

Short-Term Impact Evaluation of the Stay Alive from Education (S.A.F.E.) Program

Abstract:

Purpose: Motor vehicle crashes are a leading cause of death and injury among young people in the United States. Since many of these preventable deaths are due to risky behaviors, it is necessary to develop programs, evaluate their effectiveness, and support programs that have a positive impact on reducing injuries and loss of life. The Stay Alive from Education (S.A.F.E.) program is an educational intervention that aims to increase levels of awareness among adolescents and college students regarding the consequences of risky driving behaviors. S.A.F.E. was initiated in Miami-Dade County, Florida in 1989, and is currently used throughout Florida and in several other states. In 2001, the first scientific evaluation for S.A.F.E. program was conducted in order to assess the short-term effectiveness of the program.

Methods: A questionnaire was designed and pilot-tested with the help of a teen advisory group. The questionnaire was revised and confidentially administered to a group of male and female 10th, 11th, and 12th grade students in one Miami-Dade high school. A baseline assessment of knowledge, attitudes and behavior regarding risky behaviors and driving was conducted one week before the introduction of the SAFE program. Additional assessment was conducted one month after the S.A.F.E. intervention in order to determine whether students reported improvements in knowledge, attitudes, and behavior.

Results: T-test analyses found significant improvements in self-reported knowledge, attitudes, and behavior (all $p < 0.001$) across all students. However, further analysis found that program impact differed according to grade, gender, and race/ethnicity.

Conclusions: This short-term evaluation validated the positive impact of the S.A.F.E. program. Results will be used to further explore the longer-term impacts of the program and make changes to improve program effectiveness.

Background and Purpose:

Injuries due to motor vehicle crashes persist as the leading cause of death and injury among young people in the United States. Many of these deaths and injuries occur among teenagers, and could have been prevented had different behaviors been exercised. Several recent and highly publicized crashes involving teenagers in Miami-Dade County, Florida have increased a broad based interest in identifying new ways to reduce risk of teens while reemphasizing the need for effective programs that educate young people about the dangers of impaired driving, non-use of seat belts, and other risky driving behaviors.

The Stay Alive from Education (S.A.F.E.) program was first initiated in 1989 by two Miami-Dade paramedics who grew weary of responding to crash scenes involving young drivers and passengers. The S.A.F.E. educational intervention employs interactive discussion between emergency medical personnel and students, photographs of local crashes depicting the results of poor judgment, and a demonstration of the sequence of events including first responders and hospital procedures in a crash situation using a student volunteer. The aims of this evaluation were: (i) to design and pilot-test a standard questionnaire that could be used to determine whether the S.A.F.E. program influenced the self-reported knowledge, attitudes, and behaviors of high school students; (ii) to confidentially administer the questionnaire to a group of high school students approximately one week before and one month after participation in the S.A.F.E. program; and (iii) to analyze pre- and post-intervention questionnaires and determine whether

there was a relationship between participation in the S.A.F.E. program and estimated changes in self-reported knowledge, attitudes, and behaviors in this group of high school students. The hypothesis was that, when controlling for grade, gender, and race/ethnicity, students who participated in the S.A.F.E. program would report improvements in knowledge, attitudes, and behaviors related to risky driving practices such as non-use of seat belts and driving under the influence of alcohol/drugs. The purpose of this short-term impact study was to validate the effectiveness of the S.A.F.E. program and to obtain preliminary data that could be used as the basis for the design of a larger, long-term, multi-site study.

Methods:

This short-term impact evaluation employed a non-experimental pre-test vs. post-test design using a single intervention group. The questionnaire used in the evaluation was developed, pilot-tested, and revised with the assistance of a teen advisory group. Questions focused on demographics, knowledge, attitudes, and behavior; the post-test included an additional section dealing with the impact of program components. The evaluation was exempted from review by the University of Miami Institutional Review Board and approved by the Miami-Dade County Public Schools Office of Evaluation and Research.

The Executive Director of S.A.F.E. identified a magnet school that was willing to participate in the intervention and evaluation, and S.A.F.E.'s Director of Training/Program Development and members of the evaluation team met with school officials to arrange timelines and logistics for the intervention and evaluation. The school determined that 10th, 11th, and 12th grade students in two of the school's academies would be invited to participate. The school distributed parent/guardian consent forms, available in both English and Spanish, to the teachers of the participating classes, who were responsible for the distribution and collection of parent/guardian approval to participate in the educational intervention and evaluation. Students who obtained parent/guardian consent were eligible to fill out the questionnaire. Administration of the pre-intervention questionnaire began one week before the intervention. The post-program questionnaires were administered approximately one month after the intervention, and were distributed only to those students who had participated in the intervention. Additional days of pre- and post-intervention questionnaire administration were added in order to include and retain as many students as possible.

Questionnaires were grouped by class and matched using birth date, locker combination (if given), and demographic information. Frequency analyses were used to describe the demographic characteristics of the sample, and analysis of variance procedures were used to determine whether there were significant differences in the demographic characteristics of students who completed the study and those lost to follow-up. McNemar's chi-square tests and paired t-tests were used to analyze the pre- and post-test responses to each specific question. Overall knowledge, attitude, and behavior scores were calculated for each student who answered at least 75% of the questions, and paired t-tests were conducted to determine whether overall scores improved after the intervention. Paired t-tests were followed up using repeated measures analysis of variance procedures in order to determine whether changes in scores varied according to gender, grade, or race/ethnicity. All significance tests were conducted at the 0.05 level unless otherwise specified. Follow-up comparisons were conducted on any main or interaction effects significant at the 0.05 level, and any interaction effects significant at the 0.10 were graphed and

interpreted. Answers to open-ended questions and impact data were condensed and summarized. All statistical analyses were conducted using SPSS 10.0 for Windows.

Results:

The school invited 23 classes, with approximately 484 students, to participate in the intervention and evaluation. Of the 484 students invited to participate, 335 completed the pre-intervention questionnaire, for a response rate of 69%. A total of 284 questionnaires could be matched, for a retention rate of 84.8%. Fifty-one (15.2%) students were unavailable either on the day of the intervention or at the time of post-testing, and were therefore lost to follow-up. The population included 33.5% male and 66.5% female; 39.8% of the students were in the 10th grade, 34.2% in the 11th, and 26.1% in the 12th. A total of 29.6% of the students reported their race/ethnicity as African-American or Caribbean Native, 54.2% as Hispanic/Latino, 10.9% as White Non-Hispanic, and 5.3% as Other or Unspecified. Analyses determined that the only difference between students who completed the study and those lost to follow-up was that 10th grade students were more likely to complete the study than either 11th or 12th graders, possibly because there was a scheduled band trip on one of the dates of the S.A.F.E. presentation. Students who completed both questionnaires and those lost to follow-up did not differ according to gender or race/ethnicity. The 1/3 to 2/3 male to female ratio in the study population is generally representative of the student population in the two academies that were invited to participate.

