ORIGINAL ARTICLES

Low Aerobic Fitness and Obesity Are Associated with Lower Standardized Test Scores in Children

Christian K. Roberts, PhD, Benjamin Freed, BA, and William J. McCarthy, PhD

Objective To investigate whether aerobic fitness and obesity in school children are associated with standardized test performance.

Study design Ethnically diverse (n = 1989) 5th, 7th, and 9th graders attending California schools comprised the sample. Aerobic fitness was determined by a 1-mile run/walk test; body mass index (BMI) was obtained from state-mandated measurements. California standardized test scores were obtained from the school district.

Results Students whose mile run/walk times exceeded California Fitnessgram standards or whose BMI exceeded Centers for Disease Control sex- and age-specific body weight standards scored lower on California standardized math, reading, and language tests than students with desirable BMI status or fitness level, even after controlling for parent education among other covariates. Ethnic differences in standardized test scores were consistent with ethnic differences in obesity status and aerobic fitness. BMI-for-age was no longer a significant multivariate predictor when covariates included fitness level.

Conclusions Low aerobic fitness is common among youth and varies among ethnic groups, and aerobic fitness level predicts performance on standardized tests across ethnic groups. More research is needed to uncover the physiological mechanisms by which aerobic fitness may contribute to performance on standardized academic tests. *(J Pediatr 2010;156:711-8)*.

See editorial, p 696

Schools have been ambivalent about addressing student obesity and lack of physical fitness because these health conditions are thought to be only tangentially related to academic achievement. Optimizing student academic achievement has typically been seen as a primary goal for school boards. The suggestion that physical activity and other lifestyle behaviors may affect brain functions such as learning, memory, and decision-making is largely untested. Evidence is beginning to emerge, however, suggesting that childhood obesity and fitness may influence learning and measured academic performance.[1,2] In counseling parents about consequences of their child's weight status, it may be helpful for pediatricians to be able to address the evidence for a possible link between academic achievement and a child's body weight. Moreover, the lack of opportunities for students to engage in physical activity in the school system and the lack of measures of fitness as vital sign in pediatric medicine may contribute to this ambivalence and lack of translation of knowledge about childhood health from physician to parent.

An initial step in this process is to investigate associations between objective indices of learning and fitness and/or obesity. Along these lines, Datar et al[3] suggested that 1st grade children's standardized test scores may be associated with obesity; however, their results became statistically nonsignificant after including socioeconomic and behavioral characteristics as covariates. The study did not include an objective measure of physical fitness.[3] The present study investigated relationships between state-mandated measures of aerobic fitness, body mass index (BMI), and academic performance (ie, standardized test scores) obtained from an ethnically diverse sample of elementary, middle, and high school children from a Southern California school district.

Methods

Data were collected from 2703 youth enrolled in public schools, including 10 elementary, 2 middle, and 2 high schools during the spring. These students were evaluated as part of the statewide mandated physical performance testing (Assembly Bill 265, Education Code Section 2, Chapter 6, Section 60800). By law, California public school districts must assess all 5th, 7th, and 9th graders annually for physical fitness and body weight. To facilitate analyses, some students

BMI	Body mass index (kg/m^2)
CAT6	California Achievement Test, version 6
CDC	Centers for Disease Control
CST	California Standards Test

From the Department of Physiological Science (C.K.R.) and the Department of Epidemiology (C.K.R.), University of California, Los Angeles, CA; UCLA Center for Metabolic Disease Prevention (C.K.R.); the School of Dentistry (B.F.), University of California, Los Angeles, CA; WestEd (W.J.M.), Los Alamitos, CA; and the Department of Health Services (W.J.M.), School of Public Health and Department of Psychology, University of California Los Angeles, CA

Supported in part by a grant from the National Institute of Diabetes, Digestive and Kidney Disorders (1R01-DK063507). The sponsor had no role in the design of the study, in the collection, analysis, or interpretation of the data, or in the writing or submission of this paper. The authors declare no conflicts of interest.

 10.1016/j.jpeds.2009.11.039

otherwise eligible for inclusion were excluded from analyses because they were missing data on one or more of the study measures (n = 707). The study participants were 749 5th, 761 7th, and 479 9th graders, comprising 1012 boys and 977 girls attending a middle-to-high income Southern California school district in 2002 to 2003. The demographic characteristics of the 1989 students included in the analyses resembled the 2696 who were eligible (Table I). Fifth graders comprised 37.7% of the analytic sample but only 31% of the population, indicating some over-representation of 5th graders. Ninth graders comprised 24.1% of the analytic sample but 32.5% of the population, indicating some under-representation of 9th graders. Departures from representativeness involving ethnicity and sex were negligible, with differences in proportions of ethnic composition between the analytic sample and the population differing by less than 1% for every major ethnic group, except for African Americans, in whom the difference was 1.4%. The sex composition of the analytic sample was identical to the sex composition of the population. Ethnicity was categorized as African American, Asian/Pacific Islander (including Filipinos, Asians, and Pacific Islanders), Hispanic, and non-Hispanic whites.

