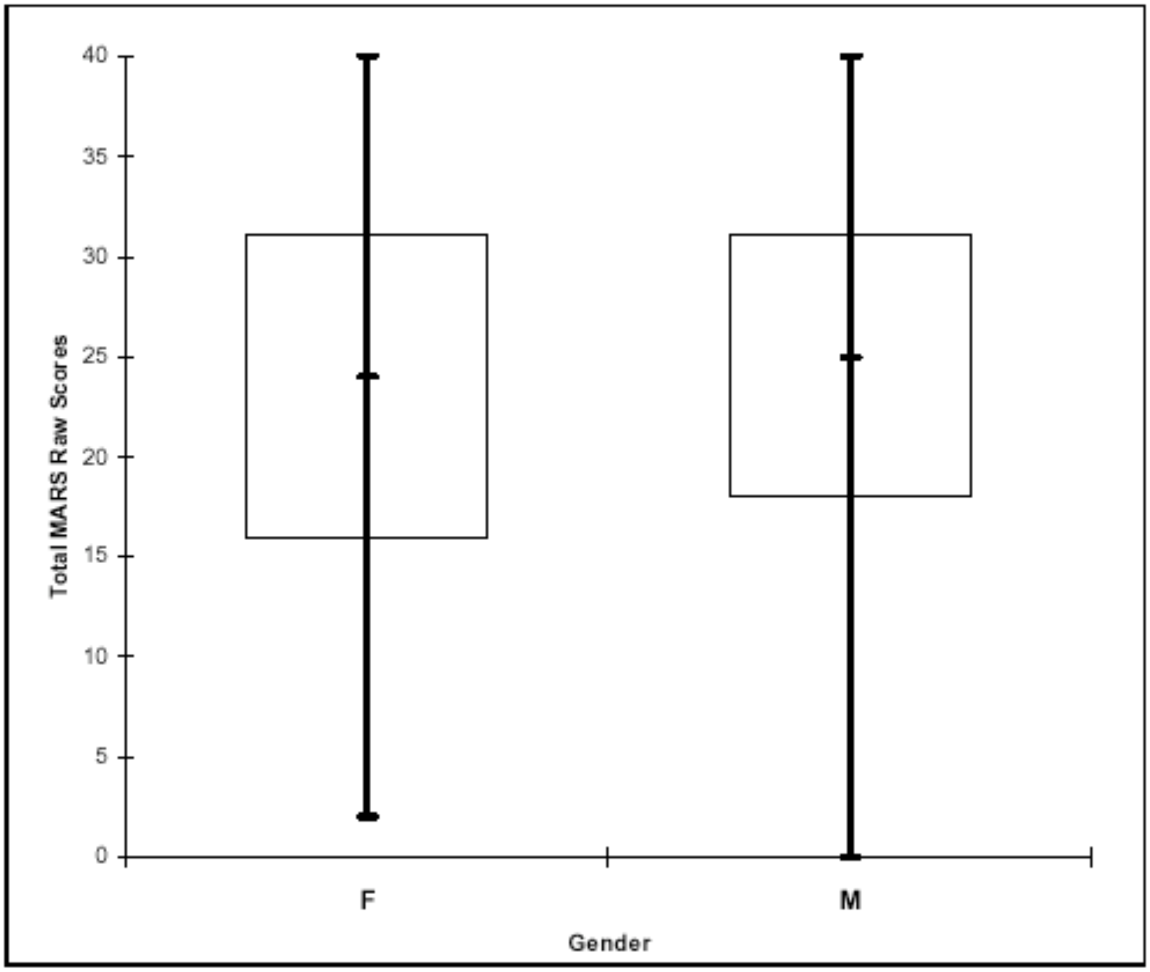


Table 9.4 Student count for Gender

Gender	Student Count
Female	233
Male	243

Figure 20.4 Distribution of sampling means by Gender

Course: 2

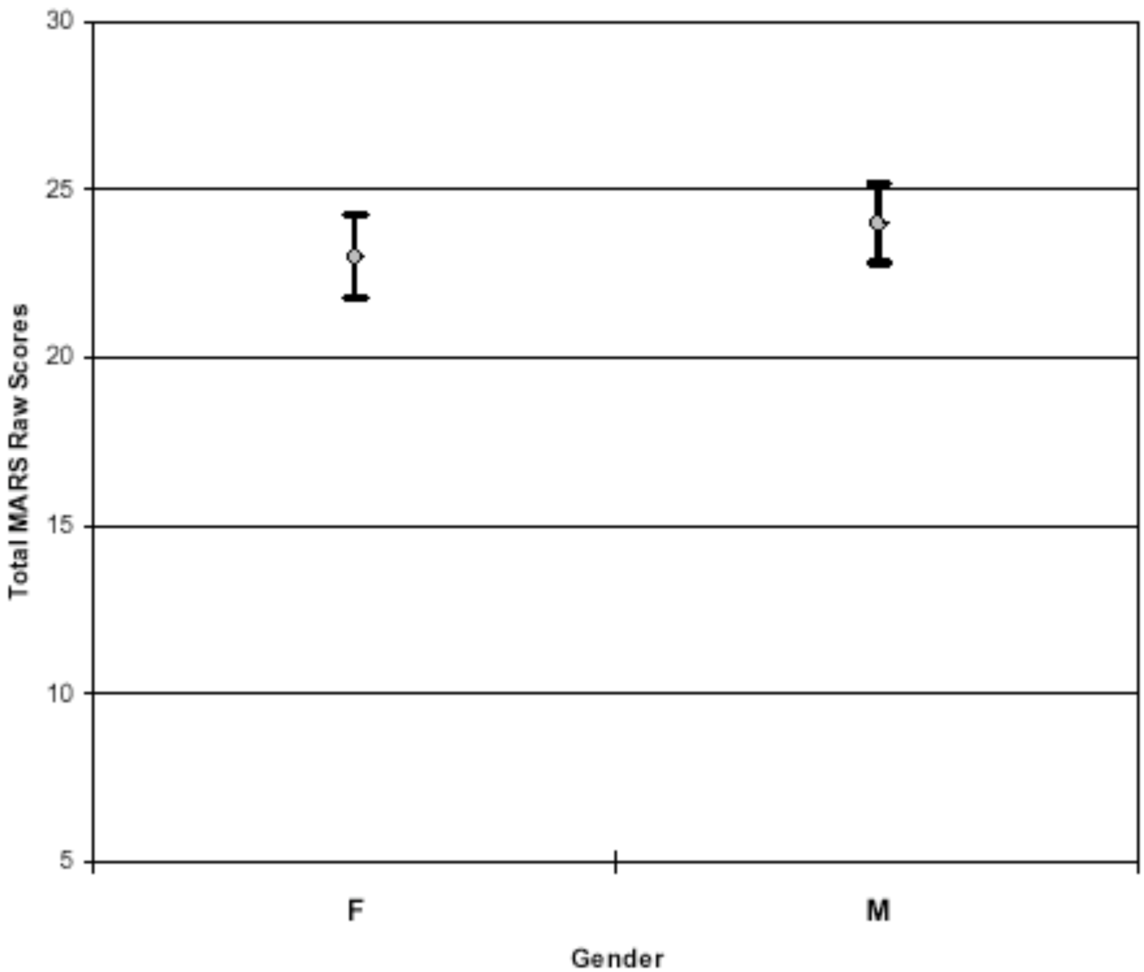


Table 20.4 Student count for Gender

Gender	Student Count
Female	233
Male	243

Distribution of sampling means
Grade 10
Gender