Answers to each question in the knowledge section were analyzed using McNemar's chi-square test; significant improvements in the percentage of students answering the question correctly were found for all questions but three. Paired t-tests also found that the mean knowledge score significantly increased across all students ($p < 0.001$), with a calculated effect size of $d = 1.17$. Analysis of variance procedures were used to determine whether knowledge changes varied by grade, gender, or race/ethnicity. The time main effect ($p < 0.001$) and the time by grade interaction effect ($p = 0.019$) were significant at the 0.05 level, and the time by grade by race/ethnicity interaction effect was significant at the 0.10 level. Follow-up tests conducted on the time by grade interaction found that the knowledge scores of 10th and 11th graders improved more than that of 12th graders.

As in the knowledge section, each individual question in the attitude section was analyzed using McNemar's chi-square test. The percentage of students reporting positive attitudes improved after the intervention for all questions but four. The change in overall attitude scores was analyzed using paired t-tests, and again found a significant improvement in reported attitudes ($p < 0.001$), with an effect size of $d = 0.55$. Analysis of variance procedures found a significant time main effect ($p < 0.001$) and a significant time by grade interaction effect ($p = 0.002$). The time by gender and time by race/ethnicity interactions were significant at the 0.10 level. Follow-up tests on the time by grade interaction found that changes in reported attitudes differed significantly between 10th and 12th graders ($p < 0.001$) and between 11th and 12th graders ($p = 0.022$).

Answers to individual questions in the behavior section were analyzed using paired t-test and McNemar's chi-square tests. Students reported improved behaviors after the intervention for all questions but one. Post-test behavior scores across all students were significantly higher than pre-test behavior scores ($p < 0.001$), with an effect size of $d = 0.47$. Analysis of variance procedures found a significant time main effect ($p = 0.004$) and a significant time by grade

interaction effect ($p=0.004$). The time by grade by race/ethnicity interaction was significant at the 0.10 level. Follow-up tests on the time by grade interaction found that changes in reported behaviors were significantly different between 11th and 12th graders ($p=0.047$) but not between 10th and 12th or 10th and 11th graders.

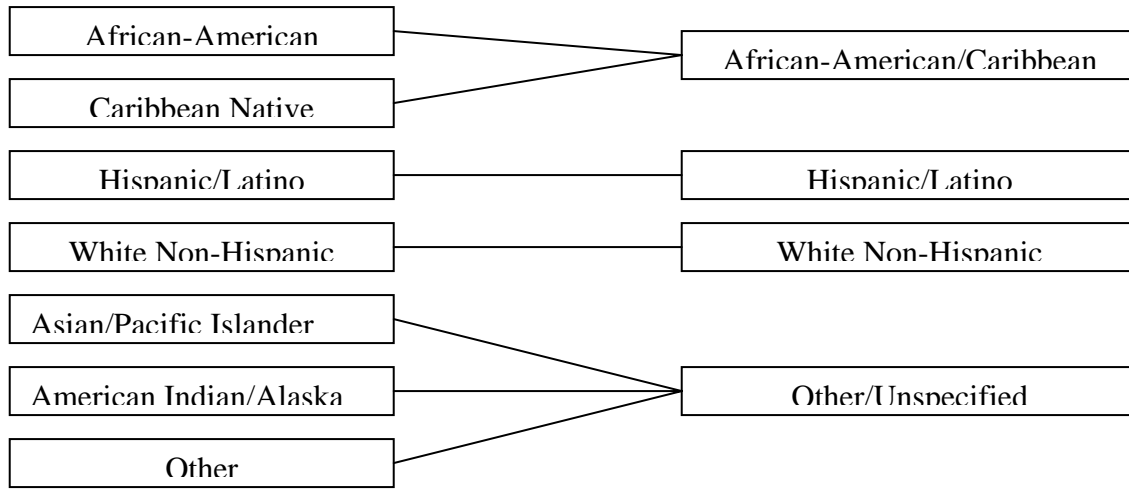
Conclusions:

Overall analyses found significant improvements in knowledge, attitudes, and behaviors (all $p<0.001$) between the pre- and post-tests across all students, but analysis of variance procedures showed that these improvements were not uniform across all grade, gender, and race/ethnicity groups. The program had a large effect in the short term on knowledge ($d=1.17$), but a more moderate effect on attitudes ($d=0.55$) and reported behaviors ($d=0.47$). It seems that the knowledge, attitudes, and reported behaviors of 12th grade participants improved less than those of 10th and 11th grade students; this may support an argument to target the program more towards younger high school students, as they seem to gain more from participation. The evaluation was subject to some limitations, including selection bias, questions about the reliability of self-reported behaviors, and failure to assess the long-term effects of the program. It would also be difficult to generalize the results of this study to the overall Miami-Dade high school population because the school in which this study was conducted was a magnet school.

The short-term impact evaluation did achieve the objectives in validating the positive impact of the S.A.F.E. program and collecting preliminary data to determine whether the S.A.F.E. program resulted in improvements in reported knowledge, attitudes, and behaviors among the participating high school students. Further investigation into the longer-term impacts of this intervention is warranted. This analysis presents the baseline data for a larger, long-term, more comprehensive evaluation of this educational intervention. Results will also be used to feed back into program design so that the effectiveness of the program can be improved.

For further communication please contact
Prof. Sherri Porcelain, University of Miami
sporcela@med.miami.edu
305.284.3128

Reclassification of Race/Ethnicity Categories



Notes on Study Sample

Total classes selected by school to participate: 23 (approximately 484 students)

Total students completing pre-intervention questionnaire: 335 (2 duplicates)

Response rate: 69%

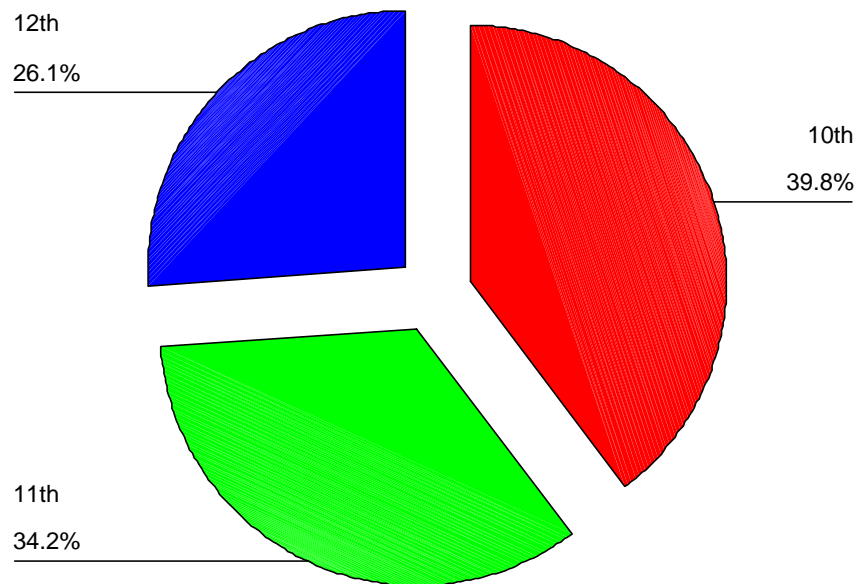
Students lost to follow-up: 51 (15.2%)

