Journal of Exercise Physiology**online**
(JEPonline)

Volume 8 Number 1 February 2005

Official Research Journal of The American Society of Exercise Physiologists (ASEP)

ISSN 1097-9751

Pediatric Exercise Physiology

PHYSICAL FITNESS AND ACADEMIC ACHIEVEMENT

JAMES B GRISSOM.

ABSTRACT

Grissom JB. Physical Fitness And Academic Achievement. ***JEP****online* 2005;8(1):11-25 . The purpose of this study was to evaluate the relationship between physical fitness and academic achievement. To do so, scores on the **FITNESS**GRAM®, a physical fitness test, were compared to reading and mathematics scores on the Stanford Achievement Test 9th edition, a standardized norm-referenced achievement test. Subjects were all 5th, 7th, and 9th grade California school children enrolled in public school in 2002 for whom there was complete data on both the physical fitness and academic achievement tests. The sample size was 884,715 students. Results indicate a consistent positive relationship between overall fitness and academic achievement. That is, as overall fitness scores improved, mean achievement scores also improved. This relationship between fitness and achievement appeared to be stronger for females than males and stronger for higher socio-economic status (SES) than lower SES students. Results should be interpreted with caution. It cannot be inferred from these data that physical fitness causes academic achievement to improve. It is more likely that physical and mental processes influence each other in ways that are still being understood.

Key Words: School Children, Activity, Grades, Learning

INTRODUCTION

The health benefits of regular physical exercise are widely acknowledged. For example, a front-page story in the Sacramento Bee (1) reported that poor diet and physical inactivity might soon overtake tobacco as the leading cause of death (2). The Health section of CNN.com reported that for cancer survivors the key ways to prevent the return of the disease included a healthy diet and exercise (3).

Even though the benefits of physical exercise are acknowledged, physical education in public schools is viewed as an extracurricular activity and physical education teachers have experienced first hand when money is tight and/or when there is pressure to improve test scores, physical education is one of the first activities to be cut back or eliminated. However, if there is evidence that physical education has a direct positive effect on important educational domains such as reading and mathematics, it could be argued that physical education is not extracurricular. Rather, it is a vital component in students' academic success.

It is not the intention of this paper to argue that the importance of physical education is its benefit to academic achievement. The overall health benefits of organized physical activity are probably much more important than possible academic benefits. However, when policy makers need to make difficult decisions about where to spend public funds and administrators need to make decisions about where to focus resources in a climate of academic accountability, a proven relationship between physical fitness and academic achievement could be used as an argument to support, retain, and perhaps even improve physical education programs.

At this time little research has examined the relationship between physical fitness and academic achievement. One study that reported a consistent and significant relationship between fitness and achievement (4) had several threats to validity. First, there were validity concerns with the academic achievement indicator. Academic achievement was based on a non-standardized, subjective five-point rating scale. Therefore, the meaning of achievement could vary from site to site. Second, the reported correlations, although statistically significant (i.e., at 0.001, 0.01, & 0.05 levels of significance) were not impressive. For example, the average correlation between sit-ups and achievement for girls age 7 to 15 was 0.13. Third, there were unexplained inconsistencies. Sit-ups showed a statistically significant relationship with achievement for 8, 9, 12, 13, 14, & 15-year old girls but not for 7, 10, & 11-year old girls. The study's conclusion of a consistent and significant relationship between fitness and academic achievement was based on the fact that there were more statistically significant correlations than there were not.

One review of research that concluded a positive relationship between physical and mental skills expressed concern that reviewed studies did not demonstrate causality (5). Most reviewed studies used correlation designs. Reviewed experimental studies had design weaknesses. One study with an experimental design in that the independent variable was manipulated did not use random assignment or matching to control for preexisting group differences (6). Another experimental study employed random assignment but failed to find a statistically significant difference in academic achievement between experimental and control subjects (7).

Another review of research, that concluded a positive relationship between fitness and achievement, expressed concern with the methodological weaknesses in studies that used correlation designs (8). For example, the Vanves study, one of the reviewed studies was an unpublished manuscript where the sample size was small and the abstract failed to explain how experimental and control subjects were matched. As for experimental studies, the reviewer argued the failure to find a statistically significant difference in academic achievement between experimental and control subjects was not problematic. Even though experimental subjects were receiving less instructional time, because they were being pulled out of class to engage in physical education, they performed as well as, if not better, than controls.

The failure of experimental designs to find statistically significant differences between experimental and control subjects is due in part to the difficulty in raising academic achievement. It is very difficult to raise student achievement, beyond what might be expected, even when that is the specific focus (9). A study intended to affect achievement indirectly would encounter even more difficulty.

Little research has examined the relationship between physical fitness and academic achievement, and the research evidence that exists is not strong. One reason research evidence may be scarce is that it is difficult to obtain valid and reliable measures for both physical fitness and academic achievement on the same subjects. Another reason is that it is difficult to obtain large representative samples of students who have both achievement and fitness scores.

The purpose of this study was to evaluate the relationship between physical fitness and academic achievement. The study attempted to address some of the weaknesses of earlier studies. For example, this study had valid and reliable measures of both physical fitness and academic achievement and a large sample size. However, the criticism that correlation studies do not demonstrate causality is a weakness of the current study.

METHODS

California Law's Education Code (EC) mandates
…during the month of February, March, April, or May, the governing board of each school district maintaining any of grades five, seven and nine shall administer to each pupil in those grades the physical performance test designated by the State Board of Education (10).

In February 1996, the California State Board of Education designated the **FITNESS**GRAM® as the required physical fitness test to be administered to California students. The Cooper Institute for Aerobics Research in Dallas, Texas developed the **FITNESS**GRAM®.

In spring 2001, the California Department of Education (CDE) began reporting the Physical Fitness Test (PFT) results each year. All students in grades five, seven, and nine are expected take the PFT. Students who are physically unable to take the entire test are to be given as much of the test as possible.