Aerobic fitness, body weight, and student demographic data were obtained from existing school records with personal identifiers removed except an arbitrary district identification number used to link Fitnessgram data, school district demographic data, and standardized test score data. Parental education, child ethnicity, and eligibility for free or reduced-price lunch status were determined by parent self-report information collected by the school district. These data were taken from the California Department of Education website (http://www.ed-data.k12.ca.us). Missing data included refusals to provide such information. The Institutional Review Board at the University of California, Los Angeles, approved the study protocol.

"Fitnessgram" refers to a comprehensive battery of physical fitness assessments devised by the Cooper Institute for Aerobics Research[4] to assess a student's overall health-related physical fitness. The Fitnessgram has been adopted by the State of California as the required physical performance test to be administered annually to all students in the 5th, 7th, and 9th grades. Physical education staff of each school district must follow strictly the measurement protocol stipulated in the Fitnessgram manual for the assessment of body composition and physical fitness.

All testing took place within the school environment and was administered by the physical education staff. For the purposes of this study, only aerobic fitness assessments of the Fitnessgram were included. Aerobic capacity was measured through use of the state-approved Fitnessgram assessment, using the mile run test. This approach has been validated as a field measure estimate of maximal oxygen uptake ($\cdot VO_{2\ max}$) in both adults and children.[5] For the Fitnessgram aerobic fitness assessment, groups of students ran around a flat quarter-mile track 4 times, and the mile time was recorded up to 15 minutes of run/walk time. Students who had not completed the mile in 15 minutes were assigned the maximum time of 15 minutes. Sex- and age-specific state standards were established for the mile run/walk times that students must achieve to qualify as falling within the "Healthy Fitness Zone."[6]

Height and body weight measurements were taken using a stadiometer and balance beam scale (Detecto, Webb City, Missouri). The school district recalibrates the scales regularly to ensure accurate measures. BMI was calculated as weight (kg)/height (m^2). For the calculation of sex-specific BMI-

Table I. Descriptive statistics for analytical sample after listwise deletion, for eligible students dropped because of listwise deletion, and for all 5th, 7th, and 9th graders in the district

	Analytical sample after listwise deletion	Analytical sample after listwise deletion		Dropped from analytical sample	District enrollment statistics
	Total, n, (%)	Male, n, (%)	Female, n, (%)	Total n, (%)	Total n, (%)
	1989 (100%)	1,012 (50.9%)	977 (49.1%)	707 (100%)	3079 (100%)
Sex, n, (%)					
Male	1012 (50.9%)	——	——	358 (50.1%)	1567 (50.9%)
Female	977 (49.1%)	——	——	356 (49.9%)	1512 (49.1%)
Grade level					
5th	749 (37.7%)	385 (37.9%)	364 (37.1%)	176 (24.6%)	953 (31.0%)
7th	761 (38.3%)	400 (39.3%)	371 (37.8%)	282 (39.5%)	1125 (36.5%)
9th	479 (24.1%)	232 (22.8%)	246 (25.1%)	256 (35.8%)	1001 (32.5%)
Ethnicity					
Asian/Pacific Islander	138 (6.9%)	61 (6.0%)	78 (8.0%)	30 (4.2%)	184 (6.0%)
African American	131 (6.6%)	75 (7.4%)	57 (5.8%)	69 (14.5%)	247 (8.0%)
Hispanic	524 (26.3%)	267 (26.3%)	262 (26.7%)	205 (29.0%)	814 (26.4%)
White	1196 (60.1%)	614 (60.4%)	584 (59.5%)	403 (41.9%)	1830 (59.4%)
Free and reduced price lunch eligibility					
Free lunch eligible	355 (17.8%)	178 (17.5%)	177 (18.0%)	103 (14.7%)	531 (20.9%)*
Reduced-price lunch eligible	125 (6.3%)	71 (7.0%)	54 (5.5%)	32 (4.6%)	157 (6.2%)*
Full-price lunch fare	1518 (76.0%)	768 (75.5%)	750 (76.5%)	568 (80.8%)	1857 (73.0%)*

NOTE. Number of participants by sex, ethnicity and grade, and school lunch subsidy status.
*Based on middle school statistics because grade-specific data were not available.

for-age percentiles, LMS parameters were provided by the CDC[7] and were used to generate *Z*-scores of BMI values that were then applied to estimate sex-specific BMI-for-age percentiles. Obesity risk classification was determined for each student. We opted to use the CDC weight status cutpoints,[8,9] in which the 85th to 94th percentile category is termed "overweight" and the 95th+ percentile is termed "obese." Recent policy statements now speak of the "obese child" rather than limiting the term "obese" to adults.[10] Hence, BMI percentile categories included <5th, ≥5th to 84th, ≥85th to 94th, and ≥95th, which correspond to the following classifications: "underweight," "desirable weight," "overweight," and "obese," respectively.