Total matched questionnaires: 284

Retention rate: 84.8%

Grade Level of Matched Respondents

Grade	Frequency	Percentage
10 th	113	39.8
11 th	97	34.2
12 th	74	26.1

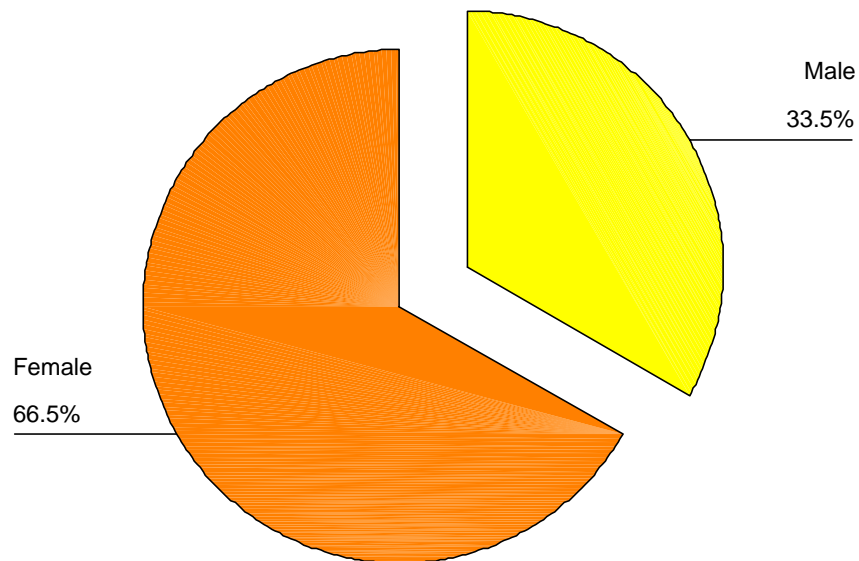


Notes:

- Statistical analyses found that the 10th and 11th grade students who completed the pre-intervention questionnaire were significantly more likely to complete the post-intervention questionnaire also.
- There was no significant difference in completion rates between 11th and 12th graders.

Gender of Matched Respondents

Gender	Frequency	Percentage
Male	95	33.5
Female	189	66.5

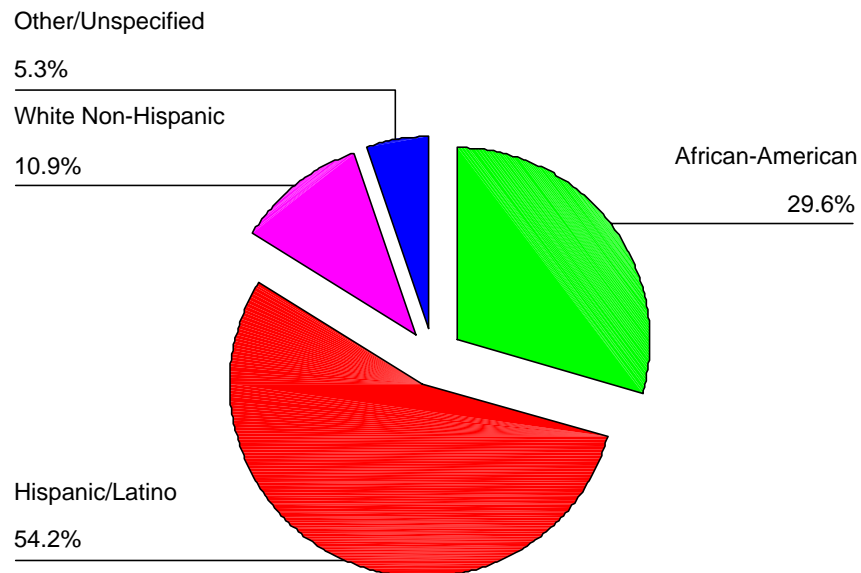


Notes:

- Statistical analyses found that there was no significant difference between male and female study completion rates.
- Statistics obtained from school officials showed that the gender distribution of the sample was representative of the student population in the two participating academies.

Race/Ethnicity of Matched Respondents

Race/Ethnicity	Frequency	Percentage
African-American/Caribbean Native	84	29.6
Hispanic/Latino	154	54.2
White Non-Hispanic	31	10.9
Other/Unspecified	15	5.3



Notes:

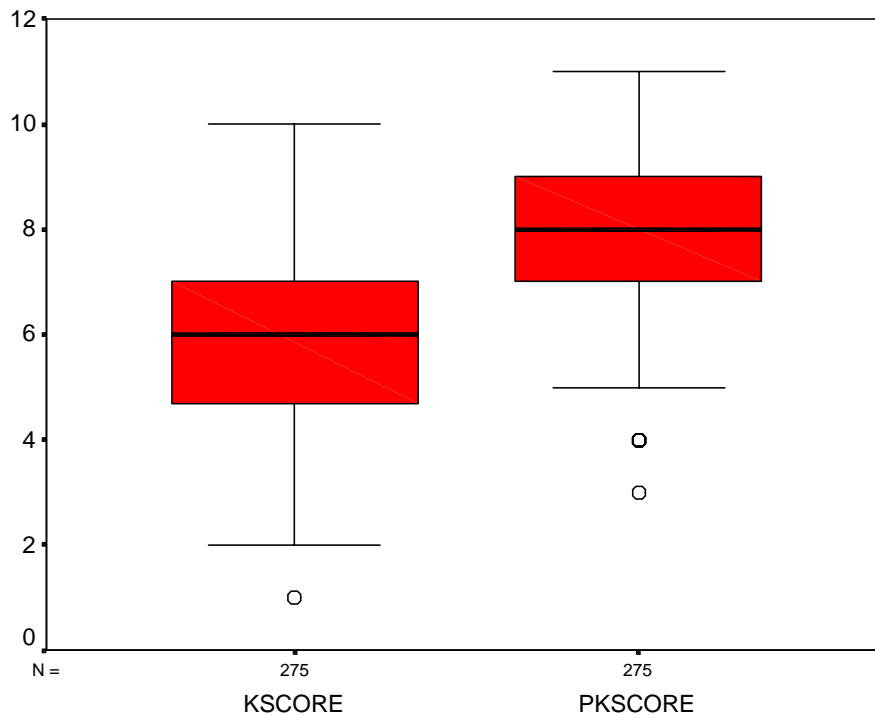
- Completion rates did not differ significantly according to race/ethnicity.

