University of Illinois at Chicago
Department of Mathematics, Statistics, and Computer Science

Statistical Analysis of the Emerging Scholars Program at the University of Illinois at Chicago

Jaime Brugueras
Luissette Hernández-González
Prof. Anatoly Libgober

ABSTRACT. The present study investigates the effectiveness of the Emerging Scholars Program (ESP) at the University of Illinois at Chicago (UIC). This program consists of workshops running parallel to the basic courses in mathematics and physics. These mathematics and physics courses are part of the requirements for most students majoring in science, mathematics, and engineering, as well as in many other disciplines. One of the main goals since the very start of the Emerging Scholars Program has been to increase the success rate of minority students participating in the introductory courses in mathematics and physics. A firm conclusion of this study is a significant positive correlation between participation in this program and the achievements of the students in these basic courses. This paper starts with the description of the Emerging Scholars Program at UIC. Then, it presents details of the background of its inspiration, the Uri Treisman's Model. Following, it presents the basic questions that we pose as the goal for our investigation as well as the methods that we used. Next it reports statistical data comparing the performance of ESP and non-ESP students, together with a brief discussion of the results.

1. Introduction

In the fall of 1989, the Department of Mathematics, Statistics, and Computer Science (MSCS) at the University of Illinois at Chicago (UIC) began an experimental program designed to improve the performance of underrepresented minority students in calculus and pre-calculus courses. As in many other Universities, data on UIC students confirmed that passing calculus was a major stumbling block for African-American and Hispanic students in entering into mathematics, science and engineering majors. For many UIC minority students with aspirations for a mathematics or science degree, the obstacle came even earlier, in pre-calculus courses. The staggering percentage of failing minority students exceeded 55% in pre-calculus courses.

In an attempt to address this problem, UIC looked to a model developed by Uri Treisman during the late 1970's at the University of California, Berkeley. Treisman's model was based on his groundbreaking study of differences between how African-American and Asian students approached their work in Calculus classes. He investigated the reasons behind the high failure rates of Black and Hispanic students in calculus courses. He started his research on the basis of four hypotheses for these high failure rates, which were socially accepted at the time. These hypotheses included lack of motivation, lack of academic preparation, lack of family support, and low income. The result of his study not only disproved these hypotheses but also shifted the blame towards another direction, the institutions. Studying twenty African-American students' routines he found that the hypotheses were wrong. After interviewing the African-American students he reported that these students have been intensely motivated since high school. The second hypothesis, lack of academic preparation, was overthrown when Treisman and his colleagues saw the paradox of African-American students with high SAT scores failing the calculus courses. High SAT scores are usually linked to good academic preparation. The third hypothesis, lack of family support, was refuted when Treisman found that the parents of these students were extremely supportive and eager for their children to go to college since "many of the parents had decided before their children were even born that their sons and daughters would go to college" (Treisman, 1992, p. 5). Also, the fact that many of the parents of these African-American students had low-income jobs did not seem to be a significant factor that would contribute to the students' failure on the mathematic courses.

Given that the original hypotheses turned out to be false, he looked more closely at the African-American students' studying habits. He found that they followed the studying routine instructed to them by the institutions studying skills courses. Treisman reported that they went to class, took notes and studied by themselves about six to eight hours a week. They did their homework and turned it in on time. Nevertheless, they were failing (Treisman, 1992). Thus, the institution was offering studying skill courses that turned out to be inefficient.

To find some answers he looked at the studying habits of the Chinese students. These studying habits turned out to be more like a lifestyle. They went to class, took notes and studied by themselves about eight to ten hours a week. Then came what is the real difference between the two groups. Treisman reported that after working individually on their homework, the Chinese students got together, ate, talked and corrected each other's work. They also worked on old exams available to the public and also had their older family members test them. Treisman also reported that they had a system which included determining which problems were designed to be easily solvable by the students and which were design to be extremely challenging, thus minimizing frustration. Treisman also found that each Chinese students knew what was their position in the class and what grade they were expecting to obtain in a particular exam. "They [Chinese students] had constructed something like a truly academic fraternity" (Treisman, 1992, p. 5).

Treisman then developed a program that modeled the successful practices of the Chinese students. The main components of Treisman's method was encouraging free discussions of the material in the courses between the students, beyond the discussion initiated by the instructor and developing the students ability to initiate such discussions as well as critical thinking about mathematical problems. The workshops were conducted by the mathematics department and decidedly were not remedial in nature. "They [African-American students] see themselves as the tutors, not the tutees" (Treisman, 1992, p. 6).

The worksheets were design to be difficult and the workshops' reputation of intensity appealed to the upcoming students from underrepresented groups. He wanted to provide a fertile ground for the African-American students to create their own academic communities and to have a comfortable environment in which they had the opportunity to share mathematical ideas. The results were very encouraging. Many students from underrepresented groups were performing as well as White and Asian students and these workshops were producing a high quantity of physicians, scientists, engineers and scholars.

2. Emerging Scholars Program at the University of Illinois at Chicago

UIC used Treisman's model of intensive workshops to implement the Emerging Scholars Program. One of the goals of these workshops is to give the students of underrepresented groups a multi-ethnic platform to collaborate with students who possess their same academic interests. ESP encourages students to communicate and share ideas about mathematical problems. This program also encourages students to get together through academic and social activities. Besides the social aspects of the ESP workshops, there is also the intrinsic academic value. At these workshops it is expected of the students to come prepared by having worked on the assigned reading and homework. In this way they are ready to tackle more challenging problems in a group environment. With these workshops the standard of academic achievement has been raised, so the students know that more is expected from them than just passing the course, they are expected to excel.

2.1 Workshop Management

The ESP workshops count with the participation of a program coordinator, faculty members, workshop instructors (teaching assistants), and undergraduate assistants. The program coordinator organizes the workshops schedules and teaching assignments. He or she also administers a general meeting with the workshop instructors prior to the beginning of each semester to emphasize the goals of the program and has additional meetings throughout the semester on an individual basis.

Faculty members get involved with the academic aspect of the workshop. They frequently collaborate with the workshop instructors during the development process of the ESP worksheets. On occasions faculty members visit the workshops to observe the dynamic between the workshop instructors and the students. Also, they encourage student participation in the ESP workshops by describing the program in their lectures.

The ESP workshops are conducted by an instructor, who is a teaching assistant in a graduate program at the Department of Mathematics, Statistics, and Computer Science at UIC. The instructor is constantly coordinating the workshop to promote an environment of work and communication. If needed, undergraduate assistants are also recruited to aid the instructors. During an ESP workshop, groups of students are formed and encouraged to try to solve problems generally more difficult than those covered in their related math courses. One of the purposes of this approach is to help the students develop problem-solving skills in a team-oriented environment. Students are encouraged to listen and to help one another while trying to solve problems presented to them in a worksheet. Following Treisman's model, the worksheet aims to challenge the students by making them use ingenuity to solve its problems. Students are also encouraged to meet in study groups outside the workshop to prepare for examinations for their related courses. This promotes the development of friendships and contacts that could endure for the rest of the students collage life and career. Having a network of peer academic support is particularly important at an urban institution such as UIC where many students commute to and from their homes everyday and could frequently underestimate the need for student interaction.