In this section, test scores are compared across gender[54]. 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.

Figure 9.5 Box and whisker plot of Total MARS Raw Scores by Language Fluency

Course: 2

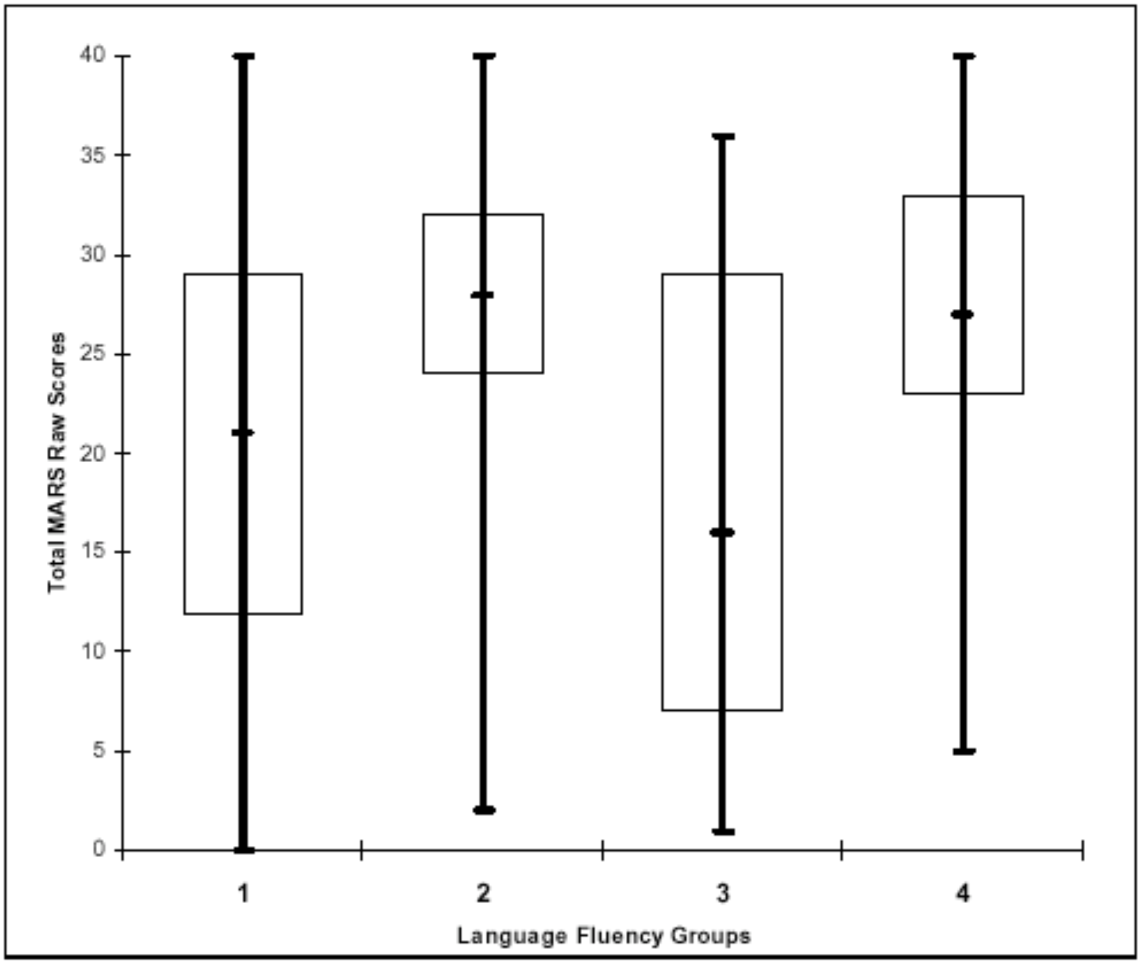


Table 9.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	275
2	Initially Fluent (I-FEP)	129
3	English Learner	11
4	ReDesignated (R_FEP)	61