Knowledge

Question-by-Question Analyses

McNemar's Chi-Square Analyses					
Question Number	Valid N	Pre-Test Correct Responses	Post-Test Correct Responses	Change	P-Value
1	282	67.7%	94.3%	+ 26.6%	0.000
2	278	83.5%	92.4%	+ 8.9%	0.001
3	281	68.3%	86.5%	+ 18.2%	0.000
4	280	53.6%	75.7%	+ 22.1%	0.000
5	277	97.8%	98.6%	+ 0.8%	0.754
6	284	26.1%	44.7%	+ 18.6%	0.000
7	284	32.0%	54.6%	+ 22.6%	0.000
8	283	82.7%	93.3%	+ 10.6%	0.000
9	282	54.3%	89.0%	+ 34.7%	0.000
10	281	68.7%	77.6%	+ 8.9%	0.005
11	281	96.1%	94.0%	- 2.1%	0.263
12	277	21.7%	24.9%	+ 3.2%	0.391
13	277	5.8%	68.6%	+ 62.8%	0.000

Pre- vs. Post- Knowledge Scores



Notes:

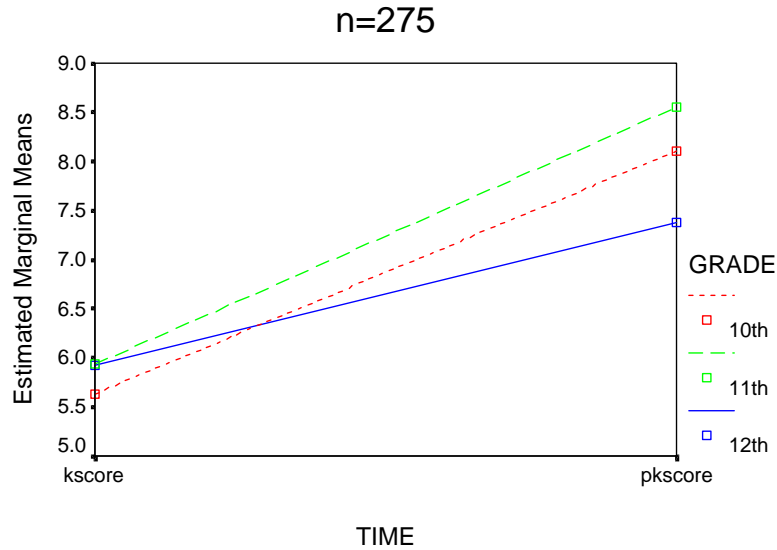
- Questions 5 and 11 were removed from the analyses and statistics recalculated because more than 90% of students answered correctly on the pre- and post-tests.
- Knowledge scores recalculated based on answering 10 of 11 questions (90%).

- Mean score increase from 5.65 to 8.07, a significant increase.
- When these two questions were removed from the analysis, effect size increased from 1.17 to 1.23.

