California
Physical Fitness Test

A Study of the Relationship Between Physical Fitness and Academic Achievement in California Using 2004 Test Results

April 2005

Prepared by the
California Department of Education

Table of Contents

Introduction

Little research has examined the relationship between physical fitness and academic achievement. Even so, available research suggests a positive relationship between physical fitness and academic achievement (Grissom 2004). That is, as one measure improves, so does the other. There is no evidence that this relationship is causal. That is, there is no evidence that improving physical fitness causes academic achievement to improve or vice versa.

This study does not address causality. It validates previous correlational research by documenting the relationship between fitness and achievement and the strength of this relationship.

Data from the 2004 Physical Fitness Test (PFT*)* and the *California Standards Tests (CSTs)* were used to study the relationship between fitness and academic achievement. PFT data were scores from the *Fitnessgram.* The *Fitnessgram* is the designated test in California to determine a student's level of physical fitness and was administered during the months of February, March, April, and May to fifth-, seventh-, and ninth-grade public school students. The *CST* scores were measures of academic achievement in English–language arts, mathematics, history–social science, and science. The *CSTs* were administered in spring 2004 to students in the second grade through the eleventh grade in California public schools.

Method

Student demographic information, such as birth date and gender, was collected as part of both the PFT and *CSTs*. The demographic information from the two testing programs was used to create matched files. Specifically, students' county/district/school codes, grade levels, birth dates, genders, and ethnicities were used to match student records from each testing program. The result was that each student record for which there was a match had individual scores for both the *Fitnessgram* and the *CSTs*. As such, these data could be used to compare scores on the PFT with scores on the *CSTs*. The matched student records file will be referred to as the matched cohort sample.

The PFT measured six aspects of fitness: (1) aerobic capacity, (2) body composition, (3) abdominal strength, (4) trunk strength, (5) upper body strength, and (6) flexibility. Student performance was classified at two levels: (1) in the healthy fitness zone, which means students met or exceeded the fitness target, or (2) needs improvement, which means students failed to meet the fitness target. Overall PFT scores ranged from zero, none of the standards were met, to six, all standards were met or exceeded. Including only students with complete data ensured comparability of overall PFT scores.[1]

---

[1]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 were 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.

Analyses first calculated the mean scale scores for the *CST* in English–language arts and the *CST* in mathematics for each overall PFT score. Second, analysis of variance (ANOVA) and linear regression were used to test the statistical significance of the relationship between the overall PFT and achievement scores.