Figure 20.5 Distribution of sampling means by Language Fluency

Course: 2

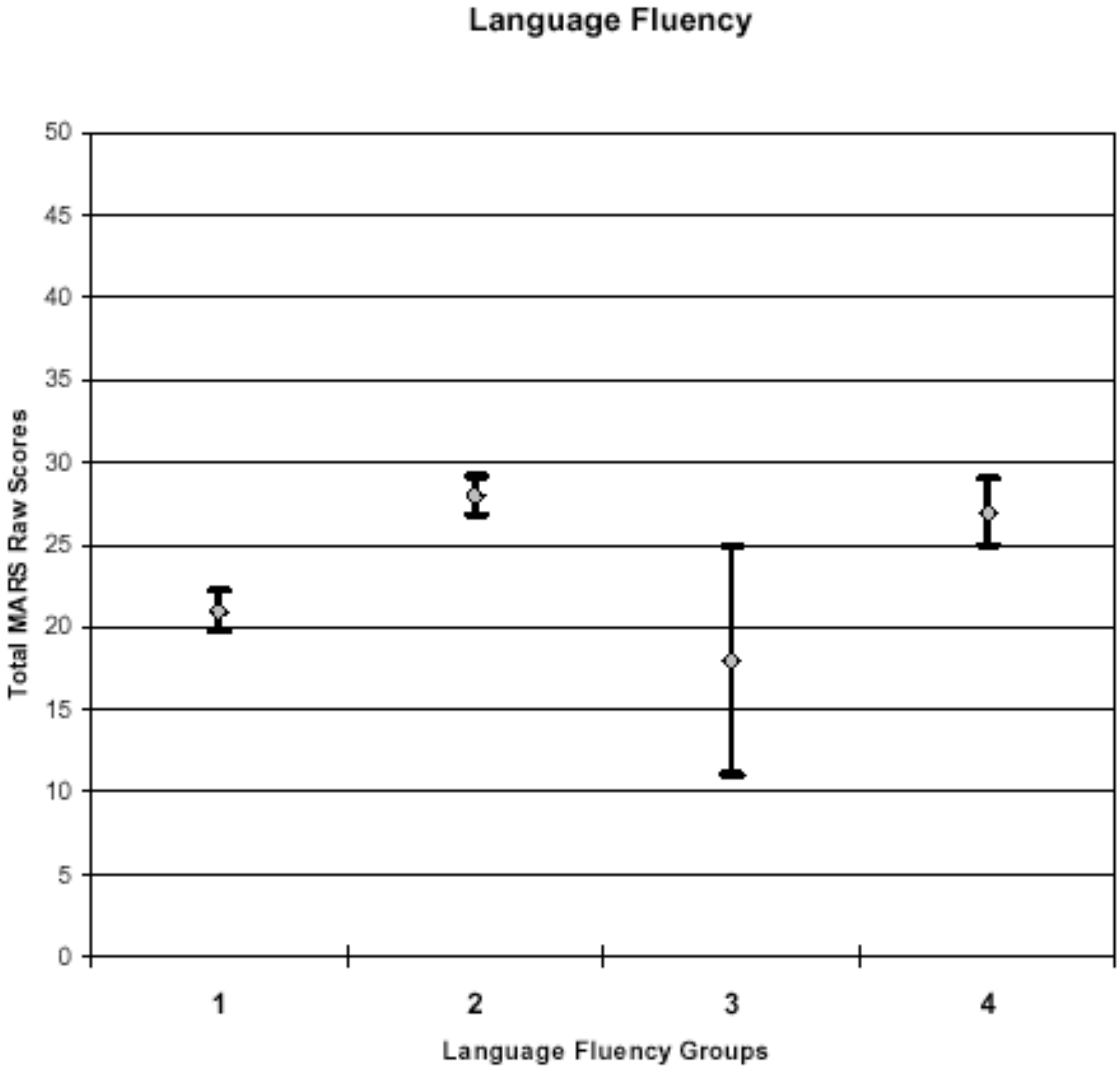


Table 20.5 Student count for Language Fluency

Language Fluency Group	Language Fluency	Student Count
1	English only	275
2	Initially Fluent (I-FEP)	129
3	English Learner	11
4	ReDesignated (R_FEP)	61

Distribution of sampling means
Language Fluency
Grade 10

In this section, test scores are compared across groups of different language fluency[55]. 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 students described as Full English Proficiency (FEP) and Re-designated FEP There is no significant difference between the scores of students as English Only and any other group.

The scores of students with Full English Proficiency (FEP) are significantly higher than those students described as English Only or English Learner. There is no significant difference between the scores of students as Full English Proficiency (FEP) and any other group.

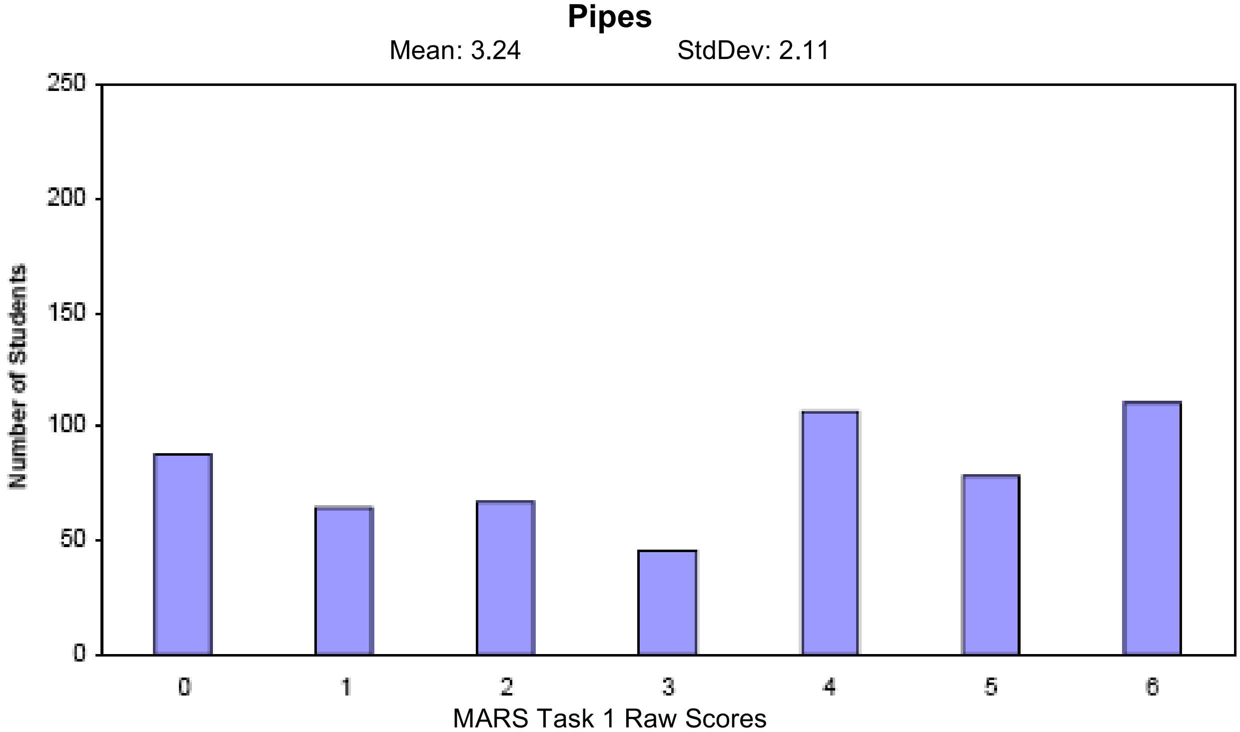
The scores of students in English Learner category are significantly lower than the scores of students described as FEP and Re-designated FEP