The **FITNESS**GRAM® provides a number of options for each performance task so all students, including those with special needs, have the maximum opportunity to complete the test. The PFT measures five aspects: 1) aerobic capacity, 2) body composition, and 3) trunk strength, 4) upper body strength, and 5) flexibility. The **FITNESS**GRAM® requires students to complete the following six tests:

- one of the options from aerobic capacity
- one of the options from body composition
- the curl-up test
- the trunk lift test
- one of the options from upper body strength
- one of the options from flexibility

The **FITNESS**GRAM® uses standards established by the Cooper Institute for Aerobics Research to evaluate fitness performance. Performance is classified as: (1) in the healthy fitness zone (HFZ) which means students met the fitness target or (2) needs improvement which means students failed to meet the target. It is possible on some tests to exceed the HFZ target. Scores exceeding the target are included with students that have scored within the HFZ.

Since the HFZ standard is considered a minimal level of achievement, students must meet all of the fitness standards to be considered fit. Overall test scores range from zero, none of the standards were met, to six, all standards were met or exceeded.

For academic achievement testing, California Law's Education Code (EC) states
…there is hereby established the Standardized Testing and Reporting Program, to be known as the STAR Program.
(b) Commencing in the 1997-98 fiscal year and each fiscal year thereafter, and from the funds available for that purpose, each school district, charter school, and county office of education shall administer to each of its pupils in grades 2 to 11, inclusive, the achievement test designated by the State Board of Education (11).

Each spring the STAR program administers four distinct standardized achievement tests: the California Standards Tests (CST), a standardized norm-referenced test, a Spanish assessment of basic skills, and the California Alternative Performance Assessment (CAPA). STAR tests were administered first in the spring of 1998. From 1998 to 2002, the standardized norm-referenced test was the Stanford Achievement Test 9th edition (SAT/9), Form T. In 2003, the standardized norm-referenced test was changed to the California Achievement Test, Sixth Edition Survey (CAT/6). The Spanish assessment of basic skills has been the Spanish Assessment of Basic Education, Second Edition (SABE/2). Tests are administered to all public school students enrolled in grade 2 through 11.

Collecting demographic information, such as birth date and gender, is part of the test administration process for both PFT and STAR. This study uses data from tests administered in the spring of 2001 and 2002. Demographic information from the PFT and STAR programs was used to create matched files. The county/district/school (CDS) code, grade level, birth date, and gender were used to match student records. Each California school has a unique CDS code. This matched file contained both the PFT scores and standardized achievement test scores for individual students. As such, these data were used to evaluate the relationship between overall scores on the PFT and the standardized achievement tests. There could be errors in the matching process but there was no reason to believe matching errors biased the results.

Matched files were created for 2001 and 2002. Sample sizes for 2001 and 2002 were 634,112 and 884,715 students, respectively. The sample size increased in 2002 because of increased PFT participation. Because sample sizes were larger, this study reports 2002 results.

The analysis had two parts. First, for each overall PFT score, which ranged from zero to six, the average SAT/9 reading and mathematics scores were calculated in normal curve equivalent (NCE) units. The NCE score is a normalized standard score with a mean of 50 and a standard deviation of 21. NCE scores range from 1 to 99. The average test score by overall PFT score provides an indicator of the relationship between fitness and achievement. Second, analysis of variance (ANOVA) was used to test the relationship between the overall PFT score and achievement scores.

RESULTS
Cohort Sample Compared to Population
To evaluate whether the matched cohort was representative of the larger population who took the PFT and the SAT/9, analyses compared PFT and SAT/9 scores for the two groups. Figure 1 shows the percent of 2002 students achieving the different number of fitness standards for the statewide population compared to the matched cohort sample. The number of fitness standards achieved and the overall PFT score are interchangeable terms.