2.2 Recruiting

The way in which students are recruited for workshop participation has changed throughout the fifteen-year period that the Emerging Scholars Program has existed in the University of Illinois at Chicago. The first recruiting effort at UIC took place during the Spring of 1989, with letters inviting 200 recipients of the President's Award Program scholarships and to approximately 50 students who obtained A or B in Pre-Calculus Algebra or Trigonometry to a Professional Development Program (PDP) [1] orientation. After the success of the PDP during the academic year of 1989-1990, the recruitment process for the following years has expanded to include announcements in class, occasional telephone calls, and emails. Prior to the beginning of each academic semester, the program coordinator sends a letter to every student enrolled in the courses for which ESP workshops are offered. This letter describes the workshops to the students and extends them an invitation to enroll. In addition, during the first day of instruction, the ESP coordinator announces the existence of these workshops at the classrooms where the courses related to the workshops are offered, to the benefit of those students who have just enrolled. At that moment, a description of the workshop is offered. Faculty members also make class announcements about the ESP workshops. In previous years, the program coordinator has contacted by telephone students that are considered to benefit the most from these workshops according to the schedule, but this is no longer practiced. Students are also notified about the workshops by email. Also, the directors of several programs that work with the academic aspects of undergraduate students are informed about the workshops through letters sent by the ESP coordinator.

Instructors highly dedicated are preferred to conduct the workshops, consequently the ESP coordinator tries to recruit teaching assistants recommended by the assistant director of graduate studies, who is in constant contact with graduate students. Also, recommendations from other teaching assistants that have participated in the program are obtained. When all instructing positions have not been filled, announcements are made about these positions. When possible candidates are identified, they are contacted by email to meet with the coordinator. At the meeting, the program is explained in detail to the T.A. including what is expected of him or her. Many teaching assistants that have participated in the program are so fulfilled by the experience that they keep working in subsequent ESP workshops. Nevertheless, due to the fact that teaching assistants are constantly completing their degree and moving to new ventures, sometimes positions are not immediately filled. In these occasions, ESP workshop instructor positions are assigned to a new T.A. as part of their teaching appointment for the semester. These situation, although rare, is not preferred since special dedication is required to conduct these workshops. Generally, the teaching assistants involved in an ESP workshop possess enthusiasm, commitment, and leadership.

Undergraduate assistants (U.A.) are generally assigned to workshops with a large amount of students enrolled. The role of the undergraduate assistant is to aid the instructor in the guidance of the groups. Undergraduate assistants are recruited by the program coordinator on the basis of their major and their performance in math courses. Preferred majors include Mathematics and Mathematics Education, but a qualified,

---

[1] In 1989 the Emerging Scholars Program was called the Professional Development Program.

dedicated student from another major within the MSCS department is also welcome to participate in the Mathematics ESP workshops. To be a qualified student, he or she must have passed the course related to the workshop, preferably with an A or a B, but a student who obtained a C is only considered if grade improvement has been made in subsequent math courses. Also, students that have participated in ESP workshops are preferred for this position since they are already familiar with the workshop dynamic.

Since the workshops implementation at UIC, the ESP scope was substantially expanded by including upper level Calculus courses, Differential Equations and courses in Physics. ESP has since served over 2,500 students in mathematical courses.

3. Statistical Analysis

The intent of this segment is to study the effects of the Emerging Scholars Program (ESP) workshops by means of performance investigation. We tried to answer the following questions:

- How does the math and physics GPA and ACT scores compare between students that participated in more than one ESP workshop with those students that did not participated in any ESP workshop.
- How does performance in mathematics(Pre-Calculus, Calculus I, Calculus II, and Calculus III) and physics (Introductory Physics, General Physics, and Natural Sciences) courses from students who participated in the ESP workshops compare with performance of those students that did not participated in these ESP workshops.
- How does performance in subsequent mathematics and sciences from students who participated in physics ESP workshops compare with performance of those students that did not participated in these ESP workshops.
- How does performance in the given mathematics and physics courses from students who participated in the ESP workshops compare with performance of those students that did not participated in these ESP workshops when grouped by ethnicity.
- How does graduation rates compare between students who participated in mathematics or physics ESP workshops with those students that did not participated in these workshops.

The data used in this study contained no personal identifiable information like name or social security number. To differentiate one student from the other in this study, an identifier assigned by the university was used.

The whole population of students who took up to 300 level math courses from 1991 to 2002 was gathered as well as all of the students who took up to 300 level physics courses from 1994 to 2002. The variables found in the data were: student id, course number, course grade, freshman year, graduation date (if any), sex, ethnicity, and ACT scores. We are interested in finding a relationship in the performance of these two types of students (ESP and non-ESP).

To start our analysis, we compile grade information for the two groups, ESP and non-ESP students, for each of the math and physics courses related to the ESP workshops to be analyzed. Following, a comparison is made of the math and physics grade point average (GPA) of these students. An overall analysis of GPA is not an accurate assessment of performance, since a student who takes the ESP program at least once is considered an ESP student. Thus, scores from non-ESP classes may be used for an ESP student's math GPA. Parametric and nonparametric methods were used for making inferences on the null hypothesis, which is that the mean GPA of the ESP sample is the same as that of the non-ESP sample. Similar methods were performed to make inferences on the sample means of the ACT score. The results show that ESP students have a significantly higher mean ACT and math GPA than the non-ESP students. One can then speculate that students joining the emerging scholars program would perform better than non-ESP students, regardless of the program. To account for this partiality, matched sampling by propensity score (Rosenbaum & Rubin, 1983a) was performed in both math and physics courses. This method allows for selection of a sample from the non-ESP group that is as 'similar' as possible as the ESP group. Therefore, direct unbiased analyses can be performed to compare the groups in each course.

As in a previous study done at the University of Wisconsin (Kosciuk, 1997), we decided to consider success if a student earned an A or a B in the course and a failure if he or she received a C, D, E, or incomplete. For each course, a 2x2 contingency table is drawn from the matched data, and a chi-square test of independence measured the relationship between a student's grade and taking the related ESP course. A compilation of graduation rates data is also presented.

3.1 GPA Comparison

A statistical model employed in our problem assumes, under the null hypothesis, that the observations for the control group are independent random variables with normal common distribution, and the treatment group are independent of themselves and the controls. The observations in our analysis included the (math or physics) GPA of each (treatment and control) student. A change in the overall level of responses is easily detectable by using the mean of each group. Therefore, an analysis that focuses on the difference of means is ideal. When there are no differences in the GPA mean of ESP students vs. non-ESP students, differences as large or larger than the one observed occur with probability less than .05 (this means a p-value of less than .05). If this happens, then we can conclude that there is little doubt that the control and treatment groups are different.

If we apply a nonparametric method to the data, the assumption that the observations come from a normal distribution may be discarded. Instead, the methods are based on data by ranks.

For our comparison of samples of ESP and non-ESP GPA scores, we applied the well known Wilcoxon rank sum test. Basically, the test grabs the $n+m$ observations, ranks them (in increasing order of GPA in our case, 1 lowest rank), and sums the ranks for the control and the treatment groups. We rejected the null hypothesis when this value was too low or too high. Figure 1 and 2 show the t-test results for mathematics and physics.

Mathematics

	N	Mean Math GPA	Standard Error
ESP students	922	3.37	0.033
non-ESP students	18795	3.18	0.008
t-test	*t* = 5.23	p < .0001	

Figure 1.

Physics

	N	Mean Math GPA	Standard Error
ESP students	598	3.49	0.038
non-ESP students	7330	3.25	0.012
t-test	*t* = 5.60	p < .0001	

Figure 2.

Both parametric and nonparametric methods(which are not shown here) give similar results. The high values for the t-test for both mathematics and physics indicate that there is probably a difference between the control and treatment group. That is, in general, the students taking ESP courses have a different (we can say higher) GPA than the ones that have not taken an ESP course ever.