The scores of students in Re-designated FEP are significantly higher than those students described as English Only or English Leaner. There is no significant difference between the scores of students as Re-designated FEP and any other group.

Course 2 Task 1 Pipe

Student Task	The task asks students to use Pythagorean Theorem to find the radius of a large pipe with 4 smaller pipes fitting inside the larger pipe.
Core Idea 3 Algebra Properties and Representations	• **Represent and analyze mathematical situations and structures using algebraic symbols..** • Solve equations involving radicals and exponents in contextualized problems such as use of Pythagorean theorem.
Core Idea 3 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.** • Draw and construct representations of two- and three-dimensional geometric objects using a variety of tools. • Visualize three-dimensional objects from different perspectives and analyze their cross sections

Frequency Distribution for Task 1 – Course 2 – Pipes



Score:	0	1	2	3	4	5	6
Student Count	88	64	67	45	107	78	110
% < =	15.7%	27.2%	39.2%	47.2%	66.4%	80.3%	100.0%
% > =	100.0%	84.3%	72.8%	60.8%	52.8%	33.6%	19.7%

The maximum score available for this task is 6 points.
The minimum score for a level 3 response, meeting standards, is 4 points.

Many students (about 84%) knew the mathematical terms for center of a circle and could use the hint to make a correct diagram. Most students (72%) could use the definition of radius to correctly find the dimensions of the square they constructed. About half the students, 52%, could interpret the purpose of the diagram, see the usefulness of finding the diagonal of the square, and apply the Pythagorean theorem accurately to find the length of the diagonal. 20% of the students could meet all the demands of the task, including visualizing the diagonal extending across the circle to find a diagonal for the large pipe, and halving the length of the diagonal to find the radius of the large pipe. 16% of the students scored no points on this task. 85% of the students with a score of 0 attempted the task.

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

- Use hints to draw a diagram connecting the centers of four circles
- Correctly apply the definition of a radius of a circle to find the lengths of the side of a square
- Recognize and use of Pythagorean theorem or the 45-45-90 right triangle relationships for finding the diagonal of a square

Areas of difficulty for geometry students:

- Finding square roots without calculators
- Working with combining expressions with radicals
- Following an extended reasoning chain through to its conclusion (Students stopped with finding the diagonal of the square, without asking or identifying the question "What is the problem asking?"
- Decomposing a complex diagram into smaller units and seeing the relationship between those parts to the question being asked

Successful students:

- Gave reasons for their calculations. They answered questions like, "Why am I doing this step? What doe this calculation represent?"
- Often made extra diagrams to show their thinking.

Implications for Instruction:

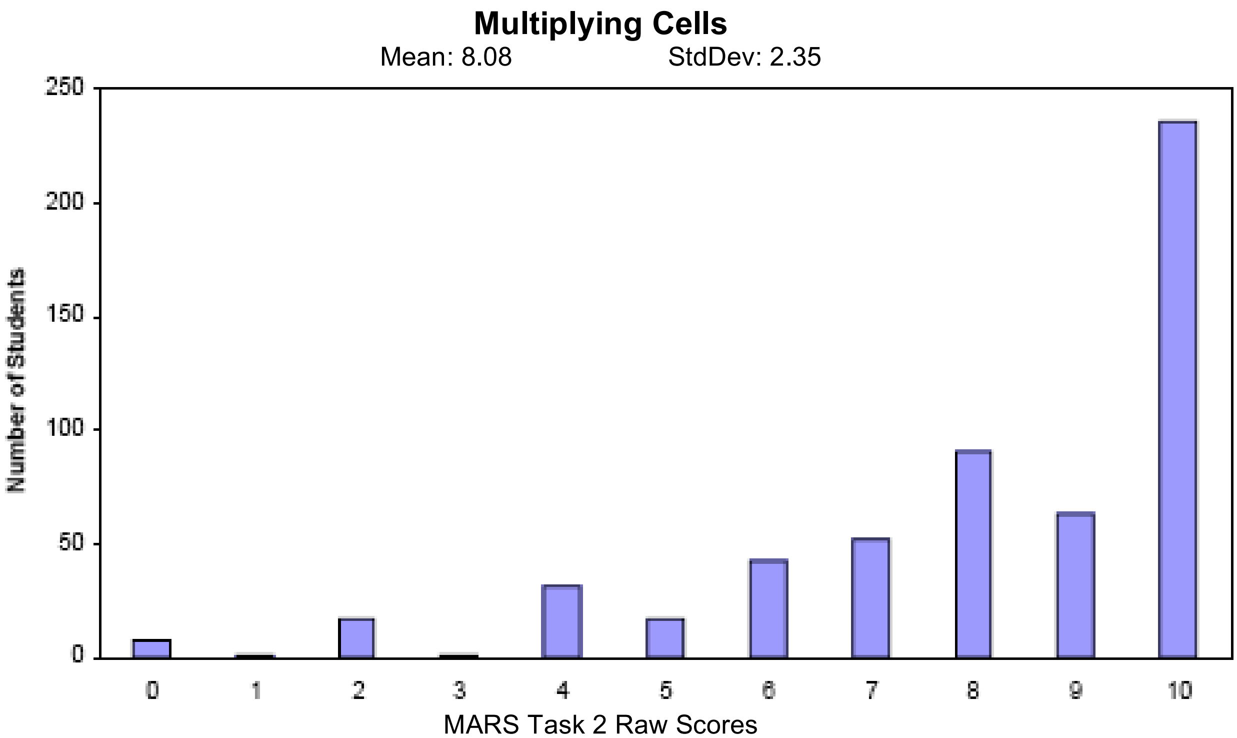
Often students showed mathematical understanding but failed to answer the questions posed. A significant number of students answered a different question than was posed, such as the length of the diagonal of the square not the radius of the pipe. Encouraging students to justify answers to questions will help students focus on addressing the question that was posed. Students often struggle with understanding the properties of circles. The concept of tangent circles is often addressed at surface level. Students need a deeper understanding of the relationships of radii of tangent circles and tangent lines. In order to correctly reason through this problem, a student must realize that the centers of the four form the vertices of a square. The hint of the problem probably assisted a significant number of students in accessing the problem. The logic that two radii from two tangent circles form a straight line is justified because they are both perpendicular to common tangent line. It is also not trivial that the quadrilateral formed by connecting the four congruent circles is a square. It would be an excellent exercise to have students justify why the hint is valid.