Knowledge

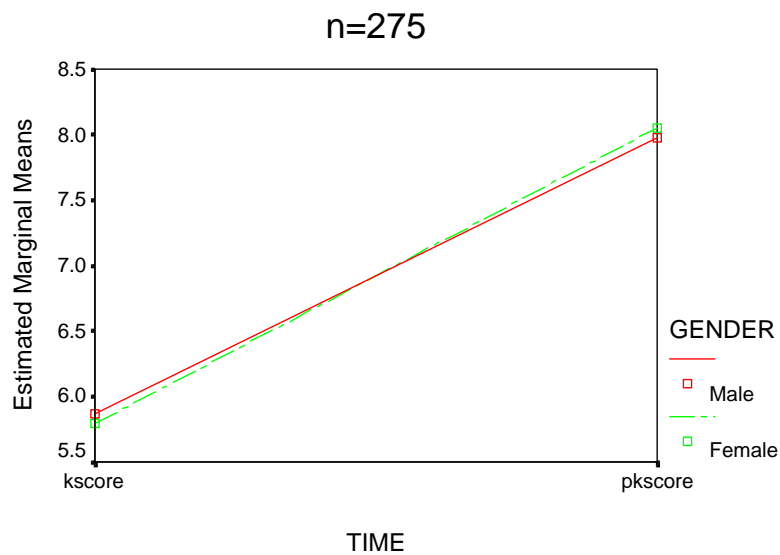
Interaction of Time and Grade

Estimated Marginal Means of KNOWSCOR



Interaction of Time and Gender

Estimated Marginal Means of KNOWSCOR

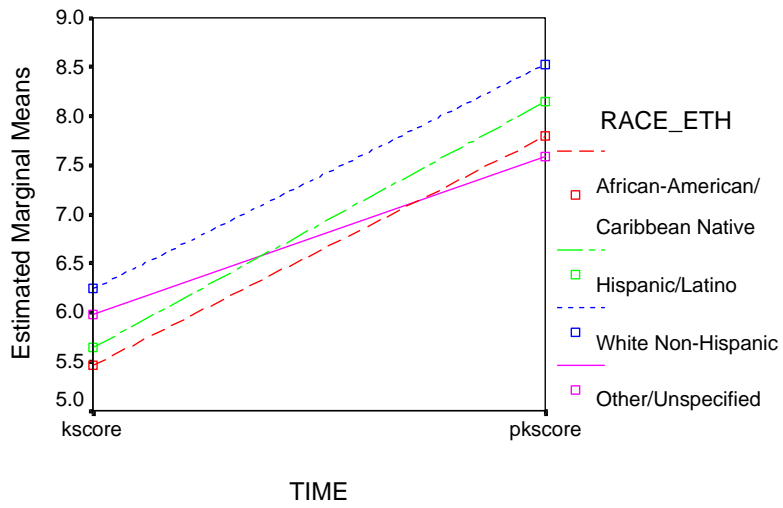


Knowledge

Interaction of Time and Race/Ethnicity

Estimated Marginal Means of KNOWSCOR

n=275

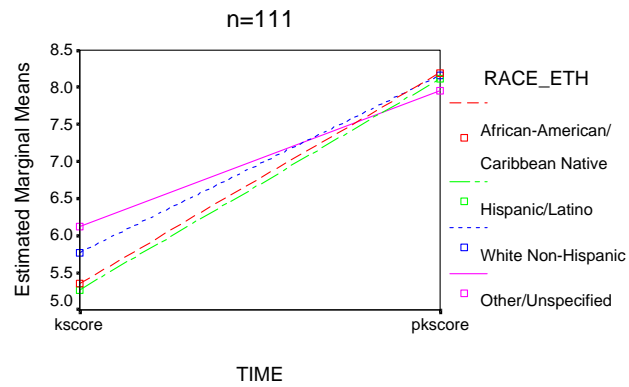


Knowledge

Interaction of Time, Grade, and Race/Ethnicity

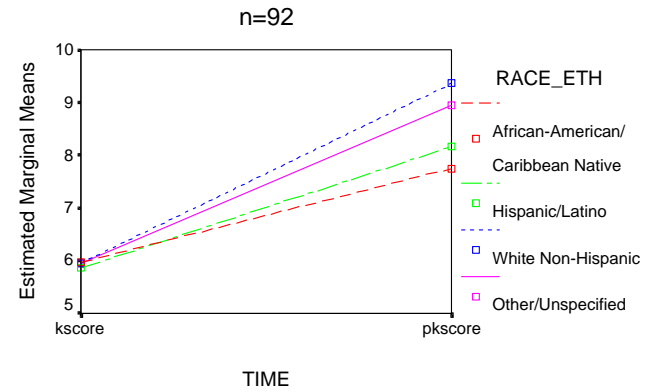
Estimated Marginal Means of KNOWSCOR

At GRADE = 10th



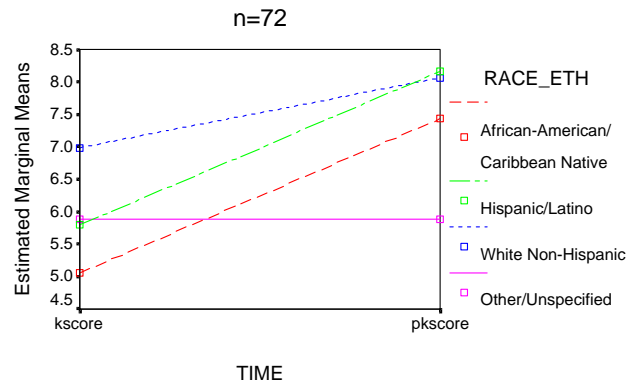
Estimated Marginal Means of KNOWSCOR

At GRADE = 11th



Estimated Marginal Means of KNOWSCOR

At GRADE = 12th



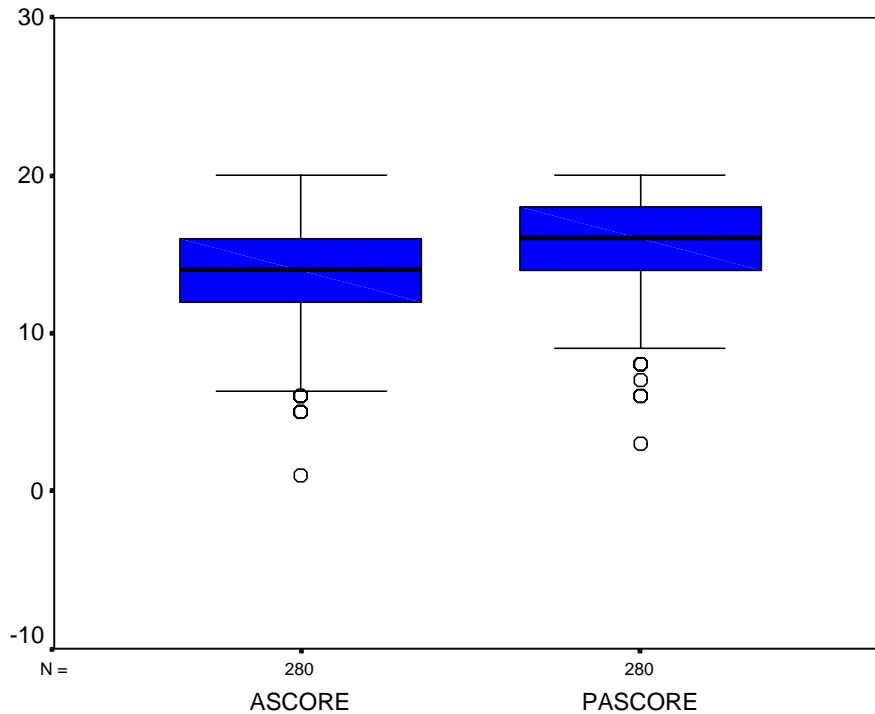
Attitudes

Question-by-Question Analyses

McNemar's Chi-Square Analyses					
Question Number	Valid N	Pre-Test Positive Responses	Post-Test Positive Responses	Change	P-Value
1	282	77.7%	91.1%	+ 13.4%	0.000
2	283	33.9%	60.4%	+ 26.5%	0.000
3	282	96.5%	93.3%	- 3.2%	0.078
4	280	75.1%	81.1%	+ 6.0%	0.046
5	280	92.1%	92.9%	+ 0.8%	0.832
6	281	53.0%	72.2%	+ 19.2%	0.000
7	280	47.1%	55.7%	+ 8.6%	0.013
8	281	38.8%	51.2%	+ 12.4%	0.000
9	281	81.1%	91.5%	+ 10.4%	0.000
10	281	63.7%	73.0%	+ 9.3%	0.002
11	281	79.0%	86.1%	+ 7.1%	0.013
12	281	60.5%	58.7%	- 1.8%	0.685
13	281	87.2%	89.0%	+ 1.8%	0.568
14	281	93.6%	93.2%	- 0.4%	1.000
15	283	97.5%	95.1%	- 2.4%	0.167
16	258	90.3%	98.1%	+ 7.8%	0.000
17A	278	80.6%	87.4%	+ 6.8%	0.005
17B	281	97.5%	98.6%	+ 1.1%	0.453
17C	279	81.7%	88.2%	+ 6.5%	0.008
17D	280	90.7%	91.1%	+ 0.4%	1.000
17E	280	97.1%	97.5%	+ 0.4%	1.000
17F	279	85.3%	91.8%	+ 6.5%	0.004
17G	275	53.1%	66.9%	+ 13.8%	0.000
17H	277	87.7%	92.1%	+ 4.4%	0.058
17I	278	58.6%	73.0%	+ 14.4%	0.000
17J	281	89.0%	92.2%	+ 3.2%	0.150
17K	280	83.2%	91.1%	+ 7.9%	0.001
17L	279	97.8%	99.3%	+ 1.5%	0.289
17M	276	47.5%	61.6%	+ 14.1%	0.000

Attitudes

Pre- vs. Post- Attitude Scores



Notes:

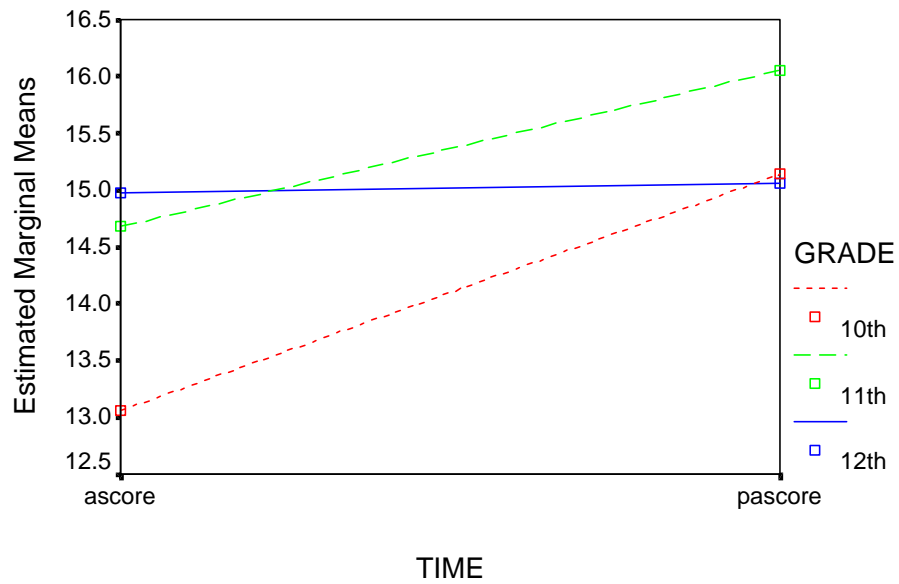
- Questions 3, 5, 14, 15, 16, 17B, 17D, 17E, and 17L were removed from the analyses and statistics recalculated because more than 90% of students answered correctly on the pre- and post-tests.
- Attitude scores recalculated based on answering 18 of 20 questions (90%).
- Mean score increase from 13.64 to 15.55, a significant increase.
- When these two questions were removed from the analysis, effect size increased from 0.55 to 0.62.