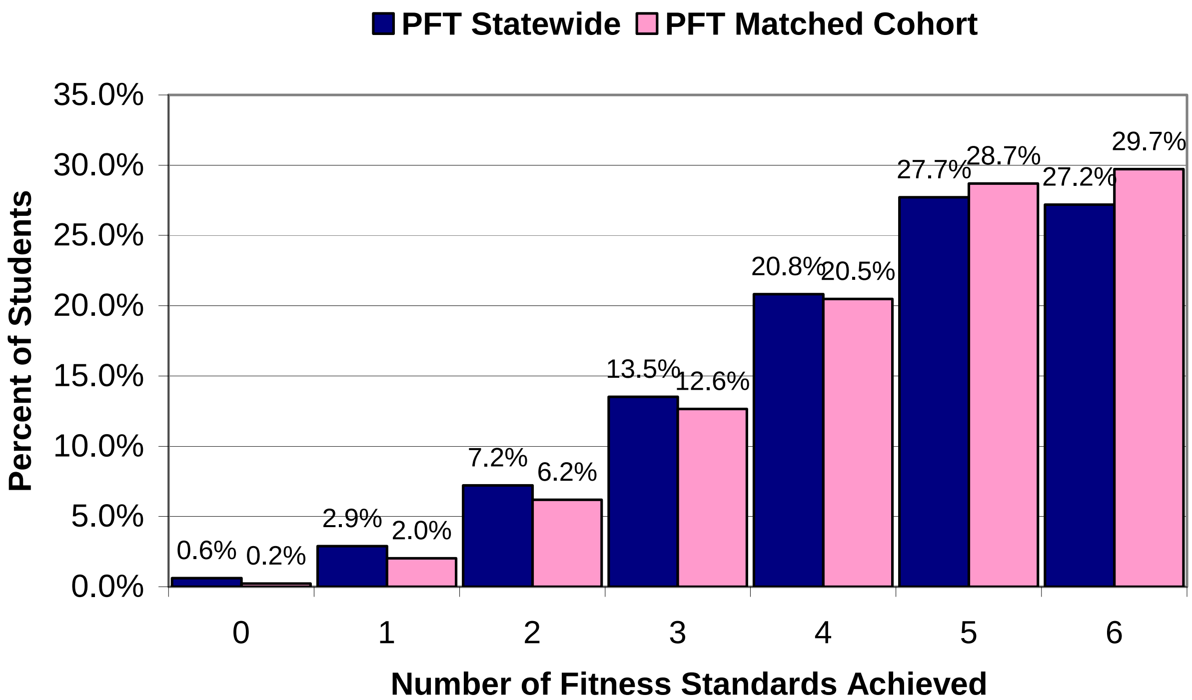

Figure 1. Summary of fitness standards achieved for 2002 grades 5, 7, & 9; n = 1,090,248 (statewide) 884,715 (cohort)

To insure comparability of overall PFT scores, only students with complete data were included. If missing data were included, the overall PFT score, except for a score of six, could have multiple meanings. A score of less than six could mean (1) the test was incomplete, (2) the student was absent, or (3) the student failed to achieve the minimum standard on one to six tests. When there was no missing or incomplete data, the meaning of scores was clear: students attempted all of the tests and were able to demonstrate minimal competency on the number of tests indicated by the score.

Figure 1 indicates a discrepancy between the 2002 statewide population and the matched cohort sample. The matched cohort demonstrated a slightly higher performance on the PFT. For example, only 27% of the population achieved all six PFT standards compared to 30% of the matched cohort.

Figure 1 is negatively skewed. That is, more students achieved higher than lower scores. Even so, the fitness standard was considered a minimal level of achievement. Therefore, only students who achieved all six PFT standards were considered minimally fit.

Next, average reading and mathematics NCE scores for the sample and population were compared to evaluate whether the sample differed markedly from the population in terms of academic achievement. Table 1 compares academic achievement scores for the population to the matched cohorts.

Table 1. SAT/9 mean test scores for 2002

	SAT/9 Reading NCE		***SAT/9 Mathematics NCE***	
Grade Level	Statewide	Matched Cohort	Statewide	Matched Cohort
5	47	48	53	55
7	47	48	51	53
9	41	43	52	54

Although the mean achievement scores were close, the matched cohort demonstrated slightly higher achievement on the SAT/9 reading and mathematics tests than the statewide population. Further analyses should be interpreted with this score difference in mind.

The Relationship Between PFT and Mean SAT/9 Scores

To evaluate the relationship between physical fitness and academic achievement, average achievement scores were calculated for each PFT score. Figure 2 shows the average reading and mathematics NCE scores on the 2002 SAT/9 by the number of fitness standards achieved (i.e., the overall PFT score) on the 2002 PFT.

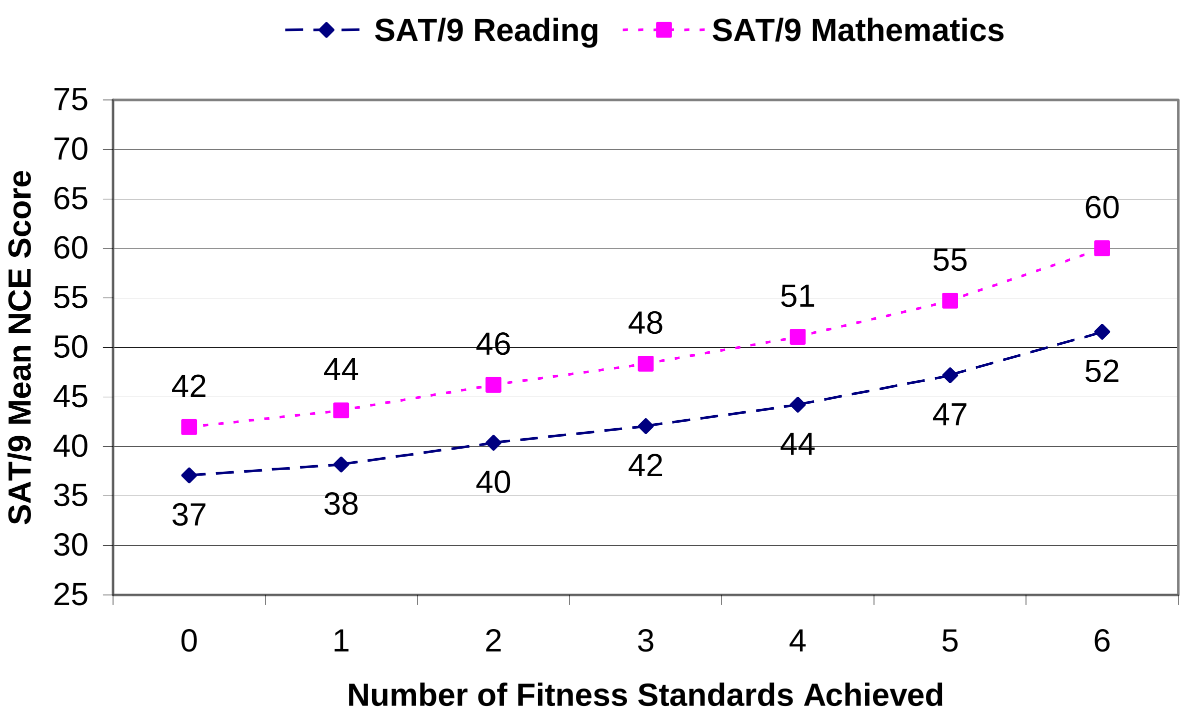

Figure 2. 2002 grade 5, 7, & 9 SAT/9 mean reading and mathematics NCE scores by overall PFT scores; n = 890,280 (reading) 888,241 (mathematics).