Results

Figure 1 shows the mean scale scores on the 2004 *CST* in English–language arts by the number of fitness standards achieved (i.e., the overall PFT score).
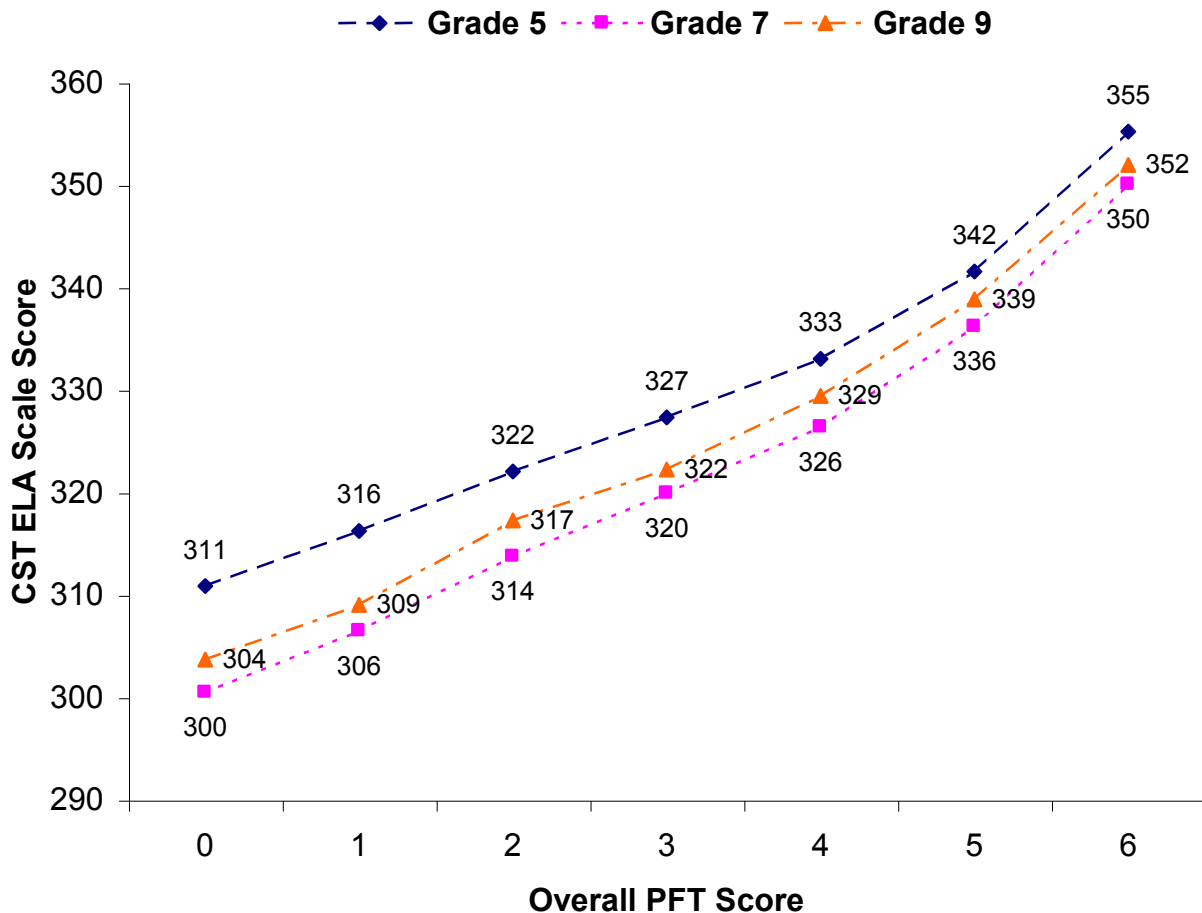

Figure 1. 2004 *CST* in English–language arts mean scale scores for grades 5, 7, and 9 by overall PFT scores. In grade 5 there were 371,198 students, grade 7 had 366,278 students, and grade 9 had 298,910 students.

As the overall PFT score improved, the mean scale score on the *CST* in English–language arts also improved. The average scale score on the *CST* in English–language arts for fifth-grade students who did not achieve any of the fitness standards was 311. The same scale score for seventh and ninth graders was 300 and 304, respectively. The average scale score on the *CST* in English–language arts for fifth-grade students who achieved all six fitness standards was 355. The same scale score for seventh and ninth graders was 350 and 352, respectively. The change in average scale scores on the *CST* in English–language arts from those who achieved none of the fitness standards to those who achieved all six was around 50 points. Results indicate a

positive relationship between academic achievement and physical fitness. As one measure improved, so did the other.

