

“Wise Schooling”: An Examination of Steele’s (1997) “Existence Proof” for the Effect of  
Stereotype Vulnerability on the Academic Achievement of Black College Students –  
and an Alternative

William Collins

University of Michigan

October, 1998

## Abstract

Steele's (1997) "existence proof" for his model of "wise schooling" for African American college students is examined. The examination of Steele's data revealed that black students who were not in Steele's model themselves participated in distinct programs designed to promote academic achievement. The different groups of students were differently prepared for college which was an important factor used by admissions officers to select students and to assign them to a given program. Although the groups of students differed in terms of their characteristics, the groups did not differ significantly on the criterion measure of First Semester Grade Point Average (FGPA). Analysis of covariance and computation of estimated least squares means resulted in findings which fail to replicate those reported by Steele. Steele's "existence proof" for a model of wise schooling for black college students based on a theory of stereotype vulnerability is found to be lacking and susceptible to misinterpretation. A more parsimonious explanation and an alternative model for promoting academic achievement in black college students are offered.

“Wise Schooling”: An Examination of Steele’s (1997) “Existence Proof” for the Effect of  
Stereotype Vulnerability on the Academic Achievement of Black College Students -- and  
an Alternative

The issue of black student academic achievement in college has been explored by Steele (1997) who reported that a “wise schooling” approach could reduce the effect of “stereotype threat” and promote academic achievement. But an examination of Steele’s “existence proof” for the effect of his version of “wise schooling” on academic achievement finds it not only to be lacking, but easily misinterpreted as well. This commentary is written from the perspective of one familiar both with the operation of Steele's model and also with the comparison groups. (Indeed, the author serves as program Director of one of the comparison groups.) However, the perspective one brings to the analysis is considerably less important than what the data themselves actually suggest. Moreover, the issue of black student academic progress at the college level is much too important for the nation as a whole to be cast as a differences in perspective or theories and instead requires the kind of considered analysis Steele attempts to provide.

Steele has maintained that a “wise schooling” approach based on his theory of stereotype vulnerability can be effective in promoting the academic achievement of college students in comparison to other approaches. Specifically, the model he and his colleagues developed at Michigan and which was called the 21st Century Program (21CP) was compared to another program (not mentioned by name in the article), the Comprehensive Studies Program (CSP), and to a control group of students who were not subject to intervention. Steele offered a graph in his article (Figure 5 in the June 1997 American Psychologist article) which he argues provided an existence proof “that an intervention derived from the [stereotype vulnerability] theory could stop or reverse a

tenacious negative trajectory in the school performance of stereotype-threatened students” (Steele, 1997). I suggest that the effect itself remains to be demonstrated and that there is an alternative and more plausible explanation for the effect shown in Steele’s Figure 5 than stereotype vulnerability and the intervention derived from it; that explanation is preparation for college work as indexed by standardized test score. An unfortunate feature of Steele’s Figure 5 is that it provides the reader with no sense of the distribution of standardized test scores within groups. In fact, it suggests a distribution that does not exist. I provide here a similar analysis as it applies to Steele’s data and in a fashion that allows the reader to understand the character of the distributions of standardized test scores for the different groups.

To begin, some consideration must be given to the local picture at the University of Michigan within which Steele’s model was tested. Michigan is a large university with over 36,000 students; in reality a number of intervention strategies exist to promote student success at Michigan, but three distinct programs include minority student retention among other objectives and form the comparison groups for our analyses. The 21st Century Program is a retention program that is based on Steele’s theory of stereotype vulnerability and which attempts to lessen or eliminate vulnerability among participants. The Comprehensive Studies Program (CSP) is a student retention program that emphasizes an intensive instructional and advising model; that is, it stresses the development of a proper work ethic as well as academic skill building among students and provides the opportunity for more contact with teachers and advisors than is typically the case. The Summer Bridge Program (SB) is a conditional admission program that allows a select group of students to begin their university studies in the summer preceding the freshman year and to develop skills in such areas as mathematics or writing prior to fall semester enrollment. It is important to note that students selected for the Summer Bridge Program typically are chosen precisely because they have relatively low standardized test scores, yet exhibit outstanding potential for college success in other