3.2 Student Performance Comparison (ESP vs. non-ESP)

Success for a student on a specific course is relative. For some students, passing a class is just as important as for somebody else to get an A. But in accordance with the literature, we restricted ourselves to defining success as earning an A or a B. (We got similar results when C was included as a successful grade). This leads to many other ways of analyzing this data.

In this section, we form 2x2 n sample size contingency tables of two variables: ESP(yes or no) and Success(yes or no).

Comparing ESP and nonESP students directly per class would give biased estimates of the treatment effect. That is, students that joined ESP could be either 'smarter', more responsible or more interested since they volunteered to join such course. As noted in the previous section, the students in ESP have a better math GPA. We also found very similar results for the ACT score; students that belonged to an ESP course were found to have a higher mean ACT score than nonESP. Measures have to be taken to reduce such bias.

3.3 Matched Pair Sampling

In this statistical study of the Emerging Scholars Program (ESP) we analyze whether students who take the complementary ESP course (treatment group) tend to receive better grades than those who are not part of the program (control group). Since this is an observational study, given that the students were not assigned to the treatment group randomly, matched pair sampling by propensity score (Rosenbaum & Rubin, 1983a, 1985) was used to reduce the bias of a nonrandom selection of students. In this case, the propensity score is the conditional probability of assigning a student to ESP given the matching covariates of ACT score, math GPA, sex, and ethnicity. The propensity score was calculated using a multivariate logistic model (Hosmer & Lemeshow, 1989) for each course. Due to a high number of missing values for the matching covariates, discarding these observations would result in a significant loss of information. Below is a table showing missing data patterns. A dot in a group implies that the data is missing a value for that variable. Notice that around 22% of the data is missing the ACT score.

a) Mathematics

A multiple imputation procedure (Rubin, 1987) was used to 'fill in' missing values with plausible values. To create the propensity scores we,

1. Created 5 complete data sets using SAS/STAT procedure, PROC MI (Yuan, 2000).
2. A propensity score was computed for each complete data set using SAS/STAT, PROC LOGISTIC.
3. Finally, SAS/STAT PROC MIANALYZE was used to generate, from the 5 data sets, a point-estimate of the propensity score and its standard error.

The point estimate of the propensity score is simply the average the 5 estimates computed from the complete data sets. i.e.

$$\overline{p} = \frac{\sum_{i=1}^{5} \hat{p}_i}{5}, \ i = 1,2,\ldots,5.$$

We found that the only significant covariates in the logistic procedures were the ACT score and the math GPA for all courses and additionally, ethnicity was significant for math 121 and 180. No interaction was significant and therefore these two simple models were implemented to compute the propensity score estimates for each course.

Matching by propensity score was implemented using a similar algorithm as the Greedy 5 -> 1 Digit Matching Algorithm (Parsons, 2001). In our case, we started the algorithm by matching the propensity scores searching for equality in the last 7 decimals. If all ESP students are matched with a nonESP student then the algorithm stops and saves the data. If there are unmatched ESP students, the propensity score is rounded to 6 decimal points and the matching algorithm continues. This is continued until the propensity score is rounded to the first decimal and all ESP students are uniquely matched. The next two tables show the frequency distribution of students before and after matching.

Number of Students per Class

Before Matching

	ESP	NonESP	Total
Pre-calculus	588	13756	14344 37.26%
Calculus	422	11568	11990 31.14%
Calculus II	316	6500	6816 17.70%
Calculus III	145	5203	5348 13.89%
Total	1471 3.82%	37027 96.18%	38498

Figure 3.

Number of Students per Class

After Matching

	ESP	NonESP	Total
Pre-calculus	588	588	1176 40.0%
Calculus	422	422	844 28.7%
Calculus II	316	316	632 21.5%
Calculus III	145	145	290 9.9%
Total	1471 50%	1471 50%	2942

Figure 4.

The treatment effect in this study is measured as a success, if the student attains an A or B, or a failure, if the student attains a C, D, E, IN, or W. The analysis of an observational study with a binary outcome using matched pair sampling by propensity score allows for the use of a chi-square test (Rosenbaum & Rubin, 1983b).

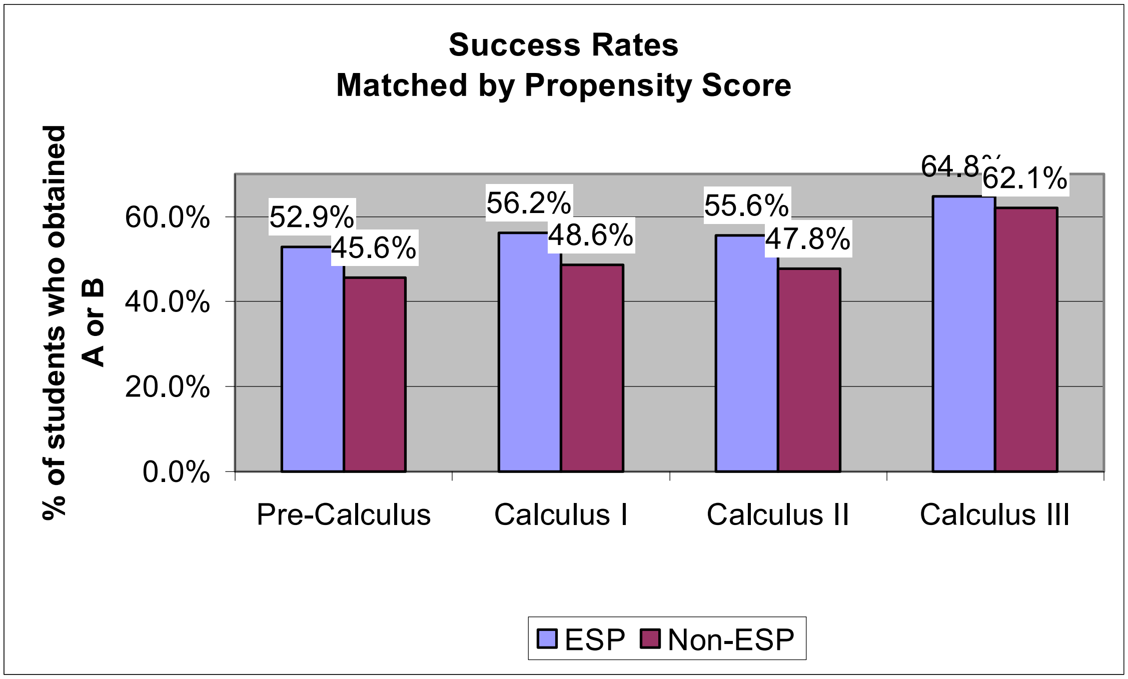

Figure 5.

Figure 5 shows that after a matched sample has been selected, the percentage of students obtaining an A or B is higher within the ESP students than the non-ESP students. In the table below the results for Fisher's chi-square test reveal that except in calculus III, at a 10% confidence level, one can say that students belonging to the ESP program have a better chance of obtaining an A or B in the respective math course. In calculus III, there is a slight higher percentage in the whole sample and in the matched sample, but there is no statistical evidence that shows this difference is due to the program.

Chi-Square Statistic and p-value

	Pre-Calculus	Calculus	Calculus II	Calculus III
Chi-Square	5.77	4.66	3.82	0.2
p-value	0.0163	0.031	0.05	0.651

Figure 7.

b) Physics

The same algorithm was used to build a matched pair sample for the physiscs courses.

As in the mathematics courses, multiple imputation was applied to the physics data because more than 30% of the data has a missing variable for one of the covariates. Figure 9 shows the frequency of ESP and non-ESP students in the physics courses before matching.