Figure 2 shows these same results using 2004 *CST* in mathematics scale scores.
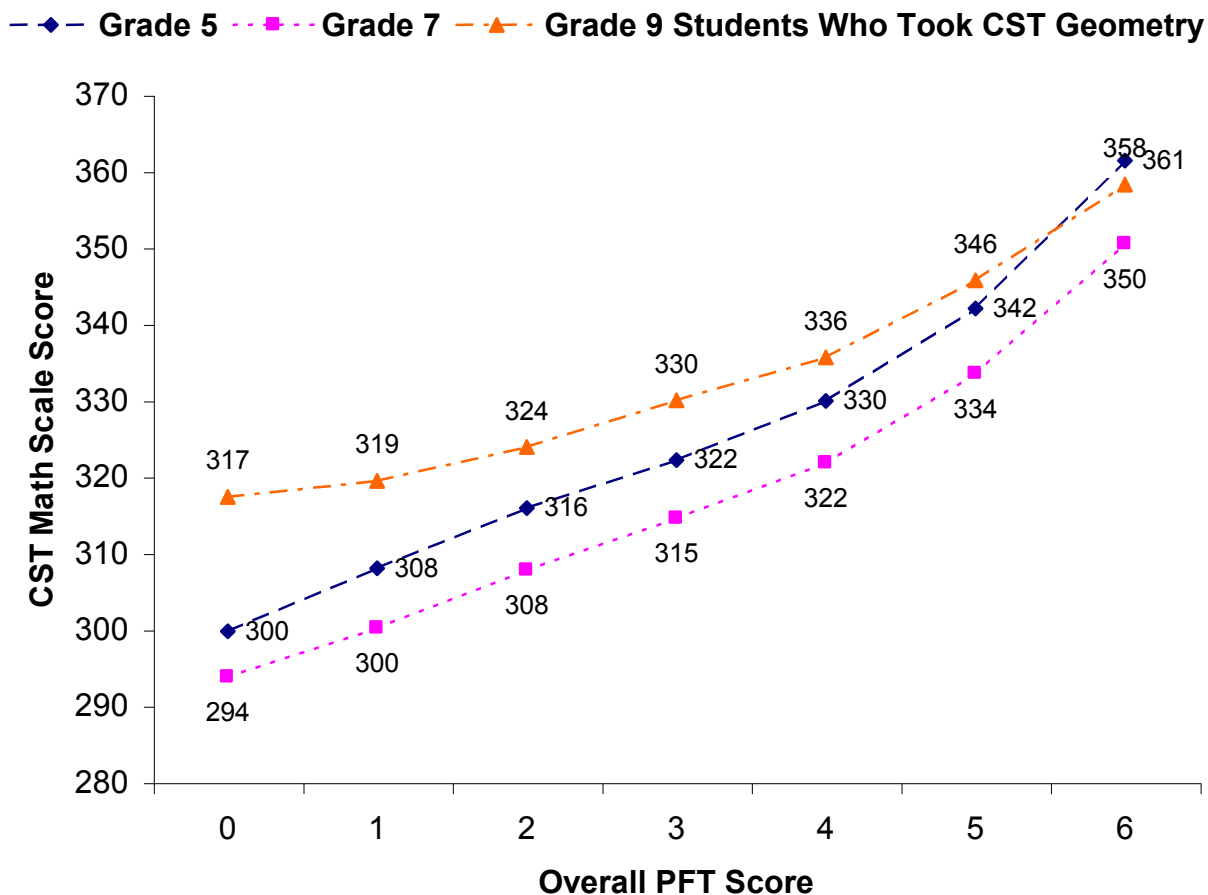

Figure 2. 2004 *CST* in mathematics mean scale scores by overall PFT scores for grades 5, 7, and 9. The numbers of students in grades 5 and 7 were the same as those in Figure 1. The number of grade 9 geometry students was 63,028.

Mathematics results were consistent with English–language arts results. That is, as the overall PFT score improved, the mean scale score on the *CST* in mathematics also improved. Results indicate that even when the measure of academic achievement changed, a positive relationship between academic achievement and physical fitness remained.

Although evidence suggests a relationship between physical fitness and academic achievement, analysis of variance (ANOVA) and linear regression were used to test the relationship for statistical significance. The results from ANOVA and linear regression were statistically significant. (See Appendix A.)