Attitudes

Interaction of Time and Grade

Estimated Marginal Means of ATTSCORE

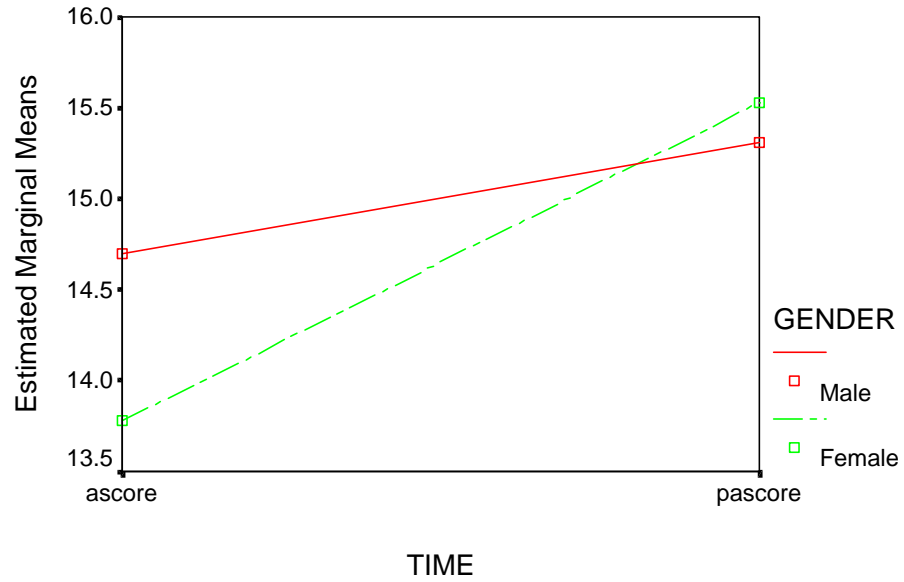
n=280



Interaction of Time and Gender

Estimated Marginal Means of ATTSCORE

n=280

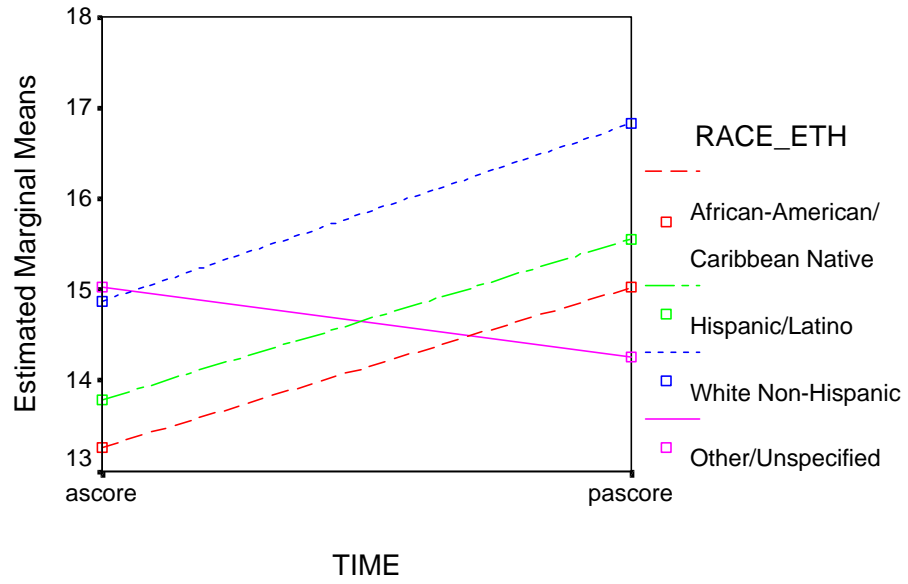


Attitudes

Interaction of Time and Race/Ethnicity

Estimated Marginal Means of ATTSCORE

n=280

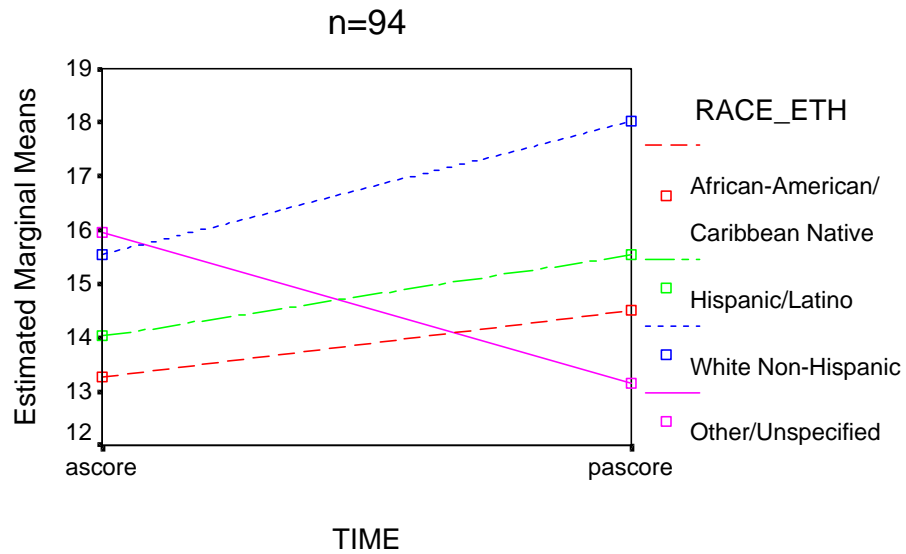


Attitudes

Interaction of Time, Gender, and Race/Ethnicity

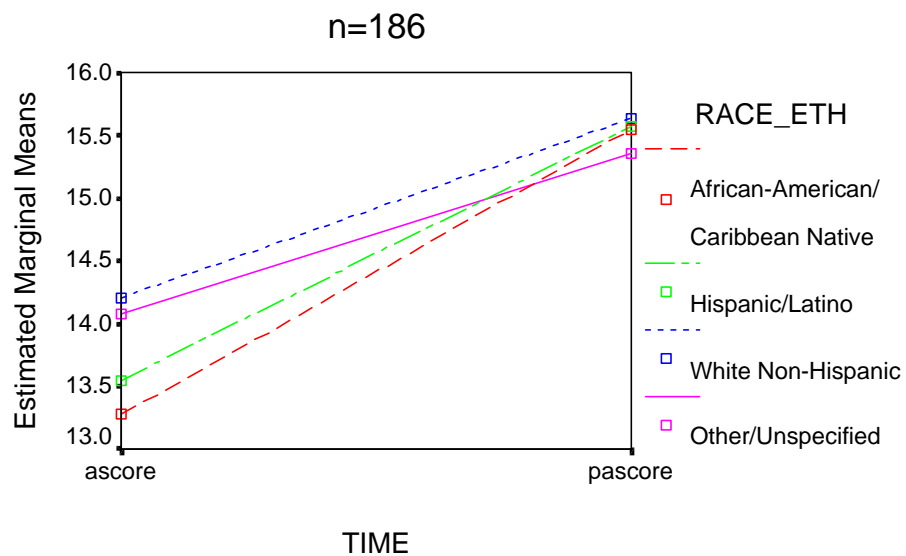
Estimated Marginal Means of ATTSCORE

At GENDER = Male



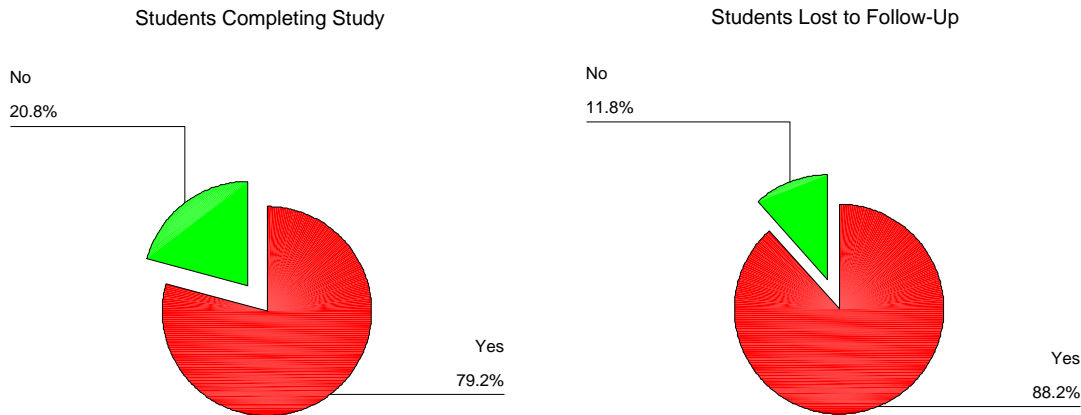
Estimated Marginal Means of ATTSCORE

At GENDER = Female



Behavior

Student Driving Status

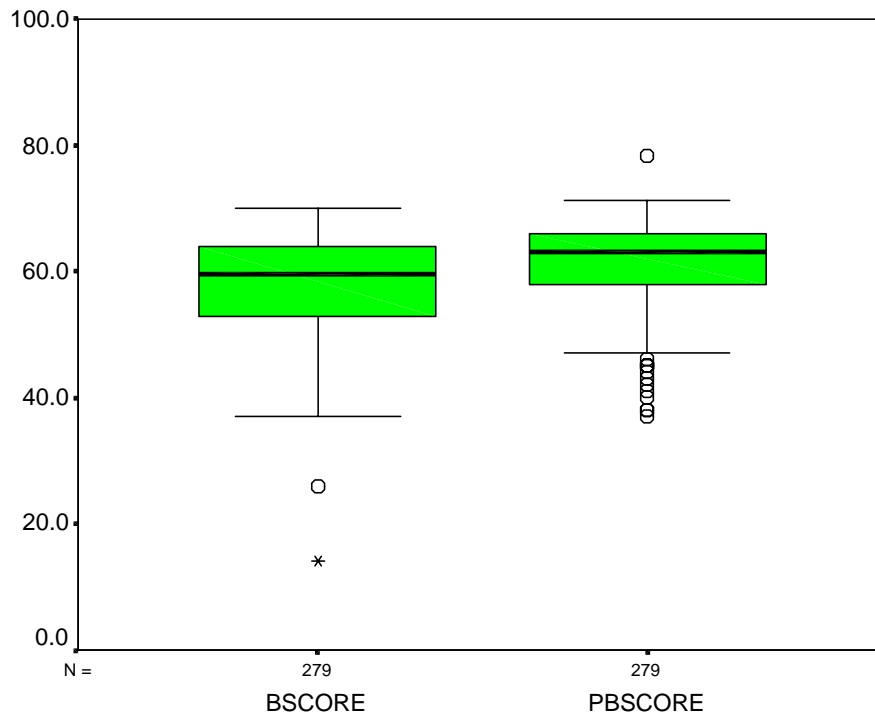


Question-by-Question Analyses

Paired T-Test Analyses					
Question Number	Valid N	Pre-Test Mean \pm Std. Error	Post-Test Mean \pm Std. Error	Change in Mean	P-Value
2	281	4.65 \pm 0.0413	4.83 \pm 0.0291	+ 0.18	0.000
3	280	3.59 \pm 0.0759	4.26 \pm 0.0567	+ 0.67	0.000
4	281	4.47 \pm 0.0522	4.70 \pm 0.0369	+ 0.23	0.000