Number of Students per Class

Before Matching

	ESP	NonESP	Total
Introductory Physics	286	3710	3996
Natural Sciences	158	1520	1678
General Physics	331	4259	4590
Total	775 7.6%	9489 92.4%	10264

Figure 9.

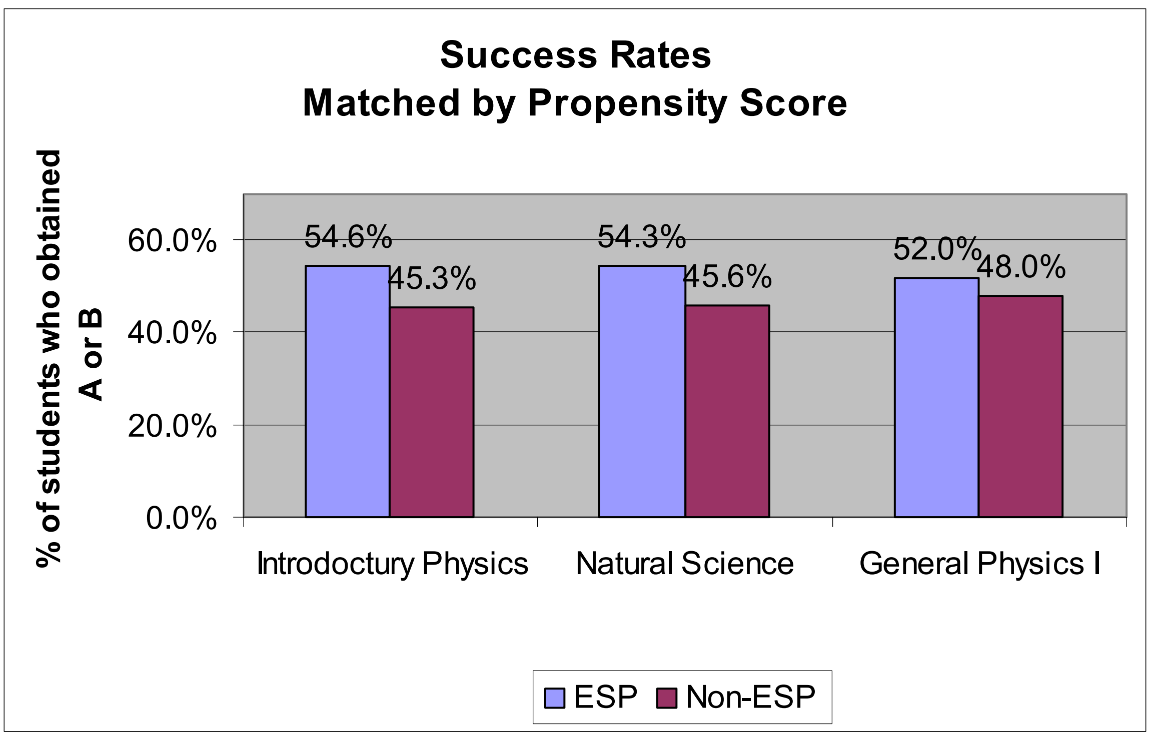

Figure 10.

Chi-Square Statistic and p-value

	Introductory Physics	Natural Science	General Physics
Chi-Square	1.88	4.88	0.602
p-value	0.10	0.027	0.44

Figure 11.

Figure 10 shows that ESP students outperformed non-ESP students. In Figure 11 the one-sided chi-square statistics and their respective p-values are shown. Except in General Physics, at a 10% confidence level, there exists statistical significance of the dependence of the performance of students to the ESP program. For General Physics, the null hypothesis is accepted, the students' grade and their participation in the program are two independent variables.

3.4 Subsequent Mathematics and Physics.

One of the ESP program's objectives is to become a better problem solver and therefore, it is expected that ESP students will do better in subsequent mathematics or physics courses. Our job was to verify this hypothesis. The data included advanced courses in mathematics where students who had ESP previously were present. As in the overall analysis, any student that took any ESP previously was considered under the ESP group. The courses chosen were Differential Equations for the mathematics students and General Physics II for the physics students. A similar chi-square analysis as in the previous section was used to interpret the results. The results are depicted in figures 12 and 13.

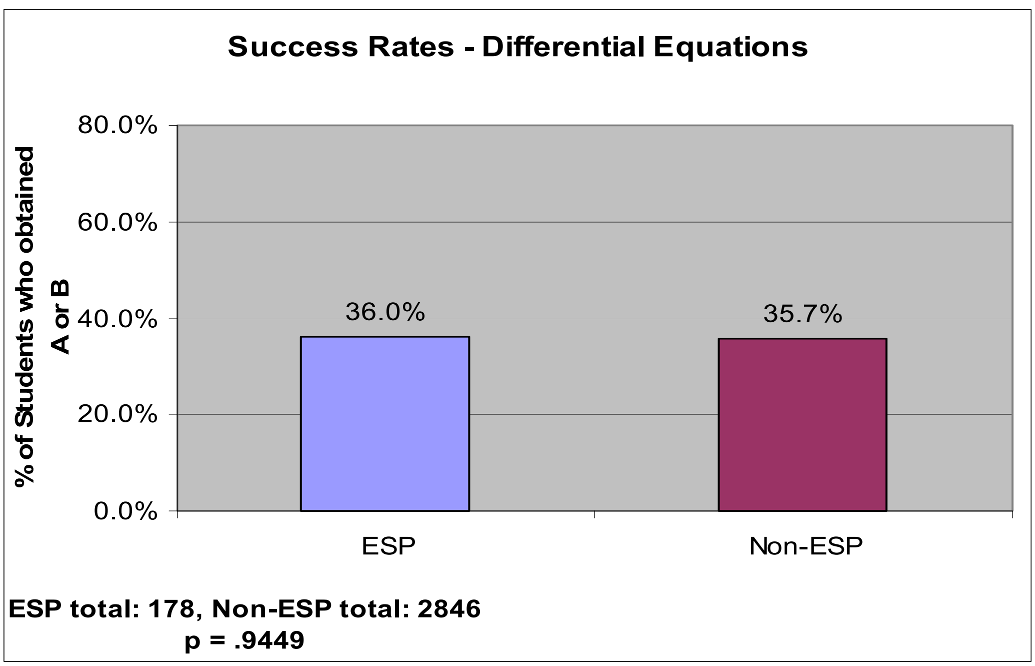

Figure 12.