Many students were successful by using the Pythagorean Theorem. Other students used the knowledge of the sides of 45-45-90 right triangles. Either approach was acceptable. Some students struggled with using irrational numbers. Experiences with converting between the exact number (integer plus a radical) and a decimal approximation is important for many students. Some students then struggled after finding the length of the diagonal. Students need more experience in problem solving and working on non-routine problems in geometry.

Course 2 Task 2 Multiplying Cells

Student Task	The task asks students to find and extend a number pattern based on powers of two and convert between pattern and time periods.
Core Idea 1 Functions	• **Understand patterns, functions, and relations.** • Understand and compare the properties of functions, including linear, quadratic, reciprocal and exponential functions.

Frequency Distribution for Task 2 – Course 2 – Multiplying Cells



Score:	0	1	2	3	4	5	6	7	8	9	10
Student Count	8	1	17	1	31	17	42	52	91	63	236
% < =	1.4%	1.6%	4.7%	4.8%	10.4%	13.4%	20.9%	30.2%	46.5%	57.8%	100.0%
% > =	100.0%	98.6%	98.4%	95.3%	95.2%	89.6%	86.6%	79.1%	69.8%	53.5%	42.2%

The maximum score available on this task is 10 points
The minimum score for a level 3 response, meeting standards, is 6 points.

Most students, about 95%, could identify the pattern for growing cells and for the powers of 2 and use that information to correctly fill out the table provided. Many students, about 86%, could also make sense of the time-intervals to extend the pattern from 2 hours to 3 hours using appropriate powers of 2. More than half the students, 67%, could also reason between hours, 20-minute intervals, the appropriate powers of 2, and calculate accurately with exponents. 42% of the students met all the demands of the task. Only 2% of the students scored no points on the task. All students in the sample who scored 0 attempted the task.

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

- Identify an exponential growth pattern
- Express the pattern in numbers and exponential notation
- Identify the time-interval and extend it appropriately
- Quantify number equivalents for exponential expressions

Areas of difficulty for geometry students:

- Carrying out calculations for higher powers of 2
- Using rules of exponents appropriately
- Working backwards from a numerical value to a power of two and then converting that power of 2 into an appropriate period of time

Successful students tended to:

- Use calculators or use rules of exponents to shorten the number of individual calculations needed
- Label their calculations to show the meaning of each step in the operation
- Show a clear understanding of the time interval and its relation to the problem

Implications for Instruction:

Geometry is an excellent context to explore non-linear growth. Linear growth dominates the algebra one curriculum. Only in the second semester of algebra one are polynomial functions introduced and used. Polynomial functions are especially useful for applications with area and volume. Exponential functions differ still from the polynomial functions. Yet many scaling problems in geometry have exponential functions. Teachers of geometry should include these functions as they explore the relationship of geometric and algebraic representations of mathematics. Examples of tasks that are both geometric and exponential are:

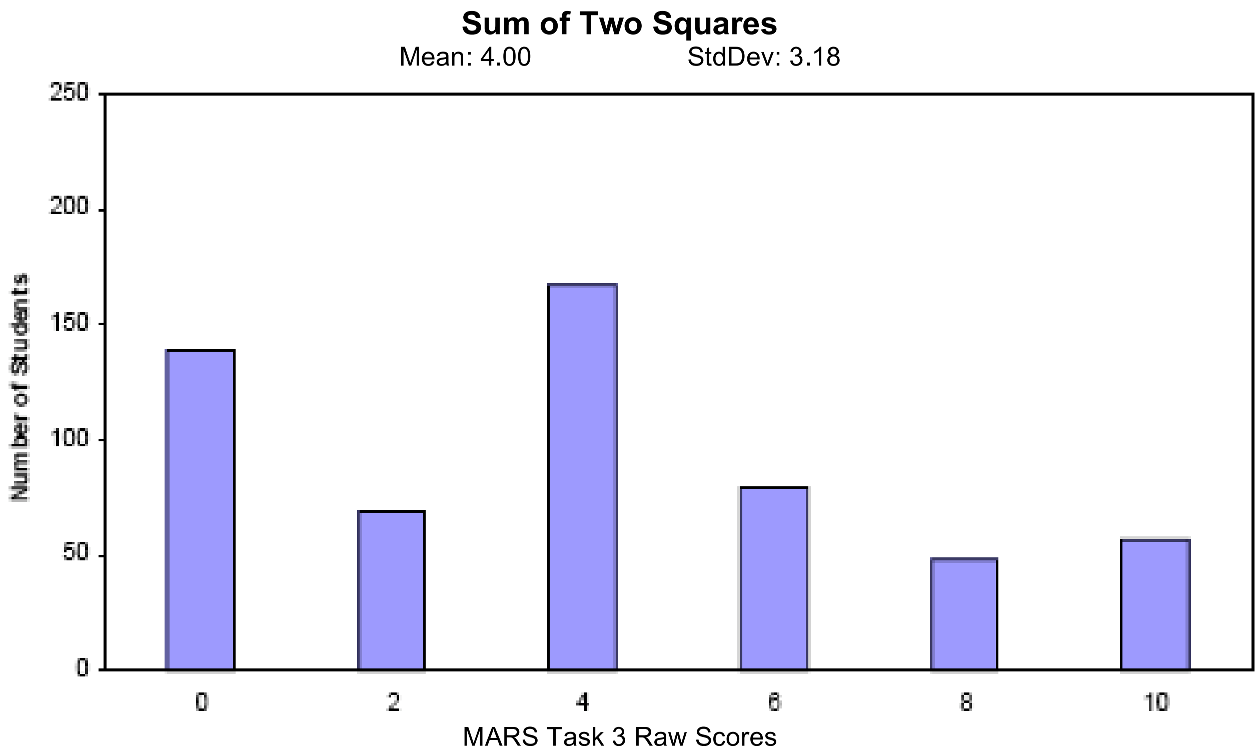
- As a figure continues to double in length, what happens to the perimeter?
- What is the functional relationship between units in different dimensions?
- How does the fractal "Sierpinski's Triangle" grow?
- Examine the change in size of a series of right triangles whose one leg grows exponentially while the other leg remains constant.
- Graph the functions $y = 2^x$ and $y = x^2$, how are they alike, how are they different and where do they intersect?

Exploring tasks described above will provide students with a visual picture of exponential growth. It will also familiarize them in non-linear functions.

Course 2 Task 3 Sum of Two Squares

Student Task	The task asks students to find and use a pattern, including making an algebraic transformation to prove why a statement is true.
Core Idea 2 Mathematical Reasoning and Proofs	**Employ forms of mathematical reasoning and proof appropriate to the solution of the problem at hand, including deductive and inductive reasoning, making and testing conjectures and using counter examples.** • Identify, formulate and confirm conjectures.
Core Idea 3 Algebraic Properties and Proofs	**Represent and analyze mathematical situations and structures using algebraic symbols.**
Core Idea 1 Functions	**Understand patterns, relations, and functions.** • Understand and perform transformations on functions.

Frequency Distribution for Task 3 – Course 2 – Sum of Two Squares



Score:	0	2	4	6	8	10
Student Count	139	69	167	79	48	57
% < =	24.9%	37.2%	67.1%	81.2%	89.8%	100.0%
% > =	100.0%	75.1%	62.8%	32.9%	18.8%	10.2%

The maximum score available for this task is 10 points.
The minimum score for a level 3 response, meeting standards, is 5 points.

More than half the students, about 62%, identify the constraints of the conjecture, follow the logic of the conjecture, substitute appropriate values into the equation and find its value, and then use knowledge of square numbers to find two perfect squares which total to that value. About half the students could find two examples to illustrate the conjecture. Only 33% of the students could see an algebraic relationship between the starting integers and the final numbers, which were squared. 10% of the students could meet all the demands of the task including expressing the relationship using algebraic notation and do an algebraic proof to show that the conjecture was true for all cases. 24% of the students scored no points on this task. 71% of the students, who scored 0, attempted the task.

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

- Identify constraints of a problem
- Substitute integers into an algebraic expression and make the corresponding calculations
- Find two perfect squares to add to a given total

Areas of difficulty for geometry students:
- Finding an algebraic relationship between starting integers and final square numbers
- Translating written relationships into algebraic or symbolic notation
- Using algebra to prove a conjecture

Successful students tended to:
- Think about finding examples that were different cases of the conjecture.
- Understand that numeric examples don't make a proof.
- Knew that an algebraic proof is being able to change one side of equation to match the conjecture.