5	280	4.47 \pm 0.0474	4.74 \pm 0.0323	+ 0.27	0.000
6	276	2.49 \pm 0.0632	2.96 \pm 0.0694	+ 0.47	0.000
7	279	4.25 \pm 0.0560	4.28 \pm 0.0602	+ 0.03	0.621
8	280	4.61 \pm 0.0450	4.63 \pm 0.0443	+ 0.02	0.832
9	278	4.72 \pm 0.0430	4.74 \pm 0.0403	+ 0.02	0.591
10	279	4.66 \pm 0.0478	4.70 \pm 0.0423	+ 0.04	0.390
11	278	4.23 \pm 0.0594	4.47 \pm 0.0524	+ 0.24	0.000
12	278	4.36 \pm 0.0572	4.48 \pm 0.0571	+ 0.12	0.127
13	281	4.34 \pm 0.0595	4.66 \pm 0.0426	+ 0.32	0.000
14	281	3.98 \pm 0.066	4.51 \pm 0.0519	+ 0.53	0.000
McNemar's Chi-Square Analyses					
Question Number	Valid N	Pre-Test Positive Responses	Post-Test Positive Responses	Change	P-Value
15A	278	75.9%	79.9%	+ 4.0%	0.169
15B	277	76.2%	72.2%	- 4.0%	0.161
15C	272	41.9%	47.8%	+ 5.9%	0.052
15D	274	86.5%	89.4%	+ 2.9%	0.215

Behavior

Pre- vs. Post- Behavior Scores



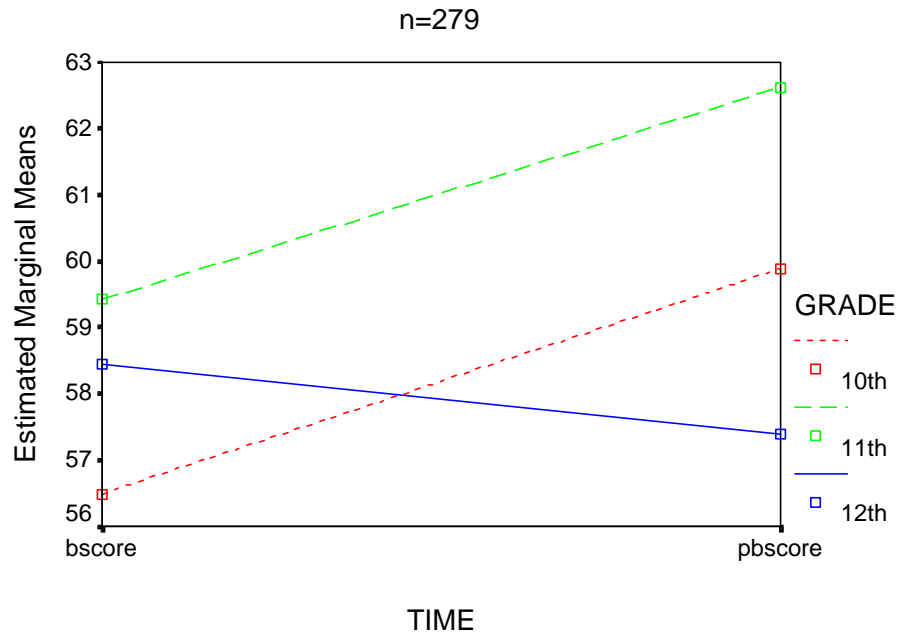
Notes:

- No questions were removed from analysis because it was felt that any change in behavior, no matter how slight, was important.
- Mean score increase from 57.84 to 60.97, a significant increase.
- Effect size of 0.47.