ways, for example, through good grades or leadership activities in high school; it also should be noted that, except for the conditional admission program, these students would not otherwise have the opportunity to enroll at Michigan. The Summer Bridge Program is a subset of the Comprehensive Studies Program and represents about ten percent of an entering class of CSP students. Students may elect to participate in any combination of the three programs described. Students are normally selected for CSP and for Summer Bridge by the admissions office. Prospective students in the 21st Century Program are identified by its staff through a separate application process for admission to a “Residential Learning Community” which includes assignment to a specific residence hall; students may also be encouraged to apply by staff via telephone.

Steele refers to CSP as a “remedial” program, which probably is not an appropriate description; rather CSP embodies a comprehensive model for facilitating academic achievement, which emphasizes intensive instruction, regularly scheduled active advising opportunities, and student development through such efforts as collaborative learning, peer advising, and freshman interest groups. It is unfortunate that the term “remedial” has developed a pejorative cachet because, whether used appropriately or inappropriately, it serves to deflect attention from any true benefits that may result from special efforts to promote student success, remedial or otherwise. Nevertheless, to the extent that the term “remedial” encompasses a focus on the development of good study habits and concern for improved academic competence, then clearly all three intervention programs qualify. To the extent that “remedial” means correcting deficiencies, then none of the programs qualifies, although I would not quibble with one who insisted on such a label for the Summer Bridge Program due to the large differences in standardized test scores its students exhibit in comparison to others in the competitive Michigan context.

Given this overall local picture, at least five groupings of black students are possible. Those who participated in the Summer Bridge Program, those who participated

in CSP but not Summer Bridge, those who were in both CSP and the 21st Century Program, those who were in the 21st Century Program but not in any CSP Program and a Control Group of black students who participated in neither CSP nor 21st Century Programs. Steele's 21CP has been in existence since 1991 and the data reported in his American Psychologist article are taken from years 1991 and 1992. The current analysis examines data from academic years 1991-92 through 1996-97. I am grateful to Claude Steele and his colleague Steve Spencer for providing copies of their data sets which included years 1991 through 1994. I have created additional new data sets for years 1995-96 and 1996-97 using all black students who were in the 21CP and the Summer Bridge Program and randomly selected students from CSP and the Control Group. An initial concern is that Steele makes no distinction between students in Summer Bridge and CSP and there should be. The assignment of subjects to groups is more appropriately indicated by the following:

- Summer Bridge (a subset of CSP; but not in 21CP) (n=262)
- CSP only (i.e., not Summer Bridge) (n=862)
- 21st Century and CSP (n=154)
- 21st Century only (n=97)
- Control Group (Blacks not in CSP or 21CP) (n=671)

This examination will be concerned with the black students in Steele's study, as they clearly are the focus of his intervention model. Steele's existence proof argues that students in the 21CP perform better academically than others and that the slope of their regression line for academic achievement is steeper than for blacks in the so-called remedial program. Yet, only the graph in his Figure 5 is offered to demonstrate this effect. Other statistics that might better give the reader a fuller picture of the nature of the variables used in Steele's analysis simply are not provided. An initial point might be to look at student characteristics before they entered college and then to establish whether

there are differences in academic achievement before proceeding to a comparison of regression lines. In other words, for the five groups of students who comprise Steele's subjects, what is the basic structure of the variables used and how do the different groups compare?