The average SAT/9 reading NCE score for students who did not achieve any of the fitness standards was 37. This same NCE score for mathematics was 42. As the overall PFT score improved the mean reading and mathematics NCE scores also improved. Results indicate a positive relationship between academic achievement and physical fitness.

Although evidence suggests a relationship between physical fitness and academic achievement, there are more formal tests. Analysis of Variance (ANOVA) was used to test this relationship. Table 2 shows these results.

Table 2. ANOVA results when the dependent variable was SAT/9 reading scores and the independent variable was the overall PFT score: 2002 Grades 5, 7, & 9

Source	df	SS	MS	F	Pr > F	R Square
Regression	6	9990420.1	1665070.0	4146.9	< .0001	0.0362
Residual	661885	265758622.5	401.5			
Total	661891	275749042.6				

ANOVA results suggest that some, or all, of the mean reading scores are statistically different from each other. Tukey's studentized range test for multiple comparisons revealed that each mean reading score for each PFT score was significantly different from all other reading score means at $P < 0.01$. For example, students who had a PFT score of six had a mean reading score of 52. The mean score of 52 was statistically greater than the mean score of 47, the mean reading score for students who had a PFT score of five. Students who had a PFT score of five had a mean reading score that was statistically greater than students who had a PFT score of four or less. Students who had a PFT score of four had a mean reading score that was statistically greater than students who had a PFT score of three or less and so on.

Table 3 shows these same results for SAT/9 mathematics scores.

Table 3. ANOVA results when the dependent variable was SAT/9 mathematics scores and the independent variable was the overall PFT score: 2002 Grades 5, 7, & 9.

Source	df	SS	MS	F	Pr > F	R Square
Regression	6	14822504.8	2470417.5	5617.3	< .0001	0.0487
Residual	658658	289670477.1	439.8			
Total	658664	304492981.9				

As with the reading scores, the mean scores are statistically different from each other and Tukey's studentized range test for multiple comparisons revealed that each mean for each of the PFT scores was significantly different from all other means at $P < 0.01$.

Even though the means were statistically significant from each other the R^2 for each of the ANOVA analyses was small; indicating considerable unaccounted test score variance. The R^2 for reading and mathematics was 0.04 and 0.05 respectively.

The correlations between overall PFT scores and SAT/9 scores were not high. Results are shown in Table 4.

Table 4. Correlation of 2002 PFT scores and SAT/9 scores: Grades 5, 7, & 9.

Variables	SAT/9 Reading	SAT/9 Mathematics
PFT score	0.186	0.217
SAT/9 reading		0.768

Even though there was a relationship between PFT scores and average achievement scores, SAT/9 scores have a lot of variance within the overall PFT scores, which explains why the R^2 data from the ANOVA analyses were so low.

Figure 3 graphically illustrates the variance of SAT/9 reading NCE scores for students who had an overall PFT of 6.

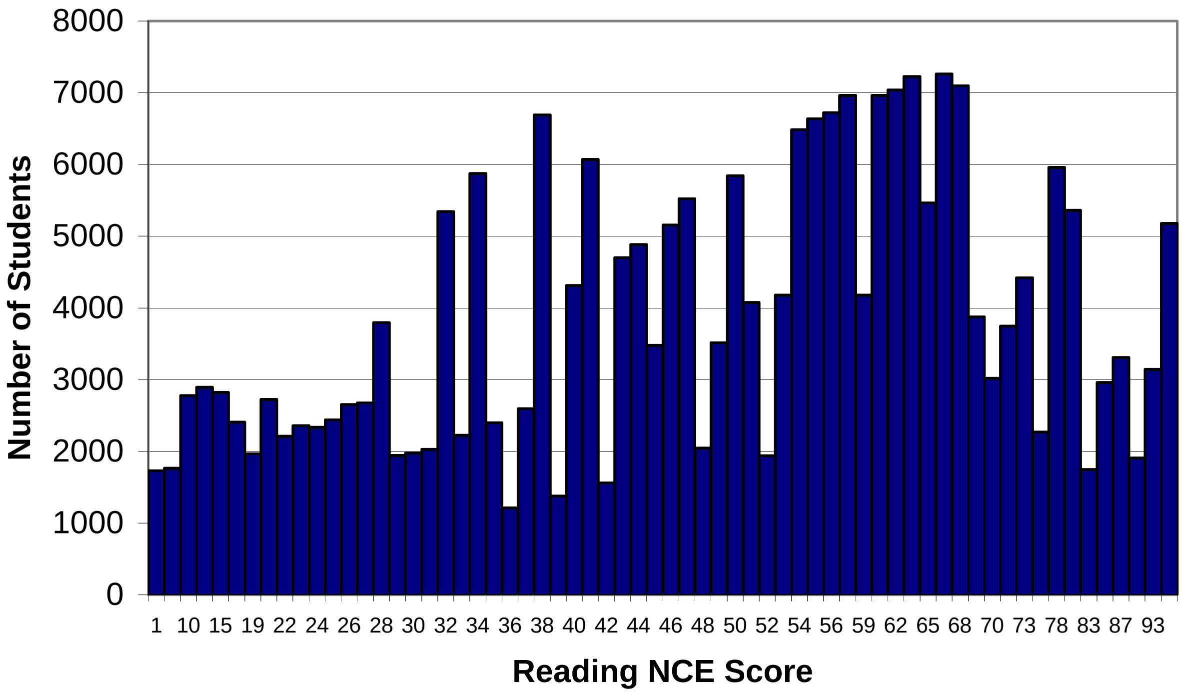

Figure 3. Distribution of SAT/9 reading NCE scores for students with a PFT of 6; n = 247,538.

Students with an overall PFT score of 6 had an average reading NCE score of 52. Figure 3 shows many of these students had scores higher and lower scores than 52. The standard deviation was 20.