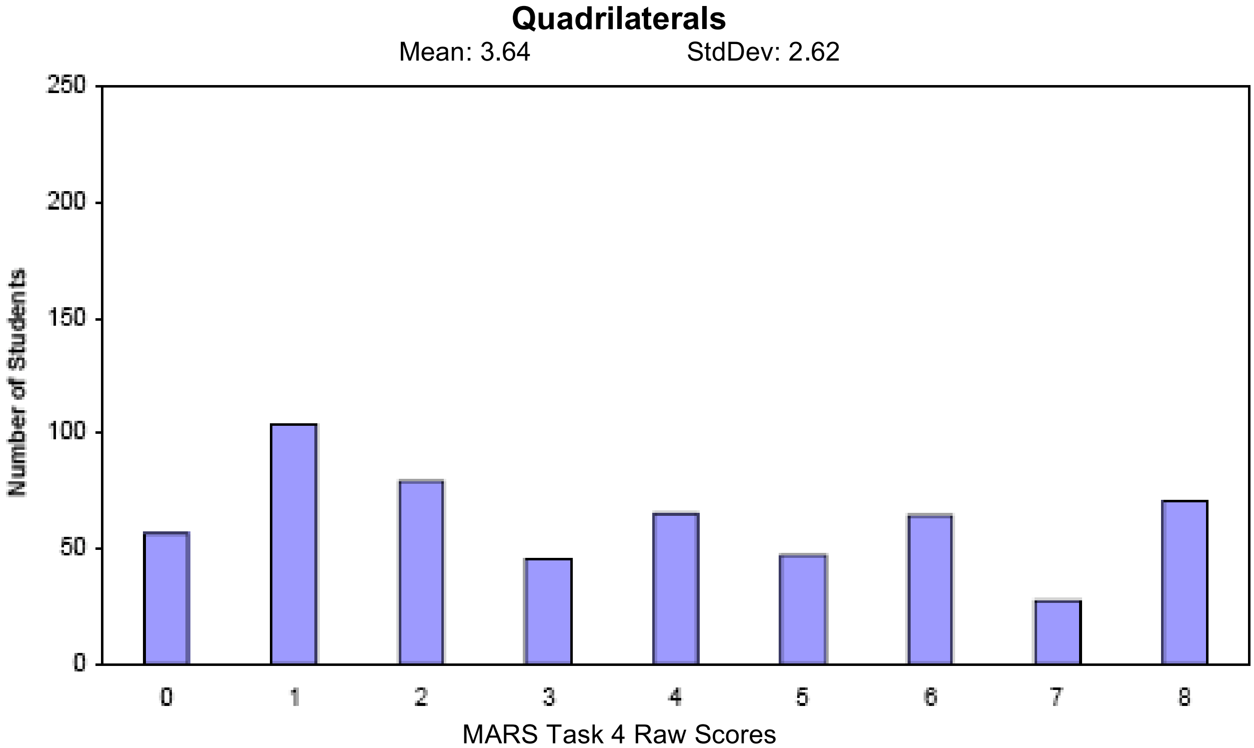
Implications for Instruction:

Justification and proof is the cornerstone of a high school geometry class. Students develop facility with proof using a variety of context including logic, algebraic and geometric. In algebra one, students should first experience algebraic proof using deductive reasoning and the axioms of numbers. There needs to be an early and increased emphasis in justification and math reasoning. Using deductive reasoning to explain numeric, algebraic and geometric conjectures should be common and central to every college prep math course. Students need to internalize deductive reasoning for themselves. Often this means letting students struggle with how to begin and what are logical steps. Students often find working backwards as a helpful strategy. Early on students may be introduced to proof by filling in missing statements or reasons, but this is insufficient to build self- reliance in reasoning and logic. Learning to justify and prove is an ongoing process that must reside within the students. Providing these learning experiences is paramount for a successful foundation in mathematics.

Course 2 Task 4 Quadrilaterals

Student Task	The task asks students to use geometric properties to solve a problem about an inscribed quadrilateral.
Core Idea 4 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.** • Explore relationships among classes of two-dimensional geometric objects, make and test conjectures about them, and solve problems involving them.

Frequency Distribution for Task 4 – Course 2 – Quadrilaterals



Score:	0	1	2	3	4	5	6	7	8
Student Count	57	104	79	45	65	47	64	27	71
% < =	10.2%	28.8%	42.9%	51.0%	62.6%	71.0%	82.5%	87.3%	100.0%
% > =	100.0%	89.8%	71.2%	57.1%	49.0%	37.4%	29.0%	17.5%	12.7%

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

Most students, about 90%, could decompose the shape and find at least one relationship between lines AC and PQ, usually that they were parallel. More than half the students could identify both relationships between the two lines. Less than half the students, 37%, could identify quadrilateral PQRS as a parallelogram and give at least some partial reasoning for why this was true. About 29% of the students could provide a complete justification for why quadrilateral PQRS was a parallelogram. Only 13% of the students could meet all the demands of the task, including reasoning about the conditions, which would make PQRS a square, and working backwards to determine what would need to be true about the diagonals of the outer quadrilateral ABCD. 10% of the students scored no points on this task. 50% of the students who scored zero did not attempt the task.

Based on teacher observations, this what geometry students know and are able to do:

- Recognize and reason about parallel lines embedded within a diagram and matching constraints of design
- Reason about the length of a line connecting the midpoint of a triangle and the base of the triangle
- Decompose shapes or figures into known or familiar shapes to help solve a problem

Areas of difficulties for geometry students:

- Recognizing a quadrilateral as a parallelogram from properties instead of diagrams
- Justifying a conjecture or hypothesis by picking out relevant facts and putting them together into a coherent logic chain
- Understanding the difference between special cases and general cases of figures
- Knowing that diagrams are often not drawn to scale, so that conjectures must be made using axioms and theorems rather than "looks like"
- Understanding properties of quadrilaterals and their diagonals
- Geometric terms like congruent, similar, rectangle, trapezoid, etc.

Implications for Instruction:

Justification is a central concept in high school geometry. It separates high school geometry from geometry learned in younger grades. Students must learn to be flexible in reasoning and communicating arguments in mathematics. Unfortunately over 50% of the students were unsuccessful in the major aspect of the course. Students need more experience in reasoning. The education researchers, Dina and Pierre van Hiele, characterized the levels of thinking students should engage in throughout their years of mathematical education. By high school, students should be at level 3; deduction in which the student proves theorems deductively and understands the structure of the geometric system. A student at this level can construct, not just memorize, proofs. Also students should be able to determine the possibility of developing a proof in more than one way. Students need to understand and determine the interaction of necessary and sufficient conditions and to clearly identify the distinction between a statement and its converse.