Determining whether the relationship between physical fitness and academic achievement was affected by student characteristics required subgroup analyses. First, the relationship between physical fitness and academic achievement was examined by gender. Figure 3 shows the relationship between 2004 overall PFT scores and *CST* in English–language arts mean scale scores by gender for fifth graders.
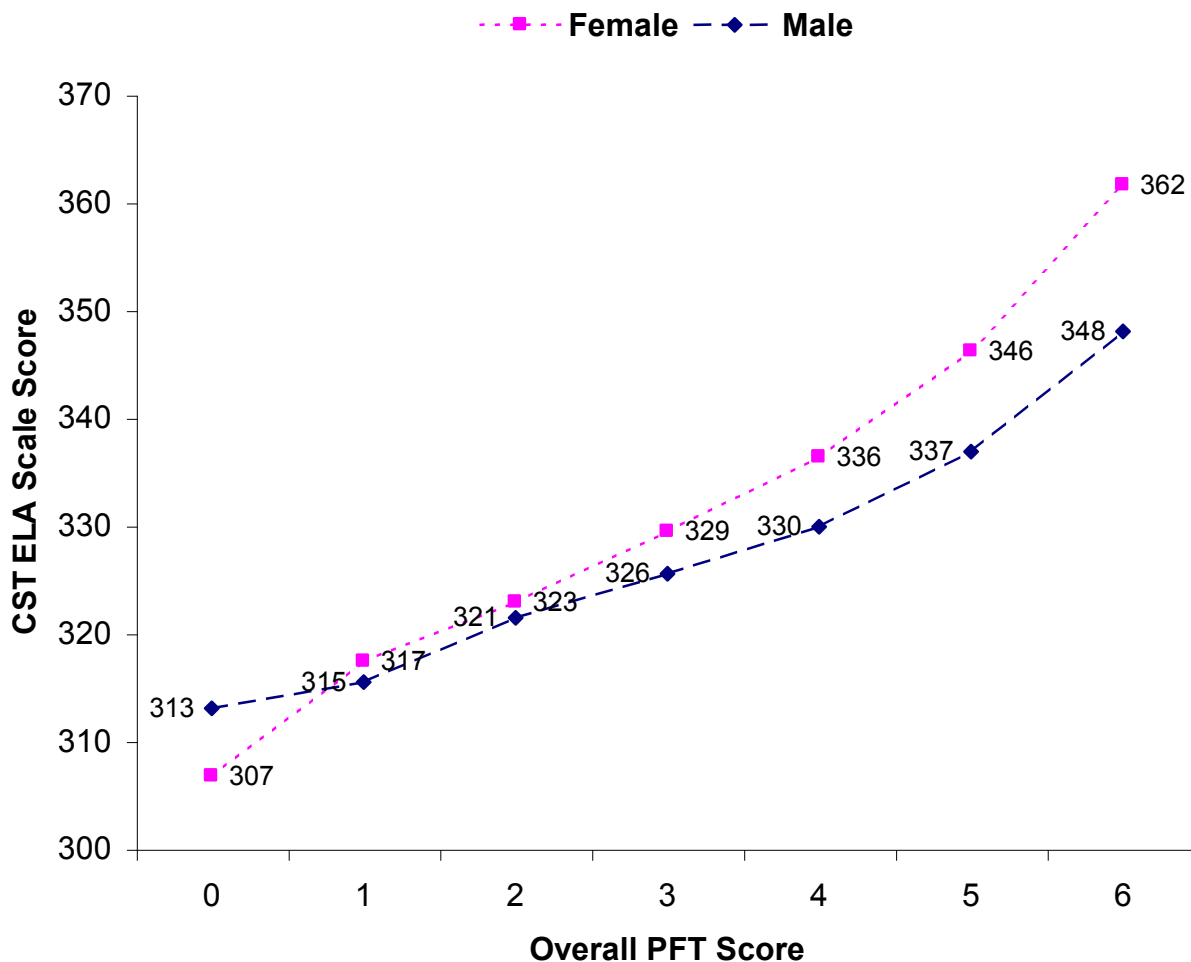

Figure 3. 2004 C*ST* in English–language arts mean scale score for grade 5 by overall PFT score and gender. The number of female and male students was 182,287 and 188,921, respectively.

Figure 3 shows that the relationship between fitness and achievement was consistent across genders. For females and males, as the overall PFT score increased, so did the mean *CST* in English–language arts scale scores.

Even though the relationship between fitness and achievement was consistent across genders, the rate of change in achievement scores was greater for females than for males. The difference in mean scale scores on the *CST* in English–language arts between the least and most fit students was 55 points for females and 35 points for males. Results using mathematics scores were consistent with those using English–language arts scores, and the results for seventh- and ninth-grade students were consistent with those for fifth graders.