Study participants (n = 412) had also participated, 4 months earlier, in a health promotion program that also included assessment of BMI. The personnel collecting these measures were nurses hired and trained to collect anthropometric and medical data. This prior assessment of BMI permitted the investigators to gauge the reliability of BMI assessment in children over 4 months. After controlling for minor differences in ethnicity and socioeconomic characteristics between the Fitnessgram sample and the health promotion subsample, the partial correlation between the 2 BMI measures was 0.93, which indicated a reasonably high level of repeatability over 4 months in a pediatric population experiencing natural growth in body weight.

California Department of Education school-level data standardized test score data from the California Achievement Tests version 6 (CAT6) and California Standards Tests (CST) were obtained from the district in both 2002 and 2003 for math and reading (CAT) or math and language (CST). CAT6 test scores are used to compare California students with those in other states and are expressed in percentiles ranging from 0% to 99%. The CST was used to categorize students as "far below basic," "below basic," "basic," "proficient," or "advanced." The CST scores are specific to California content standards; the resulting categorizations are expressed in a range from 1 to 5. By law, all students attending public schools in California were required to complete the CAT6 and are required to complete the CST by May of each year. These tests assess grade-appropriate achievement in math, reading, and language arts and other disciplines. More information is available at: http://star.cde.ca.gov/star2004/aboutSTAR_programbg.asp.

Descriptive statistics are presented as means, standard deviations, and percentages. BMI-for-age *z*-scores and percentiles were created using the Centers for Disease Control (CDC) growth chart–derived norms for sex and age.[11] Hierarchical linear regression models were estimated using maximum likelihood, regressing 2002 standardized test score performance onto BMI-for-age *z*-scores and /or mile run/walk times. The initial model consisted of a null model. With the inclusion of additional predictors, subsequent models controlled for the potentially confounding effects of income (reflected by student eligibility for free and/or reduced price school lunches), sex and ethnicity, the last of which was computed using dummy variables and treating non-Hispanic whites as the referent ethnic group. Hierarchical linear modeling was used to account for students being clustered in schools. Because some study measures had unacceptably skewed or kurtotic distributions, study participants were categorized into quintiles for mile time or BMI-for-age and test score performance. However, we did run the analysis on raw data, and no differences in the relationships were noted. We conducted tests of linear trend by treating the quintile categories as continuous variables and assigning the median score to each category in unconditional regressions.[11] For analysis purposes, only the major ethnic groups were compared; the 7 students identified as American Indian were dropped from the analyses. All data were analyzed using the STATA 10.0 Statistical software package (College Station, Texas).

Results

Table I describes characteristics of the sample by sex, grade, and ethnicity. The sample included a similar number of female and male subjects. The ethnic distribution of the children was 59% non-Hispanic white, 27% Hispanic, 7% African American, and 6% Asian/Pacific Islander. To investigate whether the Fitnessgram participants were demographically representative of the district, the distribution of the sample by ethnicity was compared with the ethnic distribution of the district. The distributions did not differ significantly by ethnicity ($P > .15$) but did vary by free and reduced price meal eligibility ($P = .03$), suggesting that the Fitnessgram participants were ethnically representative of the students comprising the school district but slightly better off, economically, than the school district as a whole (Table I). The frequency of parental completion of college was 61.5%, and only 22.8% of the children were eligible for free or reduced-price school lunches, indicating a school district with higher socioeconomic status than the average California public school district. The average California school district in 2002 to 2003 had 48.3% of students eligible for free or reduced price school lunches (obtained from http://www.cde.ca.gov/ds/sh/cw/filesafdc.asp).

Table II illustrates the mile time by healthy fitness zone standards, as established by the state of California.[6] Physical education instructors were instructed to stop the Fitnessgram mile run/walk test at 15 minutes; some physical education instructors permitted students completing the final quarter mile to complete it in 16 minutes. Observed means are therefore underestimates, given that some students who were unable to complete a mile because of the time constraint would have completed it in more time if permitted. Sixty-five percent of students had a fitness level below recommended age-specific, sex-specific standards for mile time performance. Additionally, 64% of study participants had mile times that were slower than the norms recommended by the state of California. There was a similar percentage of girls and boys classified as being in the healthy fitness zone ($P = .48$). Table II also depicts the relationship between ethnicity and achievement of California age-specific mile time