In teaching proof and justification, students should be provided with the opportunities:

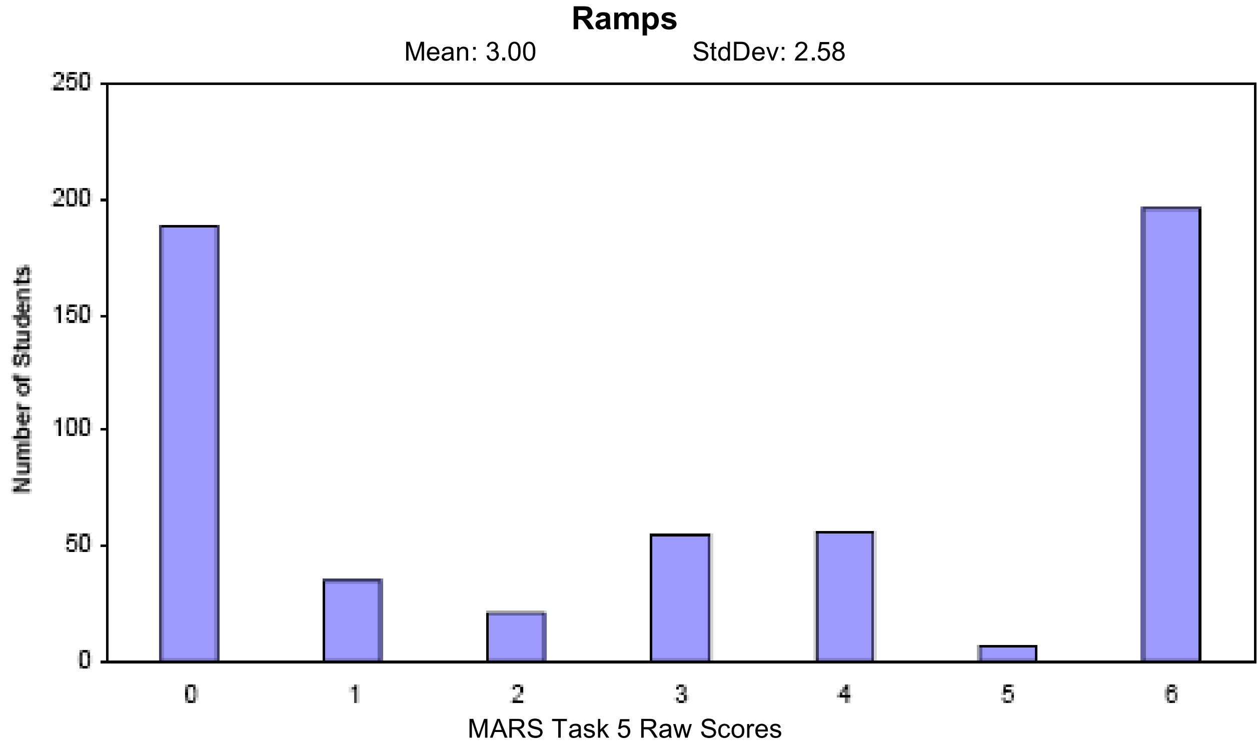
- To identify what is given and what is to be proved in a problem
- To identify information implied by a figure or by given information
- To demonstrate an understanding of meaning of undefined term, postulate, theorem, definition, etc.
- To demonstrate an understanding of necessary and sufficient conditions
- To prove rigorously the relationship developed informally at van Hiele's level 2
- To prove unfamiliar relationships
- To compare different proofs of a theorem
- To use a variety of techniques of proof
- To identify general strategies of proof
- To think about geometric thinking

Too often geometry is taught in a mechanical way. Teacher questioning is a crucial factor in directing student thinking. Students should be challenged to explain why and to think about their explanations and justifications and to consider if it could be shown another way. Thus teachers must learn to identify students' levels of geometric thought. Because the nature of a student's geometric explanations reflect that student's level of thinking, questioning is an important assessment tool. Using deductive reasoning to explain numeric, algebraic and geometric conjectures should be common and central to high school geometry. Students need to internalize deductive reasoning for themselves. Often this means letting students struggle with how to begin and what are logical steps. Students often find working backwards as a helpful strategy. Early on students may be introduced to proof by filling in missing statements or reasons, but this is insufficient to build self- reliance in reasoning and logic. Learning to justify and prove is an ongoing process that must reside within the student. Providing these learning experiences is paramount for a successful foundation in mathematics.

Course 2 Task 5 Ramps

Student Task	The task asks students to use Pythagorean Theorem to find the radius of a large pipe with 4 smaller pipes fitting inside the larger pipe.
Core Idea 4 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.**

Frequency Distribution for Task 5 – Course 2 – Ramps



Score:	0	1	2	3	4	5	6
Student Count	189	35	21	55	56	7	196
% < =	33.8%	40.1%	43.8%	53.7%	63.7%	64.9%	100.0%
% > =	100.0%	66.2%	59.9%	56.2%	46.3%	36.3%	35.1%

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

About half the students, about 56%, could use a trig function to solve a problem. Students with a score of three were equally likely to solve part 1 successfully finding the length of a ramp using the sin function or solve part 2 successfully using the tangent function to find the angle of elevation for the ramp. This problem seems to show opportunity to learn. Students were either very successful, 35% of the students could meet all the demands of the task, or scored no points. 34% of the students scored 0 points on this task. 56% of the students with a score of 0 did not attempt the problem. Less than 10% of the students received any one score point accept for 0 and 6.

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

- Use Pythagorean theorem
- Substitute values into a formula

Areas of difficulty for geometry students:

- Solving trig functions for x, when x was the denominator.
- Using trig functions to solve problems

Implications for Instruction:

This task assessed students' familiarity with basic trigonometry. Some students have not had the opportunity to learn the meaning of sine and tangent. Developing understanding of trigonometry is important. Two important fields of math come together in learning trigonometry, measurement and functions. Through trigonometry linear measures and circular measures are linked together. Without trigonometry, there would not be a way to compare the size of angle with the length of sides of polygons. Trigonometry is best learned through a circular approach. This builds conceptual understanding and connects the circle to the ratio of linear segments. It also helps students understand that there are exactly six trigonometric functions. Vocabulary often creates barriers for students learning the angular relationships. It is important that students develop meaning of the concepts to connect vocabulary terms. Rhymes, cute sayings and acronyms cannot substitute for conceptual understanding.

Appendix

NSLP

CAT/6 CST Comparisons

Stars & MAC Correlations

CST Clusters and MAC

Status of the Data

Mars Raw Score by District

Audit Analysis

Page break

CST to MAC Scatterplots

Statistical Measures

Director's Note

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 MAC 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.