The PFT student record indicated whether the students failed, passed, or exceeded the standard. If there is a relationship between physical fitness and academic achievement, then students who *exceed* the standard should have higher average scores than students who merely *pass*.

Figure 4 shows the SAT/9 mean NCE scores by whether students exceeded, met, or failed on five different PFT tests.

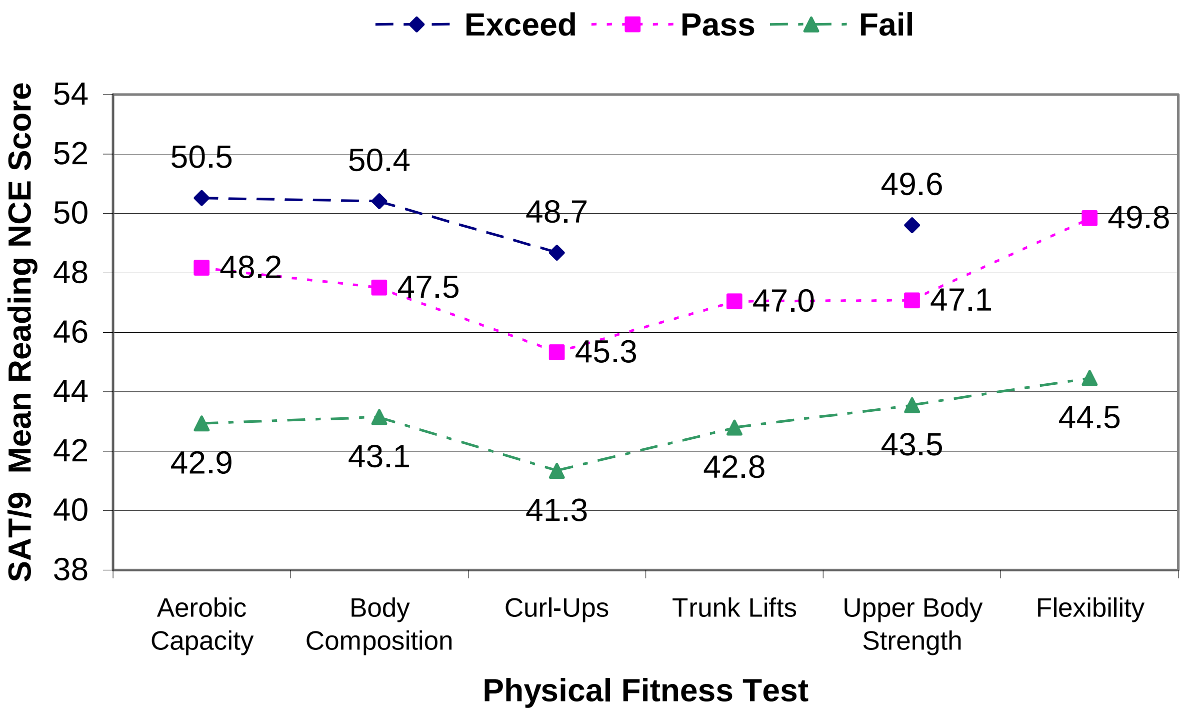

Figure 4. 2002 grades 5, 7, & 9 SAT/9 mean reading NCE score by five different PFTs; n = 890,280.

For PFT tests, except Trunk Lifts and Flexibility, students who *exceeded* on the PFT had higher average SAT/9 reading scores than students who *passed* and students who *passed* the PFT had higher average SAT/9 reading scores than students who *failed*. Physically fitter students, even when fitness was defined in different ways, had higher achievement test scores.

The *exceed* data for Trunk Lifts were not included in Figure 4 because the criteria defining *exceed* were unclear. For Trunk Lifts, while lying face down, students attempted to lift their upper body 12 inches off the floor and hold that position long enough to measure the distance. Recorders assigned a value of 0 to 99. High values in this test (e.g., 50 inches and higher) made no sense. In reality, students either met the standard by lifting their body 12 inches or they didn't. High score values were classified as *exceed* but these high values were nonsensical and were not included in this analysis.

There was no *exceed* category for the Flexibility test. There were only two categories, (1) met the standard or (2) failed to meet the standard.

The Relationship Between PFT and Mean SAT/9 Scores by Subgroups

Subgroup analyses were conducted to evaluate how the relationship between fitness and achievement was modified by other student characteristics. First, the relationship between physical fitness and academic achievement was evaluated by gender. Figure 5 shows at the relationship between 2002 PFT scores and SAT/9 mean reading scores by gender.