Behavior

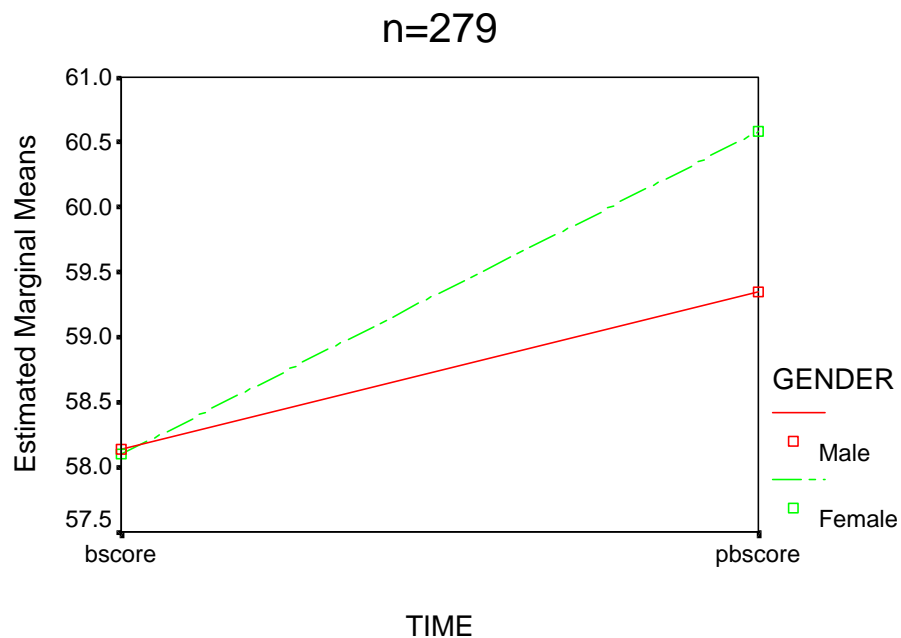
Interaction of Time and Grade

Estimated Marginal Means of BEHSCORE



Interaction of Time and Gender

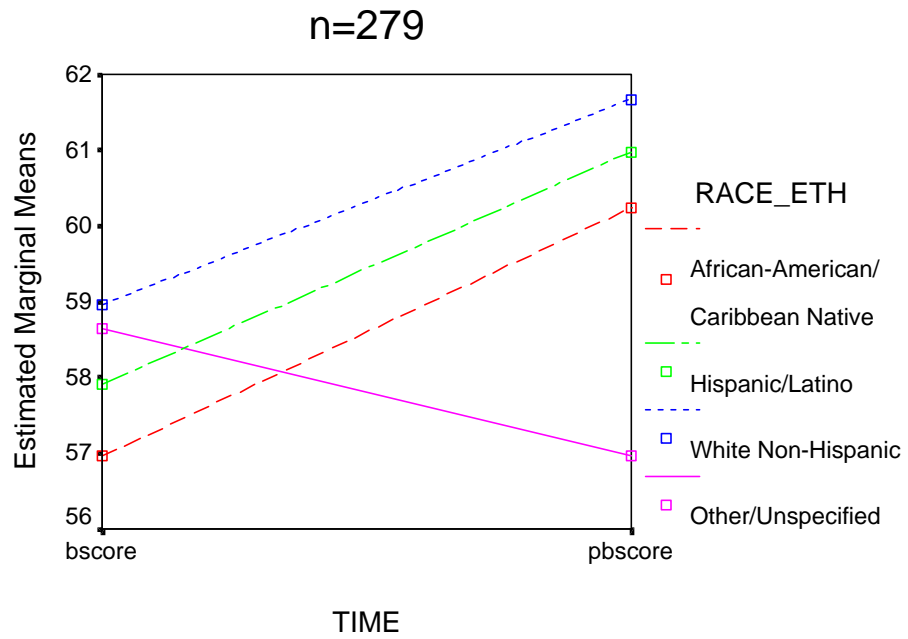
Estimated Marginal Means of BEHSCORE



Behavior

Interaction of Time and Race/Ethnicity

Estimated Marginal Means of BEHSCORE

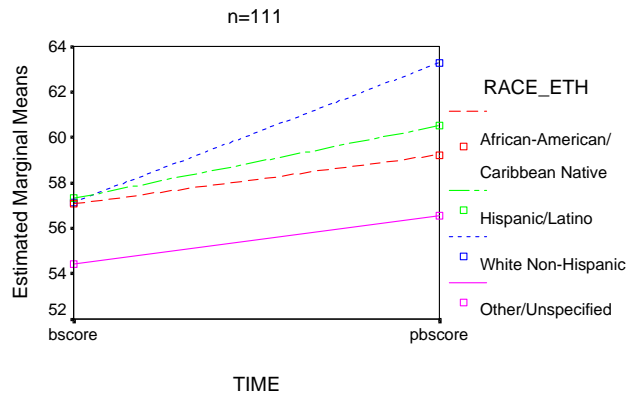


Behavior

Interaction of Time, Grade, and Race/Ethnicity

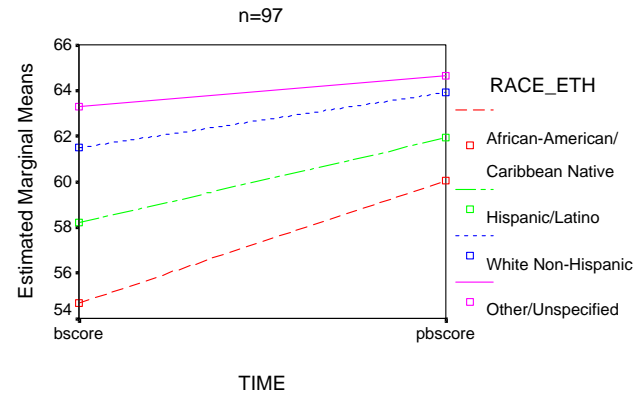
Estimated Marginal Means of BEHSCORE

At GRADE = 10th



Estimated Marginal Means of BEHSCORE

At GRADE = 11th



Estimated Marginal Means of BEHSCORE

At GRADE = 12th