-----  
Insert Figures 1, 2 and 3 about here.  
-----

Figures 1 and 2 are graphs showing mean levels of academic achievement as measured by High School Grade Point Average (HSGPA) and Standard Test Score for the five groups of students (Standard Test Score was standardized based on the national population of test takers). One sees immediately that there are substantial differences between the groups in terms of mean HSGPA and standardized Standard Test Scores; that is, there are wide discrepancies among the groups in terms of preparation for college work. Figure 3 shows mean first-semester college grade point average (FGPA) for the different groups over a six year period. One sees that the FGPA's for the different groups are rather comparable. All the groups occupy a narrow band of FGPA achievement between about 2.5 and 2.8. Observe that 21CP and Control Group students share similar profiles and have relatively high standardized Test Scores and HSGPAs. In contrast, CSP and Summer Bridge students enter college with good, but substantially lower HSGPAs and test scores. Are there significant differences between the groups on these academic achievement measures?

Table 1 shows the means obtained by the different groups on academic achievement variables. Students who participated in 21CP are observed to have higher HSGPA, Test Score, and FGPA. Table 2 is a summary table based on the results of ANOVA and comparing 21CP to the other groups on the three academic achievement variables. During none of the years examined was there a significant difference between

21CP and Control Groups on any of the variables. For five out of six years there was a significant difference between 21CP and Summer Bridge students on HSGPA and Test Score (in each case  $p < .05$ ). Similarly, there was a statistically significant difference on Test Score between 21CP and CSP students for five out of six years examined. However, no significant differences were found between 21CP and other groups for any year on the criterion variable of FGPA.

-----  
Insert Table 1 about here  
-----

Although these data suggest that the academic achievement of students in the 21st Century Program and the comprehensive program is mediated by levels of pre-college preparation, Steele emphasizes the difference in slopes of the regression lines for GPA vs. standardized test score as the really important issue and this question requires closer attention. At base, Steele asserts that stereotype vulnerability depresses the academic performance of black students and also that programs designed to address specific academic needs, such as the Summer Bridge or comprehensive program described here, can have the effect of accentuating both stereotype vulnerability and its depressive effects on achievement. As proof he offered a graph, his Figure 5, depicting first-semester grade-point average (FGPA) as a function of program and race controlling for high school GPA (HSGPA). The graph depicts a linear relationship between variables, reflecting the assumption of the ordinary least squares regression analysis; the graph also suggests a wide distribution of subjects along the entire regression line, which would mean that there were large numbers of subjects from each group at the extremes (that is, two standard deviations beyond the mean in Steele's Figure 5).

However, it should be pointed out that the University of Michigan is a highly selective institution and standardized test scores for all groups of students are higher than



national averages. But when the standardization is based on the local population the well-known difference of one standard deviation in mean test score between blacks and whites is apparent. Thus, all of the blacks groups have a mean test score that is lower than the local mean. Yet, the mean standardized test score for black students who participated in the 21st Century-only was well above the national mean (indeed no students were below it), while in contrast the mean standardized test score for participants in the Summer Bridge Program was below the national mean. Therefore, for whatever reason, the students who elected to join the 21st Century Program tended to be exceptionally well prepared before entering college in comparison to other black students in the study, while the Summer Bridge participants, in contrast, were chosen for that program precisely because they were not so well prepared, but exhibited qualities other than high test scores. Thus, within the local population the students in the 21st Century Program were concentrated at the high end of academic preparation as measured by HSGPA and test score, while Summer Bridge students were concentrated at the lower end. Steele's analysis, illustrated by the graph in his Figure 5, obscures any group differences that may exist in the distributions of students along the dimension of standardized test score and creates an inaccurate impression of the relationship between FGPA and test score by program and race.