Next, the relationship between physical fitness and academic achievement was examined by socioeconomic status (SES). The National School Lunch Program (NSLP) served as a proxy for SES. Whether students received free or reduced lunch was an indicator of being economically disadvantaged or lower SES. Non-NSLP participation was an indicator of not being economically disadvantaged or higher SES. Figure 4 shows these results for fifth graders.
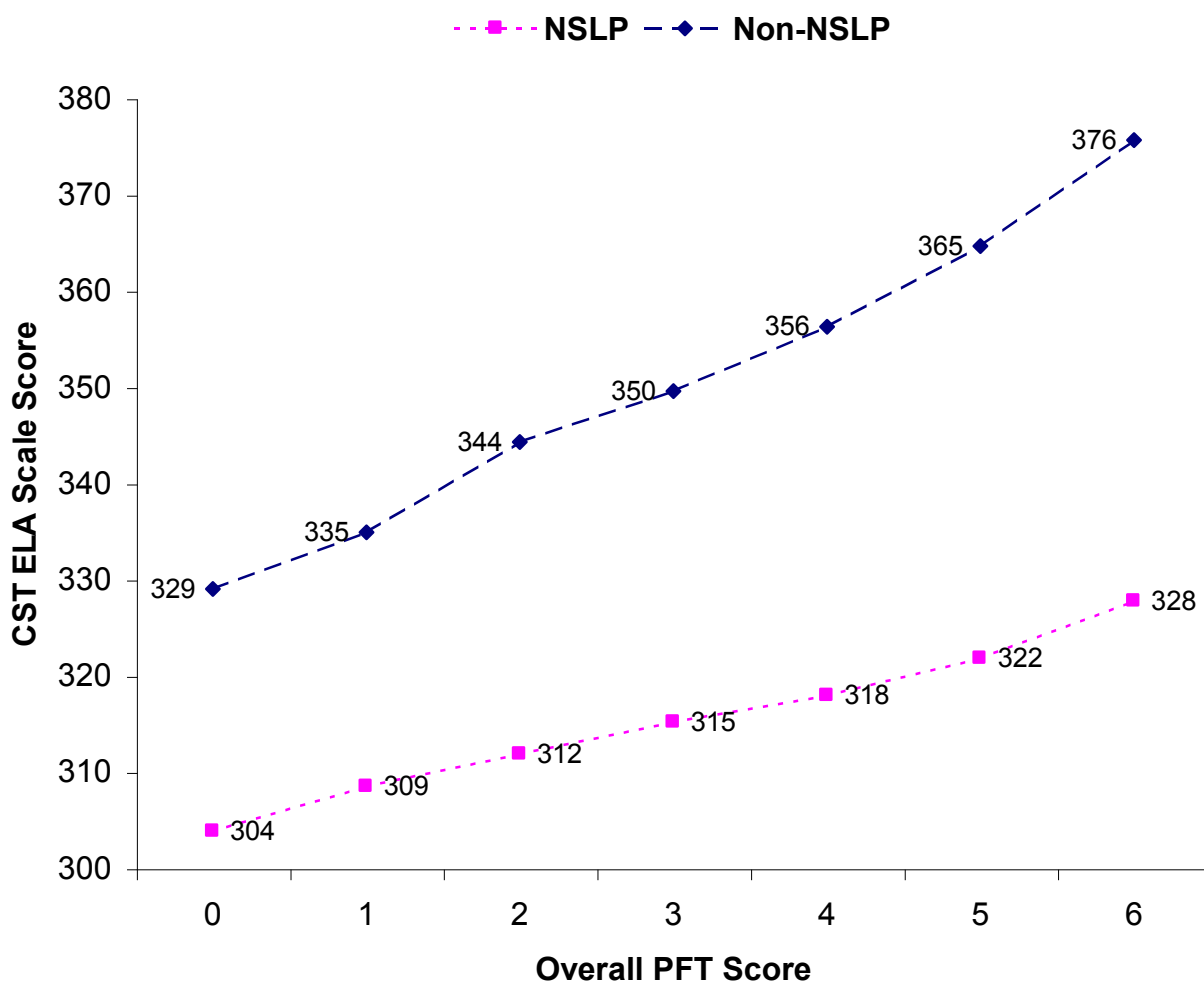
Figure 4. 2004 *CST* in English–language arts mean scale score for grade 5 by overall PFT score and NSLP. The number of NSLP and non-NSLP students in grade 5 was 203,726 and 167,472 respectively.

Figure 4 indicates that as the PFT score increased, so did the mean academic achievement for both non-NSLP and NSLP students. However, the rate of change in achievement scores was greater for non-NSLP students than for NSLP students. The difference in mean scale scores on the *CST* in English–language arts between the least and most fit students was 47 points for non-NSLP and 24 points for NSLP. Results using mathematics scores were consistent with those using English–language arts

scores, and results for seventh- and ninth-grade students were consistent with those for fifth graders.