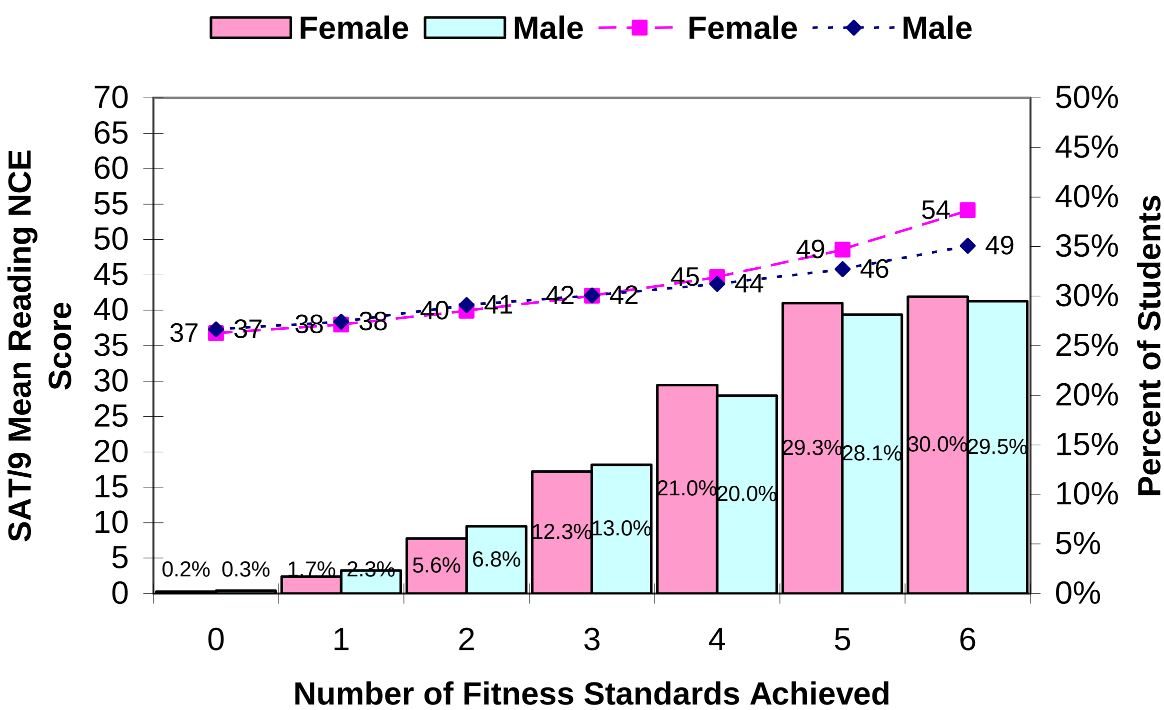

Figure 5. 2002 grade 5, 7, & 9 SAT/9 mean reading NCE score by overall PFT score and gender, n = 436,139 (female), 448,576 (male)

Figure 5 shows that female and male students demonstrated comparable levels of fitness. For example, 30 percent of females and 29 percent of males had an overall PFT score of 6. There were modest test score differences by gender. The overall mean reading NCE score of 48 for females was slightly higher than overall average of 45 for males. Even though there were test score differences, the relationship between fitness and achievement was consistent across gender. For females and males, as the overall PFT score increased so did the mean reading achievement scores.

Even though there were no differences in fitness and little difference in test scores, there appeared to be an interaction effect. The rate of change in achievement scores was higher for females than males. The difference in mean reading NCE scores between the least and most fit students was 17 NCE points for females and 12 NCE points for males.

To evaluate whether results were consistent across socio-economic status (SES), the relationship between physical fitness and academic achievement was evaluated by NSLP. NSLP is the national school lunch program. NSLP indicates whether or not students received free or reduced lunch and serves as a proxy for SES. Participation in NSLP is an indicator of lower SES. No NSLP participation is an indicator of higher SES. Figure 6 shows these results.

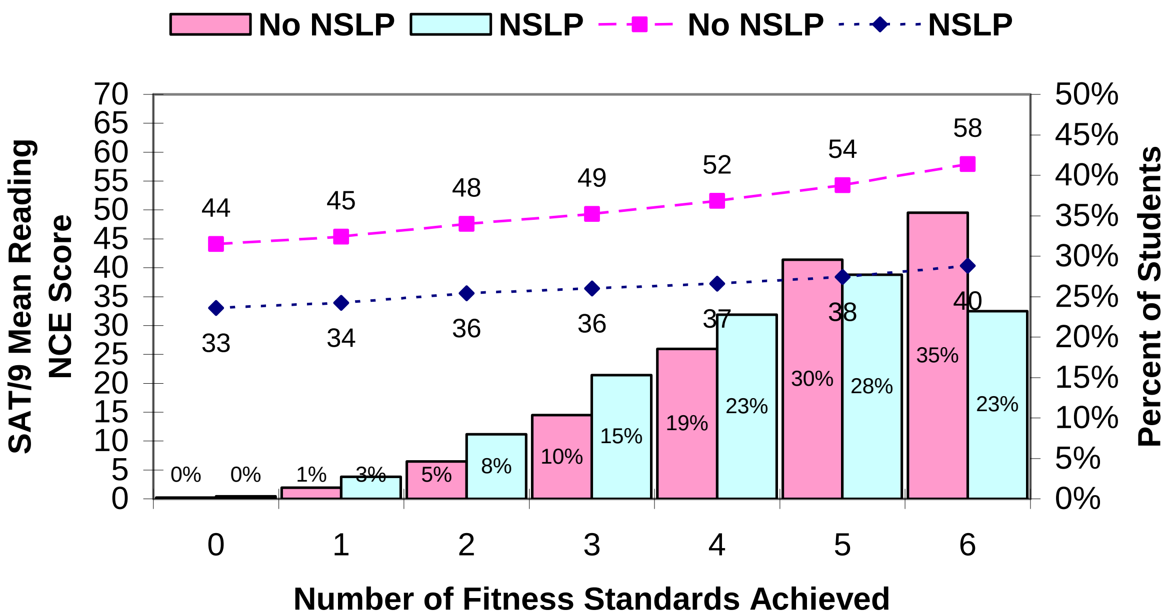

Figure 6. 2002 grades 5, 7, & 9 SAT/9 mean reading NCE score by overall PFT score and NSLP, n = 477,300 (no NSLP), 412,980 (NSLP)

Results indicate that a greater proportion of no NSLP (i.e., higher SES) students achieved the higher PFT scores (i.e., scores of 5 & 6) than NSLP (i.e., lower SES) students. Figure 5 also indicates that test scores were different for different levels of SES. Students classified as No NSLP had higher achievement scores than NSLP students. Even so, as the PFT score increased so did the mean achievement for both SES levels. However, the rate of change in achievement scores was greater for no NSLP students than for NSLP students. The difference in mean reading NCE scores between the least and most fit students was 14 NCE points for no NSLP and 7 NCE points for NSLP.