-----  
Insert Figure 4 about here.  
-----

Figure 4 is a set of bar graphs of the adjusted mean FGPA for the five groups over a six-year period. The adjusted mean FGPA was calculated for each year with the use of ANCOVA in which FGPA was the criterion variable and HSGPA and test score were the covariates. Effectively, each bar represents an estimate of what the mean FGPA's for the groups would have been if each had had a common standardized test score and common HSGPA identical to the actual means across all groups. The results of the ANCOVA for

these groups of students consistently show that students in the Summer Bridge Program demonstrate the highest gain in predicted FGPA and not students in the 21CP. This is significant not only because it fails to replicate Steele's findings, but also because students in the Summer Bridge Program are the most academically at-risk, are required to participate in their program as a condition of admission, and therefore should be the most susceptible to heightened stereotype threat as posited by Steele.

Figure 5 is a graph similar to Steele's Figure 5, but one which properly assigns students to groups and is based on six years of data. This graph shows that students in the Summer Bridge Program have a regression line that has both the highest elevation and the flattest slope. The second highest elevation, and the regression line with the steepest slope, is for students in the CSP. Steele's 21 CP group is third. Interestingly, both the Control Group and students who participated in both 21CP and CSP were observed to have negative slopes over the six-year period.

-----  
Insert Figure 5 about here  
-----

This examination of students in the 21CP in comparison to other programs suggests that the existence proof for Steele's intervention model is lacking. What the analysis actually suggests is that, although the concept of stereotype vulnerability is intellectually appealing, its impact on black student achievement in a real school context is questionable. More importantly, there appears to be a more parsimonious explanation for the differences that do exist: students who are better prepared tend to perform better academically; and programs that help students improve their preparation for academic work or which pointedly seek to develop their academic abilities lead to improved performance. A complex theory of stereotype vulnerability simply is not needed to account for the differences in academic achievement that have been observed.

An alternative to a student success model based on overcoming stereotype vulnerability is the comprehensive model described here and for which the present data ironically provide an effectiveness existence proof. The comprehensive approach acknowledges the different circumstances from which students may emerge as they seek to realize their potential through higher education. Steele (1997) is almost certainly correct in his assumption that sustaining success in school requires identification with school achievement and that one must perceive good prospects for achievement in the schooling domain as well. Likewise, those who pursue higher education clearly identify with schooling. But realizing one's potential in the face of substantial disparities in preparation is a daunting task; it is rather like running a footrace but starting many meters behind the other runners. To win the race, you must first close the gap. The comprehensive model emphasizes doing so early and places a positive focus on such effort while being honest with students about what is required of them in terms of commitment to their goals. The comprehensive model includes intensive instruction, both academic and personal advising, the development of sound study habits, and active involvement in the total university community. Many programs adhering to similar models exist nationwide and they are unabashedly eclectic, welcoming --indeed, even seeking out-- effective concepts and approaches wherever they may arise. A notion like stereotype vulnerability is certainly worthy of consideration as the basis for one among many tools these programs have shown are required for meeting the challenges they face. But the true practical significance of the concept remains to be demonstrated.

## References

Steele, C. M. (1992, April). Race and the schooling of black Americans. *The Atlantic Monthly*, 69-79.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52 (6), 613-629.

List of table captions.

Table 1. Means on academic achievement variables for selected groups.

Table 2. Years in which there was a statistically significant difference between 21CP and other groups on academic achievement variables over a six year period.

List of figure captions.

Figure 1. Mean HSGPA for selected groups over six years.

Figure 2. Mean standardized test score for selected groups over six years.

Figure 3. Mean college FGPA for selected groups.

Figure 4. Bar graph of FGPA for selected groups adjusted for HSGPA and Test Score over six years.

Figure 5. FGPA as a function of Test Score and HSGPA for selected groups.

Table 1. Means on academic achievement variables for selected groups.

	<u>HSGPA</u>	<u>Test Score (standardized)</u>	<u>FGPA</u>
SB (n=262)	2.98	-.49	2.52
CSP (n=862)	3.18	-.002	2.54
21CSP (n=154)	3.15	.08	2.46
21CP (n=97)	3.43	.64	2.80
Control (n=671)	3.38	.62	2.77

Table 2. Years in which there were statistically significant differences between 21CP and other groups on variables of interest over a six year period.