Although evidence suggests the relationship between physical fitness and academic achievement was different for females and males and different for higher SES and lower SES students, ANOVA was used to test whether these differences were statistically significant. ANOVA results indicated that the rate of increase in mean achievement scores by PFT scores was significantly greater for females than for males and significantly greater for students not receiving NSLP (i.e., higher SES) than for students receiving NSLP (i.e., lower SES). (See Appendix B.)

Discussion

Results validate earlier studies. There was a strong positive relationship between physical fitness and academic achievement. The relationship between fitness and achievement was stronger for females than for males and stronger for higher SES students than for lower SES students.

However, neither ANOVA nor regression results indicate causality. That is, it cannot be inferred from these data that improved physical fitness caused an increase or improvement in academic achievement or vice versa.

However, it is possible that better general health and/or better living conditions were responsible for both higher fitness levels and higher levels of academic achievement. Previous research has shown that as SES improves, so does overall health (Evans 2004) and as SES improves, so does academic achievement (Herrenkohl, Herrenkohl, Rupert, Egolf, & Lutz 1995). Results from this study indicate a stronger relationship between fitness and achievement for higher SES students. This cumulative evidence indicates that conditions that improve general health promote both a healthy body and improved intellectual capacity.

Appendix A

ANOVA results for fifth–grade students when the *CST* in English–language arts scale scores are the dependent variable and overall PFT scores are the independent variable.

Source	df	SS	MS	F	Pr > F
Regression	6	50979850	8496642	3090.0	< .0001
Residual	368172	1012365755	2750		
Total	368178	1063345605			

ANOVA results and multiple comparison tests indicated that students who had an overall PFT score of six had a mean scale score on both the *CST* in English–language arts and mathematics that was significantly greater than students who had an overall PFT score of five or less. For example, fifth-grade students who had an overall PFT score of six had a mean scale score of 355 on the *CST* in English–language arts. The English–language arts mean score of 355 was statistically greater than the mean English–language arts score of 342, which was the mean scale score on the *CST* in English–language arts for students who had an overall PFT score of five. Students who had an overall PFT score of five had a mean *CST* scale score that was statistically greater than students who had an overall PFT score of four or less and so on. There was one exception to this pattern: Students who had a PFT score of one tended to have a *CST* mean scale score that was not statistically greater than students who had a PFT score of zero.

Results from linear regression analyses showed that as overall PFT scores increased, *CST* scores also increased in a statistically significant pattern. The regression equation indicated that for every unit increase in overall PFT score, the scale score on the *CST* in English–language arts increased eight points for fifth graders, ten points for seventh graders, and nine points for ninth graders.

Appendix B

ANOVA results for fifth-grade students when the *CST* in English–language arts scale scores are the dependent variable and overall PFT scores, gender, and NSLP are the independent variables.

Source	df	SS	MS	F	Pr > F
Regression	20	213450342	10672517	4623.1	< .0001
Residual	368158	849895263	2309		
Total	368178	1063345605			
Overall **PFT** *Score*	6	21986377.2	3664396.2	1587.3	< .0001
Gender	1	310443.9	310443.9	134.5	< .0001
NSLP	1	10978627.9	10978627.9	4755.7	< .0001
Gender*PFT Score	6	943467.5	157244.6	68.1	< .0001
NSLP*PFT Score	6	2608492.8	434748.8	188.3	< .0001
Contrast	**df**	**SS**	**MS**	**F**	**Pr > F**
Six compared to Scores Less than Six	1	8808564.4	8808564.4	3815.7	< .0001
Five compared to Scores Less than Five	1	3589235.8	3589235.8	1554.8	< .0001
Four compared to Scores Less than Four	1	1432484.7	1432484.7	620.5	< .0001
Three compared to Scores Less than Three	1	582932.3	582932.3	252.5	< .0001
Two compared to Scores Less than Two	1	212324.2	212324.2	92.0	< .0001
One compared to Zero	1	27469.8	27469.8	11.9	0.0006

References

Grissom, J. B. (2004). Physical fitness and academic achievement. Unpublished manuscript.

Evans, G. W. (2004). The environment of childhood poverty. *American Psychologist*, 59 (2), 77–92.

Herrenkohl, E. C., Herrenkohl, R. C., Rupert, L. J., Egolf, B. P., & Lutz, J. G. (1995). 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*, 19 (2), 191–203.