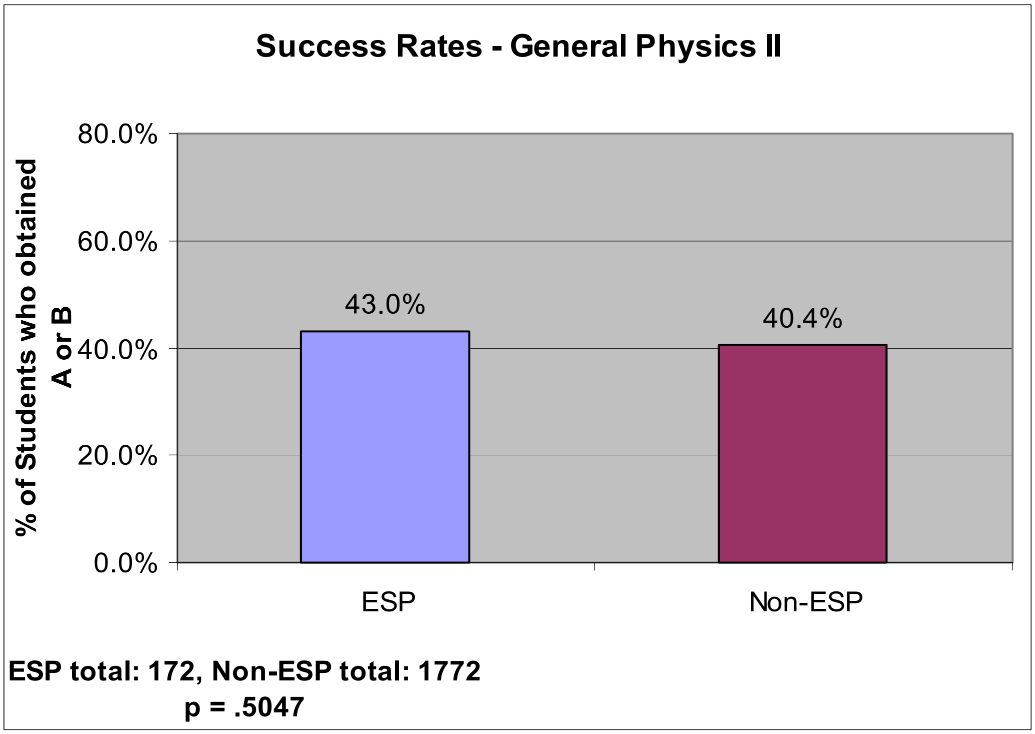

Figure 13.

Course	Chi-Square	p-value
Math 220 (Differential Eqns.)	.0048	0.9449
Physics142 (Physics II)	.4451	0.5047

Figure 14.

There is no statistical evidence of any relationship between success in subsequent mathematics or physics courses and pertaining to an ESP or non-ESP group. That is, according to this study, students who took ESP not necessarily have a better chance of earning an A or a B on subsequent mathematics courses than the non-ESP students. So far, we have found that ESP students tend to do better only when they take a simultaneous ESP course. Later, we will see how the graduation rates between ESP and non-ESP students differ.

3.5 Analysis by Ethnicity

Figures 15 through 20 show the success rates between ESP and non-ESP groups for each course after adjusting by ethnicity. For each ethinicty, a contingency table was constructed to count the number of succesful students in each group and a Fisher's chi-square statistic was calculated to measure statistical significance between the treatment and control group.

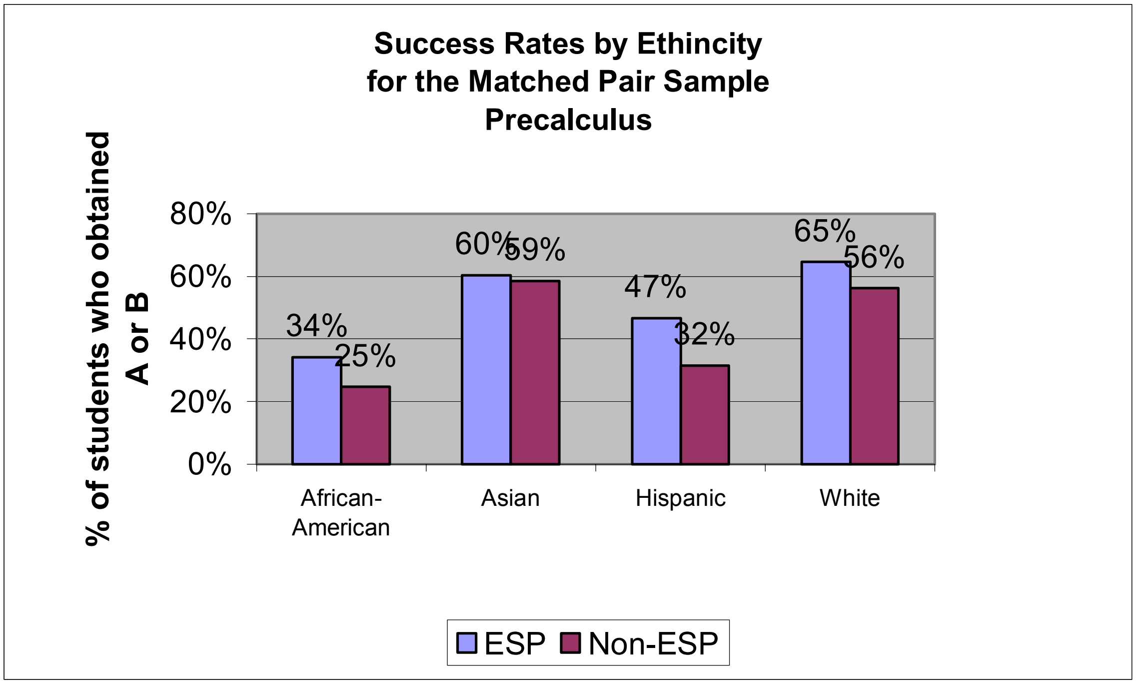

Figure 15.

In precaculus, there is statistical evidence to say that african-american students and hispanic students benefit from the program. For the asians and whites, even though Fisher's test does not find difference due to anything except randomness, the graph shows a slight superior percentage for ESP students.

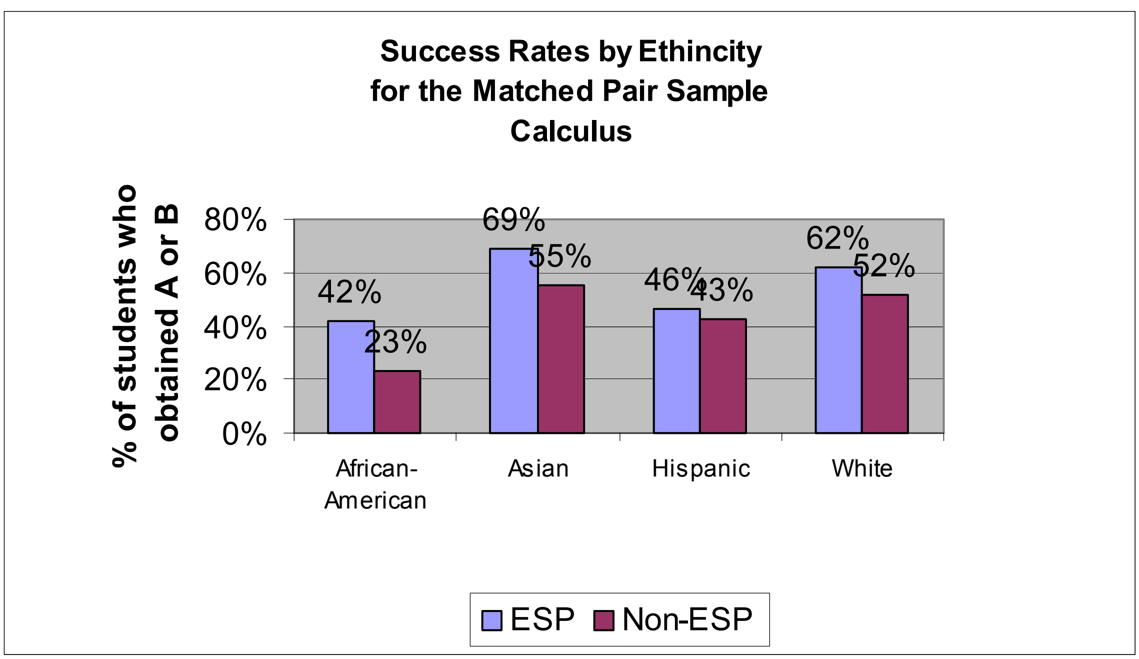

Figure 16.