Although evidence suggests the relationship between physical fitness and academic achievement was different for females than males and different for higher SES than lower SES students, there are more formal tests. Analysis of Variance (ANOVA) was used to test the relationship between physical fitness and academic achievement by gender and SES. Two variables were used as indicators of SES, NSLP and parent education. Parent education was coded into five levels: (1) Not a high school graduate, (2) high school graduate, (3) some college (includes AA degree), (4) college graduate, (5) graduate school or post graduate training. Less formal education is an indicator of lower SES and more formal education is an indicator of higher SES.

For these analyses, rather than multiple comparisons, planned orthogonal polynomial contrasts were used to test linear, quadratic, cubic, quartic, and quintic trends. Results are shown in Table 5.

Table 5. ANOVA results when the dependent variable was SAT/9 reading scores and the independent variables were overall PFT score, gender, NSLP, and parent education; 2002 Grades 5, 7, & 9.

Source	df	SS	MS	F	Pr > F	R Square
Regression	48	71502698.8	1489639.6	4827.1	< .0001	0.2593
Residual	661843	204246343.8	308.6			
Total	661891	275749042.6				
PFT Score	6	2218592.6	369765.4	1198.2	< .0001	
Gender	1	28401.0	28401.0	92.0	< .0001	
NSLP	1	789460.3	789460.3	2558.2	< .0001	
Parent Education	4	1742978.4	435744.6	1412.0	< .0001	
Gender*PFT Score	6	262654.8	43775.8	141.9	< .0001	
NSLP*PFT Score	6	71565.6	11927.6	38.7	< .0001	
Parent Ed.*PFT Score	24	121711.4	5071.3	16.4	< .0001	
Contrast	**df**	**SS**	**MS**	**F**	**Pr > F**	
Linear	1	160572.5	160572.5	520.3	< .0001	
Quadratic	1	301.9	301.9	1.0	0.3226	
Cubic	1	933.7	933.7	3.0	0.082	
Quartic	1	719.0	719.0	2.3	0.1269	
Quintic	1	978.5	978.5	3.2	0.075	

Results indicate a statistically significant linear relationship between overall PFT scores and SAT/9 reading scores. The linear trend is significant at $P < 0.0001$. As the PFT score increases the average SAT/9 NCE reading score increases. The R^2 was 0.26; indicating considerable unaccounted test score variance. There may even be variables, not included in the model, that mediate the relationship between physical fitness and academic achievement. However, available data suggest a strong linear relationship between fitness and achievement.

There were significant interaction effects. ANOVA results and Figure 5 indicate that the rate of increase in mean achievement scores by PFT scores was greater for females than males. ANOVA results and Figure 6 indicate that the rate of increase in mean achievement scores by PFT scores was greater for students not receiving NSLP (i.e., higher SES) than for students receiving NSLP (i.e., lower SES). The interpretation of the interaction effect of parent education with PFT score is comparable to NSLP. That is, the rate of increase in mean achievement scores by PFT scores was greater for students with more highly educated parents (i.e., higher SES) than for students with less educated parents (i.e., lower SES).

ANOVA results using SAT/9 mathematics NCE scores as the dependent variable are shown in Table 6.

Table 6. ANOVA results when the dependent variable was SAT/9 mathematics scores and the independent variables were overall PFT score, gender, NSLP, and parent education: 2002 Grades 5, 7, & 9

Source	*df*	*SS*	*MS*	*F*	*Pr > F*	*R Square*
Regression	48	67442782.9	1405058.0	3903.8	< .0001	0.2215
Residual	658616	237050199.0	359.9			
Total	658664	304492981.9				
PFT Score	6	4912795.0	818799.2	2274.9	< .0001	
Gender	1	3980.9	3980.9	11.1	< .0001	
NSLP	1	635533.6	635533.6	1765.8	< .0001	
Parent Education	4	1178744.3	294686.1	818.8	< .0001	
Gender*PFT Score	6	83536.2	13922.7	38.7	< .0001	
NSLP*PFT Score	6	104521.4	17420.2	48.4	< .0001	
Parent Ed.*PFT Score	24	188040.7	7835.0	21.8	< .0001	
Contrast	*df*	*SS*	*MS*	*F*	*Pr > F*	
Linear	1	386110.0	386110.0	1072.8	< .0001	
Quadratic	1	0.4	0.4	0.0	0.9744	
Cubic	1	4482.4	4482.4	12.5	0.0004	
Quartic	1	42.0	42.0	0.1	0.7326	
Quintic	1	26.3	26.3	0.1	0.7869	

Results were comparable to reading scores. That is, there was a significant linear relationship between PFT scores and SAT/9 mean NCE mathematics scores and there were significant interaction effects. The interpretation of the interactions is the same as with reading scores. That is, the rate of change in mean SAT/9 mathematics scores by overall PFT scores was greater for females than males and greater for higher SES than lower SES students.

DISCUSSION

When the overall PFT score was compared to mean SAT/9 reading and mathematics scores, there was a consistent positive relationship between physical fitness and academic achievement. ANOVA analyses revealed, as overall PFT scores increased, mean achievement scores also increased in a statistically significant way. There was a statistically significant positive linear relationship between fitness and achievement.

ANOVA analyses also revealed statistically significant interaction effects. For example, as PFT scores increased, mean achievement scores increased at a greater rate for females than males. This indicates that relationship between fitness and achievement was stronger for females than males. Also, as PFT scores increased, mean achievement scores increased at a greater rate for higher SES students than for lower SES students. This indicates that relationship between fitness and achievement was stronger for higher SES students than for lower SES students.

Results need to be interpreted with caution. First, it is possible that the relationship between fitness and achievement was mediated by variable(s) not included in this study. For example, higher SES is generally indicative of higher academic achievement. Higher SES is also indicative of better general health (12). It's possible that better general health or better living conditions were responsible for both higher fitness levels and higher levels of academic achievement.