<u>Group</u>	<u>Variable</u>		
	<u>HSGPA</u>	<u>Test Score</u>	<u>FGPA</u>
21CP vs.			
Summer Bridge	'91, '93, '94 '95 & '96	'91, '92, '93 '94 & '96	-
CSP	'94, '95, '96	'91, '92, '93, '94 & '96	-
21CSP	'91	'93 & '94	-
Control	-	-	-

(No significant differences were found between 21CP and other groups on FGPA for any years between 1991 and 1996; nor were there any significant differences between 21CP and the Control Group on any variable).

Fig. 1 - Mean HSGPA for Selected Groups over Six Years

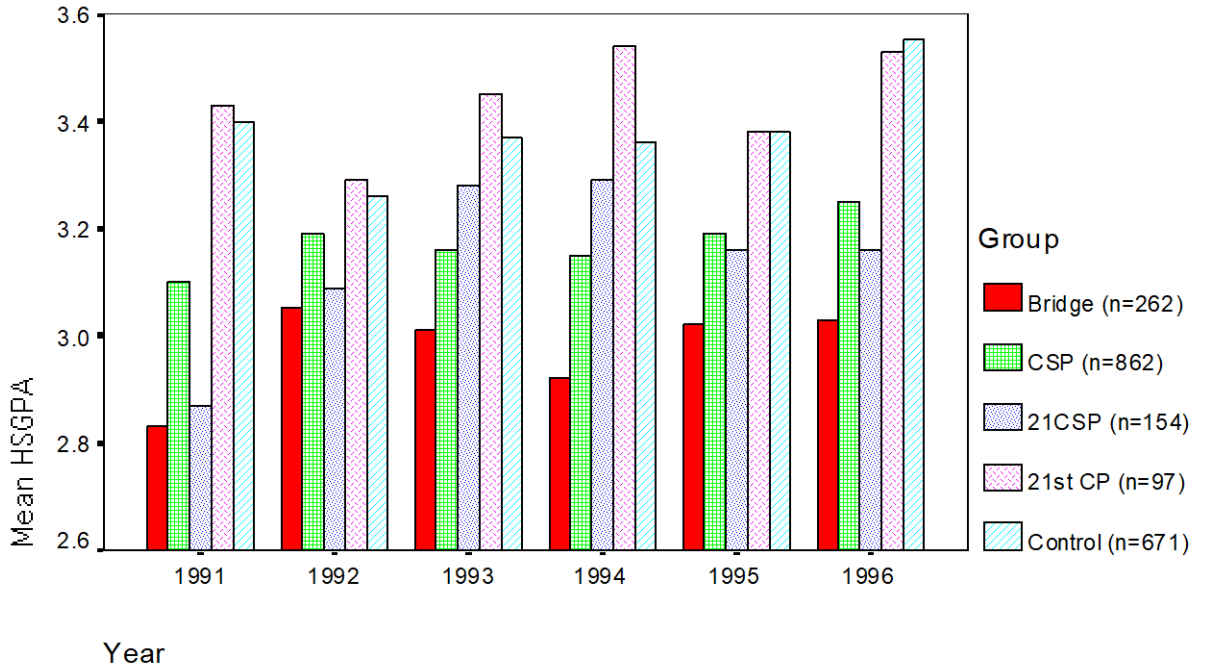




Fig. 2 - Mean Standardized Test Score  
for Selected Groups by Year.