In Figure 16 the percentage of students who obtained and A or B in Calculus are displayed. Again, the ESP students show superior performance. The african-americans and asians show statistical significance according to the Fisher's statistic.

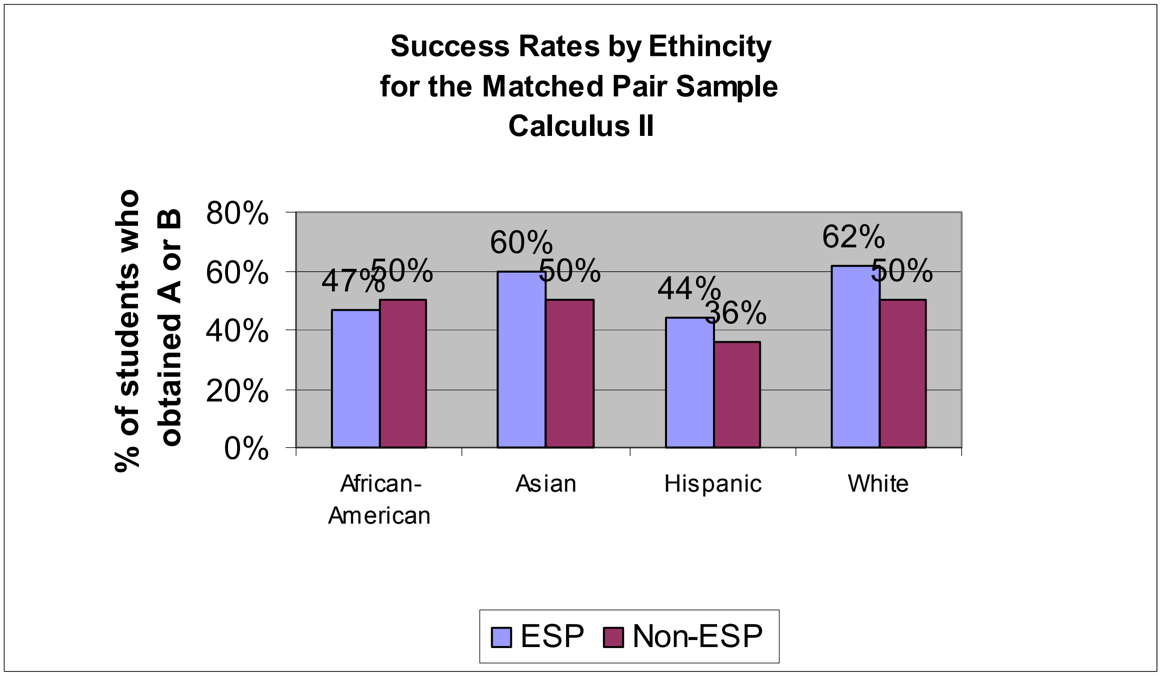

Figure 17.

Except for african-american students, all other ethnicities show a considerable difference in percentage of succesful students. Fisher's chi-square test still accepts the null hypothesis of independence between the program and the students' grades.

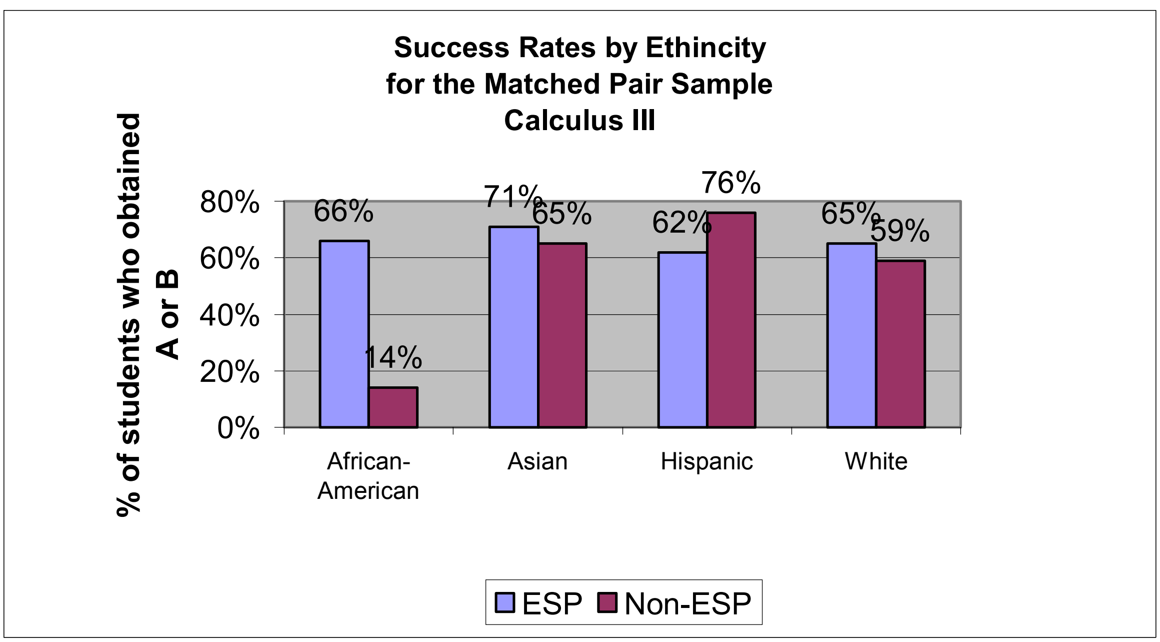

Figure 18.

The results for calculus III as shown in Figure 18 are mixed. The large difference in african-americans may be due to a small sample observed and the chi-square statistics are not relevant. The other ethnicities show no statistical evidence of dependence.

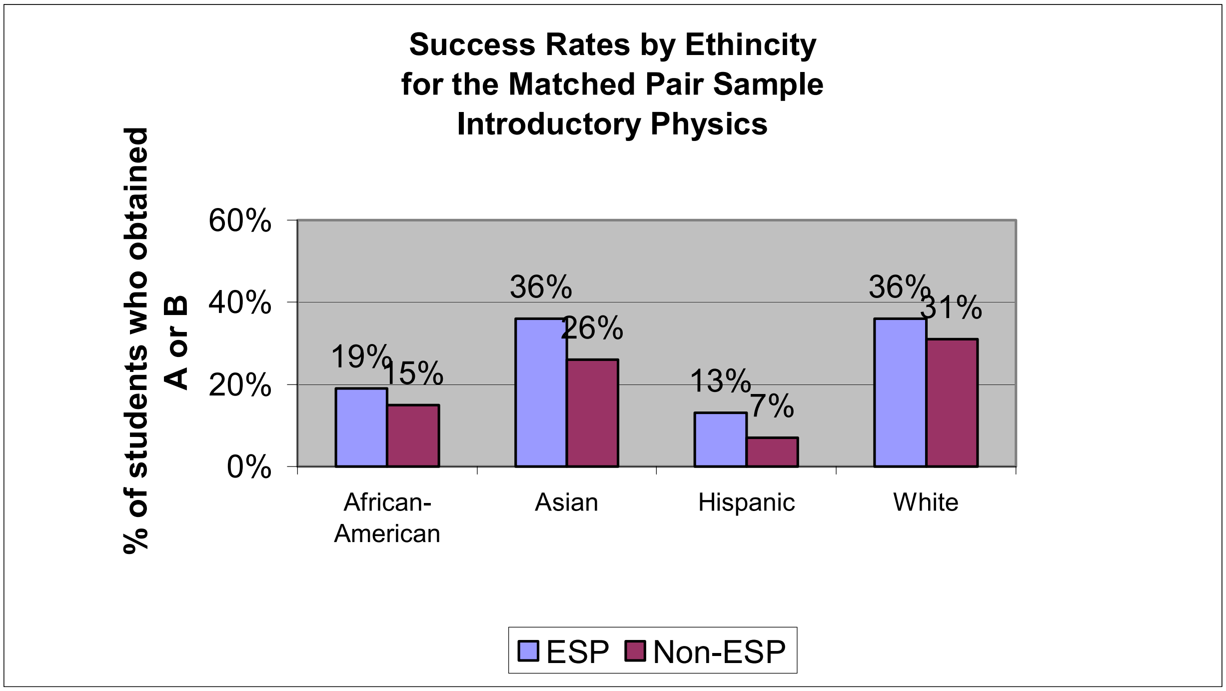

Figure 19.