Second, a major axiom of social science research is that correlation is not causality. It cannot be inferred from these data that physical fitness increased or improved academic achievement. There was no time or logical ordering that automatically lead from one event to the other. It is just as logical to believe that mental capacity affects physical ability. For example, there is evidence that mental stress can lower the effectiveness of the immune system (13).

This study is seen as preliminary. Even so, the notion that conditions that promote a healthy body also promote a healthy mind seems a tentative conclusion. It's known for example, that the greatest risk factor for low social functioning (e.g., inability to attend to tasks) was low SES (14). That's because lower SES students suffer more family turmoil, live in more chaotic households, have fewer and less supportive social networks, have fewer cognitive enrichment opportunities, and live in more polluted, unhealthy environments than higher SES students (15). Although the dynamics of these processes are still being understood, studies in biology (16) and developmental psychology (17, 18) suggest that mental and physical processes are mutually dependent and it is often difficult to determine what is causing what.

Studies in biology and developmental psychology also suggest that biological systems are never static but are constantly changing and change is dependent on context. For example, the highland people of New Guinea develop at a physically different rate than comparable Chinese and English populations due to the high altitude, little water, and a low protein diet (19). Research indicates that activities that positively enrich both mental and physical process promote the positive development of the whole system (17). Enriched mental environments produce brain growth throughout life and that when people are physically and mentally relaxed they learn best (20). Although improved aerobic capacity by itself is not going to improve reading achievement, physical and other activities that promote good health seem to promote intellectual capacity.

Public funded nutritional (e.g., NSLP) and physical education programs may play a major role in the health of students especially lower SES students. The home environments of lower SES students may be such that even with these programs their general health will be lower than higher SES students and as a result will on average demonstrate lower academic achievement and lower physical fitness than higher SES students. However, it's worth considering that eliminating health related programs to focus attention on improving test scores might impact the health of lower SES students to such a degree that their academic achievement is impacted negatively.

Even though it cannot be inferred from correlation data that physical fitness causes academic achievement to improve, correlation and/or naturalistic designs may be the best models for preliminary studies. First, these models offer the best opportunity to build theory about phenomena *by better understanding the constructs, what they consist of, and how they relate to other constructs* (21). Second, the difficulty of raising achievement may limit the ability of experimental designs to find a causal relationship even when one exists. This is not an argument against using experimental designs. It is simply an argument that experimental designs may be premature until the relationship between physical fitness and academic achievement is better understood.

REFERENCES

1. Griffith D. Obesity on track to be No. 1 killer. ***Sacramento Bee***. 2004, March 10, p. A1, 15.
2. Mokdad, AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States. ***JAMA* 2004**;291(10):1238-1245.
3. Carney K. (2003, December 12) Study: Exercise may help ward off cancer's return. *CNN.com, Health*. Retrieved March 29, 2004, from http://www.cnn.com/2003/HEALTH/12/12/hln.fit.cancer.exercise/index.html.
4. Dwyer T., Sallis JF, Blizzard L, Lazarus R, Dean K. Relation of academic performance to physical activity in children. ***Ped Exerc Sci*** 2001;13:225-237.
5. Keller S. An historical review. ***J Phys Ed Recr Dance*** 1982;53(9):26-28.
6. Ismail A, Gruber J. *The predictive power of coordination and balance items in estimating intellectual achievement*. Proceedings of the first International Congress of Psychology of Sports, Rome, 1965.
7. Keller S. *Endurance fitness training and the elementary school child: Effects on physical and psychological well-being*. Unpublished doctoral dissertation, Concordia University, Montreal, 1981.
8. Shepard RJ. Curricular physical activity and academic performance. ***Ped Exerc Sci*** 1997;9:113-126.
9. Shepard LA, Flexer RJ, Hiebert EH, Marion SF, Mayfield V, Weston TJ. (Effects of introducing classroom performance assessments on student learning. ***Education Measurement: Issues and Practices*** 1996;15(3):7-18.
10. California Law (2003). Education Code, Section 60800. Web site: http://www.leginfo.ca.gov/calaw.html.
11. California Law (2003). Education Code, Section 60640. Web site: http://www.leginfo.ca.gov/calaw.html.
12. Adler N, Cohen S, Cullen M, McEwen B, Paxson C, Scwartz J. (2002). MacAurthur network on health and SES. Web site: http://www.macses.ucsf.edu.
13. Borysenko J. *Minding the body, mending the mind*. Reading, MA: Addison-Wesley, 1987.
14. Herrenkohl EC, Herrenkohl RC, Rupert LJ, Egolf BP, Lutz JG. Risk factors for behavioral dysfunction: The relative impact of maltreatment, SES, physical health problems, cognitive ability, and quality of parent-child interaction. ***Child Abuse & Neglect*** 1995;9(2):191-203.
15. Evans G W. The environment of childhood poverty. ***American Psychologist***, 2004;59(2):77-92.
16. Weiss P. The living system: Determinism stratified. In A. Koestler & J.R. Smythies (Eds.), *Beyond reductionism: New perspectives in the life sciences* (pp. 3-42), Boston: Beacon Press, 1969.
17. Gollin ES. Development and plasticity. In E.S. Gollin (Ed.), *Developmental plasticity: Behavioral and biological aspects of variations in development* (pp. 231-249). New York: Academic Press, 1981.
18. Schneirla TC. The concept of development in comparative psychology. In D.B. Harris (Ed.), *The concept of development* (pp. 78-108). Minneapolis: University of Minnesota Press, 1957.
19. Gajdusek DC. Physiological and psychological characteristics of stone-age man. ***Engineering and Science*** 1970;33:26-33, 56-62.
20. Caine RN. Caine J. *Teaching and the human brain*. Alexandria, VA: Association for Supervision and Curriculum Development, 1991.
21. Smith ML, Glass GV. *Research and evaluation in education and the social sciences* p.198. Needham Heights, MA: Allyn and Bacon. 1987.