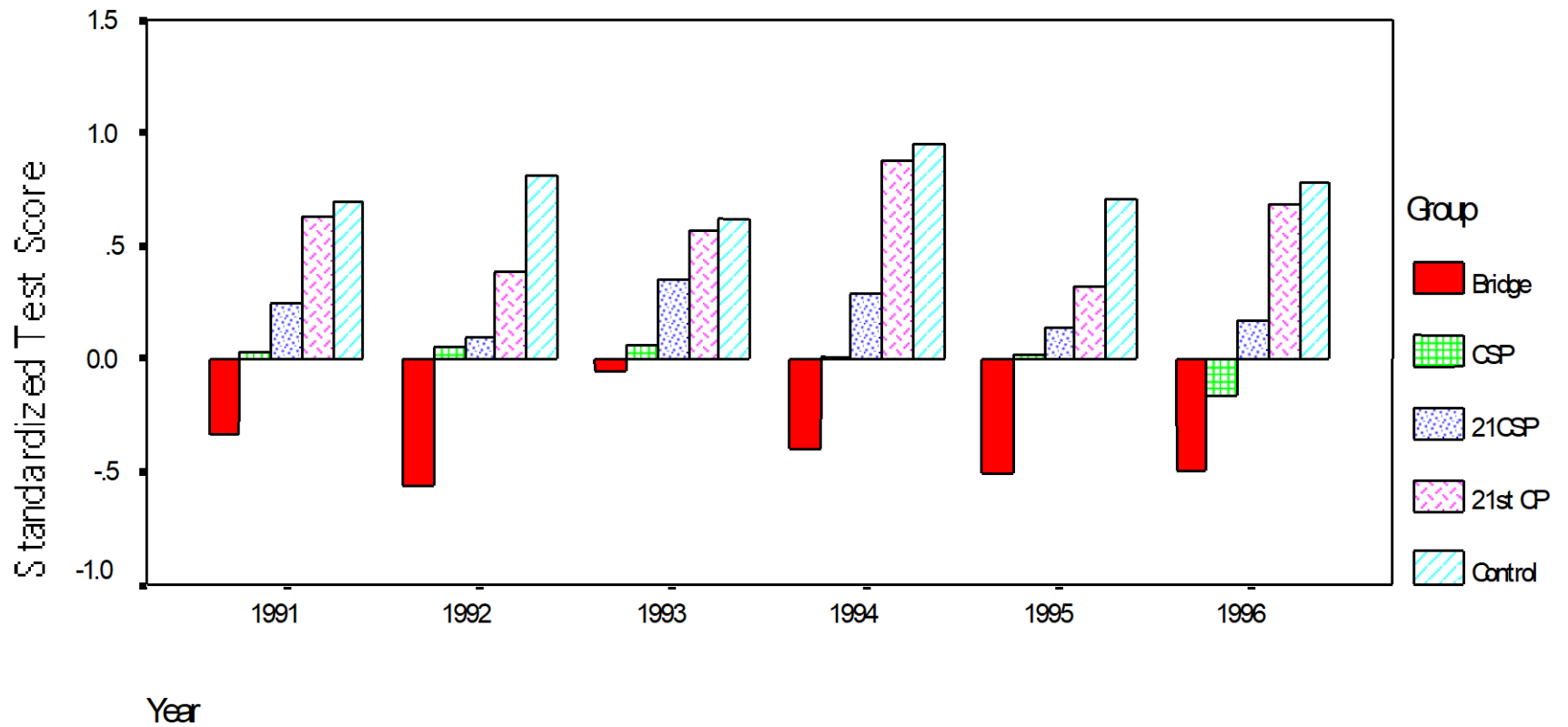
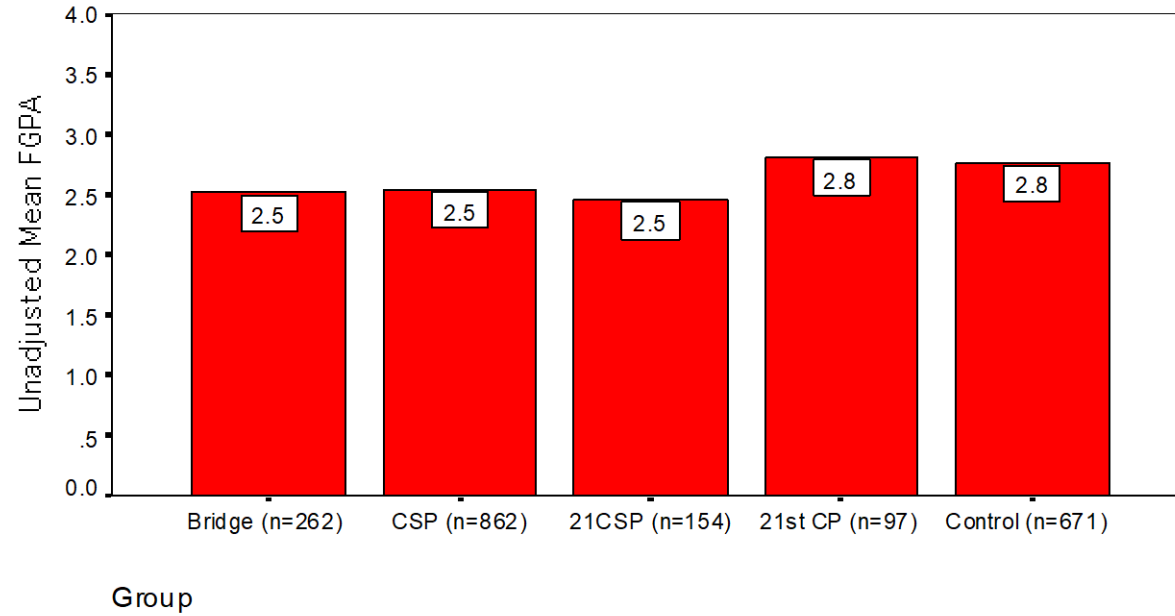