Even though the results in Figure 19 look inclined towards the ESP students having better performance for all races, the results are not statistically significant. That is, after adjusting for ethnicity, the chi-square p-value for the four ethnicities is above 0.15.

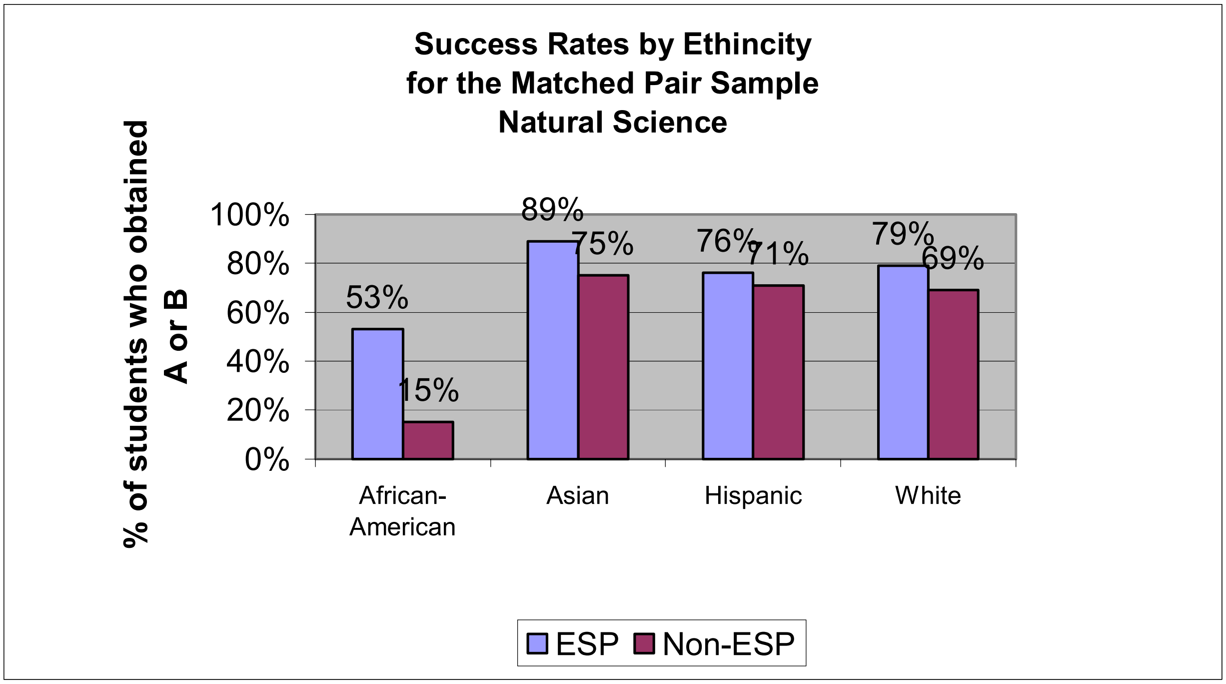

Figure 20.

As in Introductory Physics, Figure 20 shows better performance of ESP students in Natural Science. A small sample of only 28 african-american students was observed. Only 2 where successful non-ESP students and this might lead to incorrect inferences from Fisher's chi-square statistic.

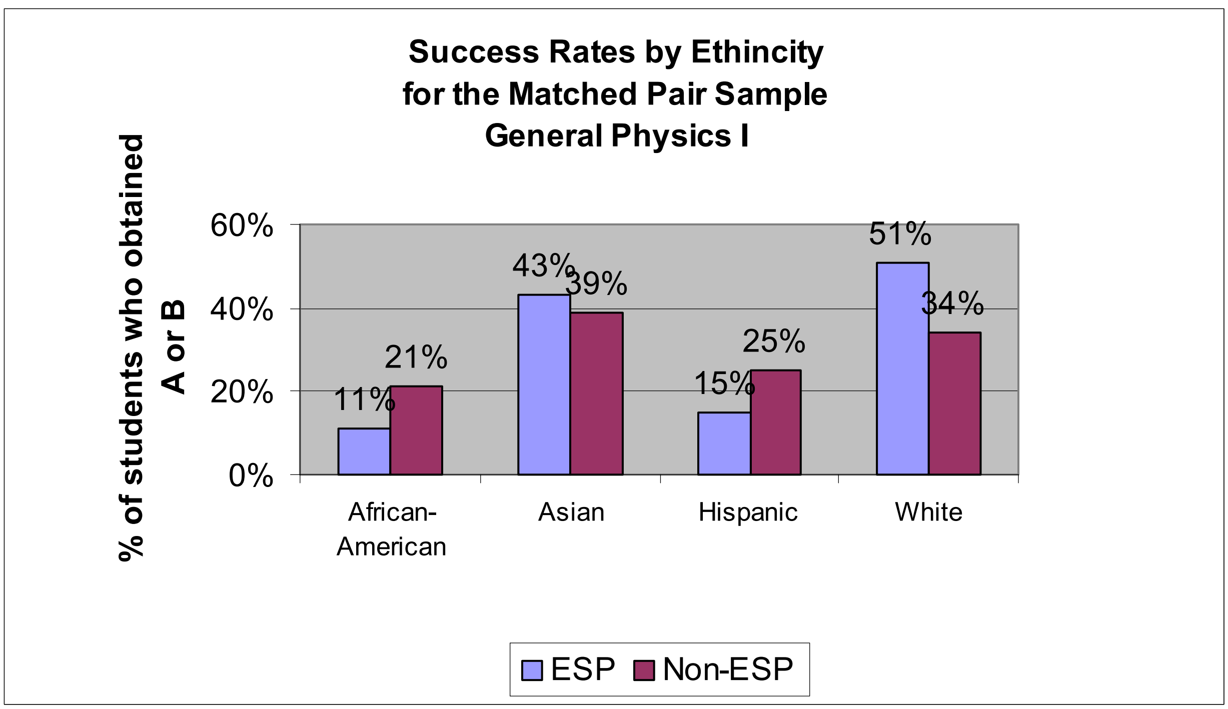

Figure 20.

In Figure 20 there is notably no influence of the ESP program on the students' grades. Although there is a statistically significant difference in white students from the emerging scholars program, the other ethnicities show no difference.

3.6. Graduation Rates

In order to try to analyze graduation tendencies in ESP and non-ESP students, we grouped the students by several categories. These include students who graduated within four to six years since entering the university as freshmen, students who took between six to eight years to graduate, students who graduated in more than eight years, and students that have spent more than six years of study without graduating. We call the GPA Criteria this: students that have spent more than six years of study without graduating we divided in two groups using their cumulative GPA to try to take into account the fact that students may transfer to other universities. Figures 21 and 22 show the results.

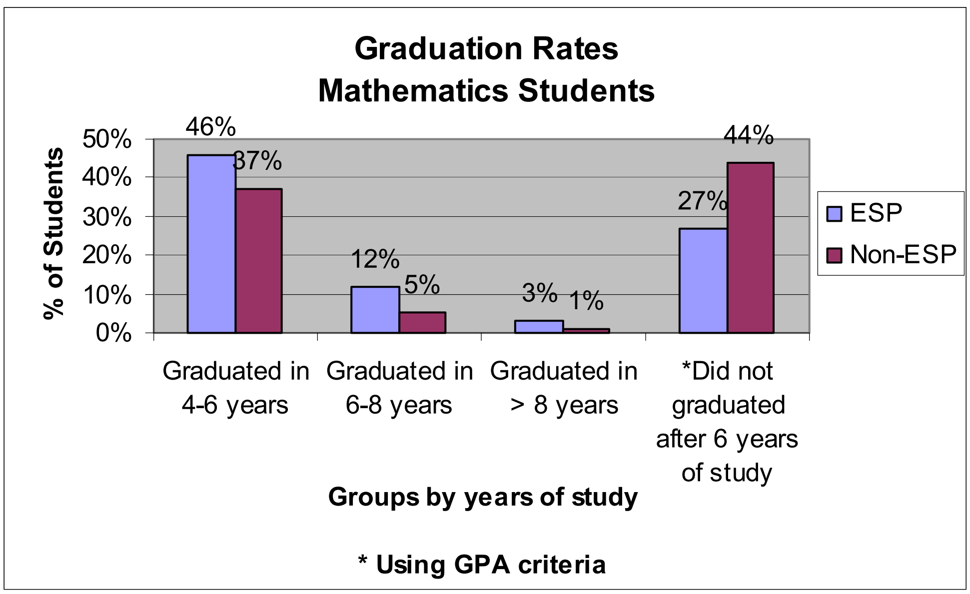

Figure 21.

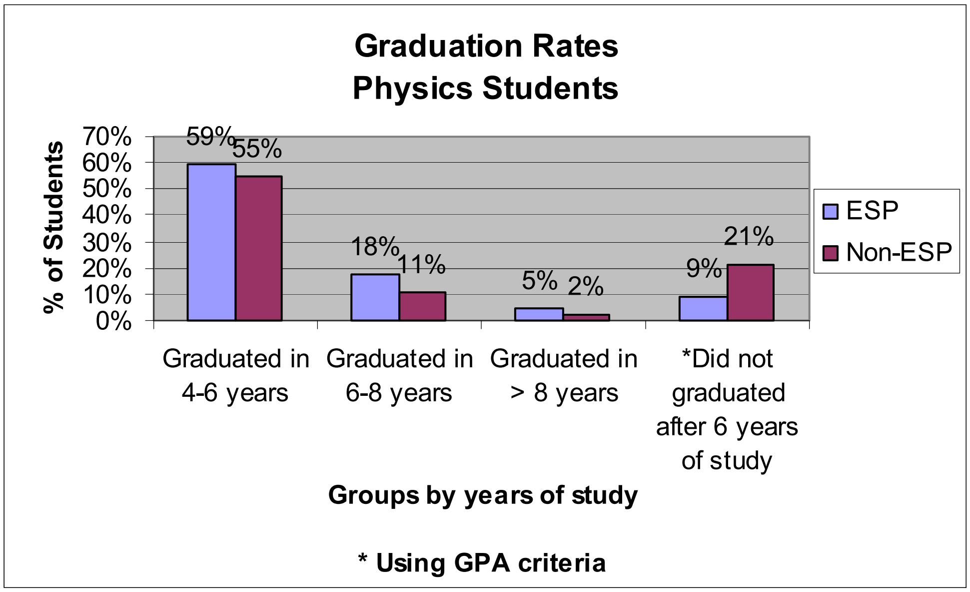

Figure 22.

From figure 21 we see that 58% of the mathematics ESP students graduated within 8 years of entrance to the university while 42% of the mathematics non-ESP students graduated within 8 years of entrance to the university. From figure 22 we see that 77% of the physics ESP students graduated within 8 years of entrance to the university while 66% of the physics non-ESP students graduated within 8 years of entrance to the university.

3.7 Conclusion

The Emerging Scholars Program students in mathematics and physics are more likely to earn a grade of A or B in the course than those students not in the program. This reflects the immediate influence that the program has on its students. A direct comparison of the performance of all ESP and non-ESP students would have given biased results. Thus, a matched sample by propensity score was built for each course since (1) the students were not assigned randomly to take the program and (2) the mean math GPA and ACT score for ESP students is greater than those that have never been part of the program. Figure 21 and 25 show that all courses,physics and math, show a statistically significanct difference between ESP and non-ESP students except in two courses: Calculus III and General Physics. The results in section 5 show that students who have been in the program do as well as students outside the program in subsequent advanced mathematics courses(differential equations). Results in section 6 show that the program benefits african-american and hispanic students considerably in precalculus and calculus, while white and Asian ESP students are significantly benefited in General Physics, Calculus I and II. The authors of the paper have several more ideas on improving this study. We want to find a model that predicts the probability that a student will receive an A or B in the course when his/her information is given. A logistic regression might be useful to predict such probability. We are also interested in including several new variables in our study and then determining if they are significant or not. A backward selection model might be suitable for such kind of analysis. Finally, this study reflects many positives but also points for improvement in the methodology of the program. Hopefully, future teaching assistants and counselors in the program will consider these results to better the program.

Acknowledgements

This work was carried out with the financial support of the Center for Excellence in Teaching and Learning to which we express our appreciation for their assistance. Without their help this project would not have been possible. Also, we want to express our deep appreciation to the initiators of this program at UIC, Prof. John Baldwin, Beatriz Jamaica, and Marty Gratzman, for their work in the ESP program and for their help in our work in this statistical study. We would also like to thank the Director of Computing Richard Chang, along with the graduate students Limin Wang and Barry Booton for their assistance.

References:

Cody, Ron and Pass, Ray. 2002. *SAS Programming by Example*. 7th ed. SAS Publishing.

Der, Geoff and Everitt, Brian S. 2002. *A Handbook of Statistical Analyses using SAS*. 2nd ed. Chapman & Hall/CRC.

Hogg, Robert and Craig, Allen. 1995. *Introduction to Mathematical Statistics*.5th ed. Prentice Hall.

Hosmer, D. and Lemeshow, S. (1989), *Applied Logistic Regression*: John Wiley and Sons, Inc.

Kosciuk, Steve. 1997. "Impact of the Wisconsin Emerging Scholars First-Semester Calculus Program on Grades and Retention from Fall '93-'96". Advanced Research Projects Agency.

Parsons, L. (2001), "Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques," Ovation Research Group, Seattle, WA. Paper 214-26.

Rice, John A. 1995. *Mathematical Statistics and Data Analysis*. 2nd ed. International Thompson Publishing.

Rosenbaum, P.R. and Rubin, D.B. (1983a), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," Biometrika, 70, 41-55.

Rosenbaum, P.R. and Rubin, D.B. (1985), "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," The American Statistician, Vol. 39, 33-38.

Rubin, D.B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: John Wiley and Sons, Inc.

Treisman, Uri. (1992), "Studying students studying calculus: A look at the lives of minority mathematics students in college." College Mathematics Journal 23(5), 362-372.

Yuan, Y. (2000), "Multiple Imputation for Missing Data: Concepts and New Development," SUGI 26 SAS Institute Inc. P267-25.

SAS/STAT User's Guide. 1999. Version 8, Cary, NC: SAS Institute Inc.