Fig. 3 - Mean FGPA\* for Selected Groups for Six Year Period.



\*First-semester Grade Point Average in College

Fig. 4 - Mean FGPA for selected groups adjusted for HSGPA & Test Score for each of six years.

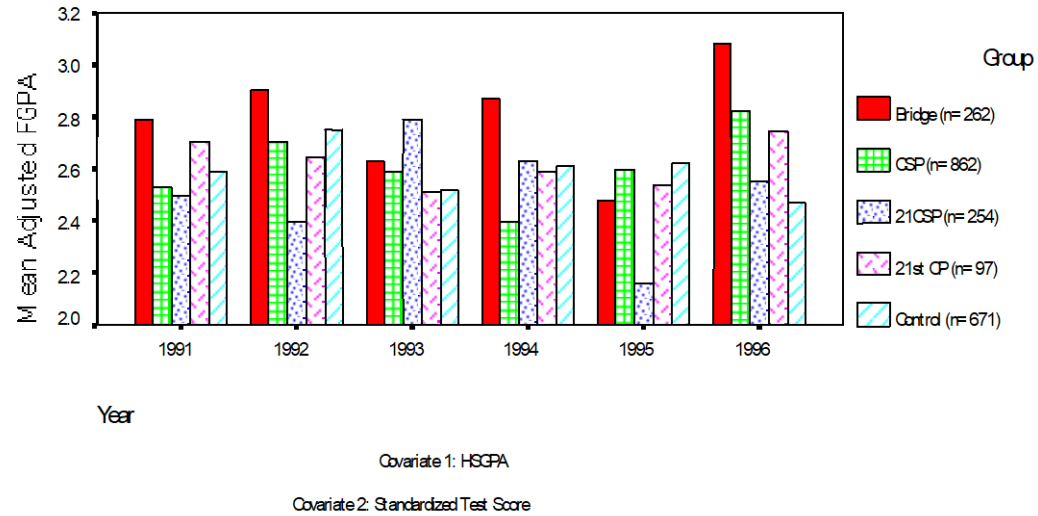


Fig. 5 - FGPA as a function of HSGPA and Test Score  
for selected groups over six-years (1991-97).