Table II. Fitness and BMI-for-age classifications by sex and ethnicity

Healthy fitness zone status†	Exceeded state aerobic fitness standards	Met state aerobic fitness standards	Did not meet state aerobic fitness standards	Mean mile time Mean (min) [95% CI]
Sex				
Male, n (%)	95 (6.9)	373 (27.2)	902 (65.8)	9.72 [9.02-10.43][a]
Female, n (%)	67 (5.5)	379 (31.1)	774 (63.4)	10.98 [10.56-11.39][a]
Ethnicity				
Asian/Pacific Islander, n (%)	12 (7.2)	58 (34.9)	96 (57.8)	10.18 [9.57-10.79][a]
Hispanic, n (%)	39 (5.6)	175 (25.3)	477 (69.0)	10.69 [9.94-11.45][a,b]
African American, n (%)*	6 (3.3)	40 (21.5)	140 (75.3)	10.81 [10.28-11.35][b]
White, n (%)	105 (6.8)	476 (30.9)	959 (62.3)	10.16 [9.65-10.68][a]
Total, n (%)	162 (6.3)	752 (29.0)	1676 (64.7)	10.34 [9.81-10.87]

BMI-for-age‡	Underweight	Desirable weight	Overweight	Obese	Mean BMI percentile Mean [95% CI]
Sex					
Male, n, (%)	24 (2.0)	779 (66.2)	191 (16.2)	183(15.6)	64.1 [60.6-67.5][a]
Female, n (%)	20 (1.8)	799 (70.5)	184 (16.2)	130 (11.5)	63.0 [58.0-68.0][a]
Ethnicity					
Asian/Pacific Islander, n (%)	5 (3.3)	123 (80.4)	19 (12.4)	6 (3.9)	52.8 [47.6-57.9][a]
Hispanic, n (%)	3 (0.5)	316 (51.8)	127 (20.9)	164 (26.8)	74.4 [71.3-77.4][c]
African American, n (%)	5 (3.2)	88 (55.7)	32 (20.3)	32 (20.3)	70.1 [64.3-76.0][c]
White, n (%)	31 (2.4)	1047 (75.7)	196 (14.2)	109 (7.9)	59.1 [56.8-61.5][b]
Total, n (%)	44 (1.9)	1578 (68.3)	375 (16.2)	313 (13.6)	63.5 [59.6-67.5]

Means with different subscripts (a, b) were significantly different from each other after Bonferroni correction for multiple comparisons, $P < .05$; means with the same subscript were not significantly different.

*African American aerobic fitness differed from that of whites and Asian/Pacific Islanders, $P < .05$.

†Fitness classification based on 2002 State of California fitness standards for sex and age. 2002 mile run/walk time standards are identical to 2008 standards, which are available at: www.cde.ca.gov/ta/tg/pf/documents/healthfitzone08.pdf.

‡BMI-for-age obesity status established using CDC sex- and age-specific percentiles (http://www.cdc.gov/nchs/nhanes/growthcharts/datafiles.htm), with underweight <5th percentile; 5th percentile > desirable weight < 85th percentile; 85th percentile ≤ overweight < 95th percentile, and obese ≥ 95th percentile.

performance. African American students were less likely to achieve California fitness standards than Asian American and non-Hispanic white students ($OR_{Asians} = .49$, $P = .04$; $OR_{whites} = .58$, $P = .005$).

Obesity status and mean BMI-for-age percentiles (converted from *z*-scores for ease of interpretation) are also included in Table II. The combined prevalence of overweight and obese (BMI ≥85th percentile) was 31.8% for boys and 27.7% for girls (Table II). Mean BMI-for-age percentiles varied by ethnicity, with 16.3% of Asians classified as either at risk for overweight or obesity, 22.1% of non-Hispanic whites, 40.6% of African Americans, and 47.7% of Hispanics. The percentages of students classified as overweight or obese was greater among African Americans and Hispanics than among Asians or non-Hispanic whites (all comparisons $P < .003$, after Bonferroni correction). Additionally, Asian/Pacific Islanders had slightly lower BMI-for-age percentiles than non-Hispanic whites ($F(1,11) = 10.65$, $P = .04$). Hispanics and African Americans did not differ significantly from each other, nor did boys' BMI-for-age differ on average from that of girls.

With respect to age-specific fitness standards, students who failed to run the mile in the appropriate time interval established as appropriate for each age and sex scored significantly lower on the CAT6 and CST math, reading, and language California standards tests compared with those students who fell in the healthy fitness zone (Table III). Tests for linear trends revealed that decreasing quintiles of aerobic fitness scored progressively lower on CAT6 math and reading (linear trend, $P_{math} < .0001$; $P_{reading} = .001$) (Figure 1) and on CST math and language tests (linear trend, $P_{math} < .0001$; $P_{language} < .0001$) (Figure 2; available at www.jpeds.com).

Table III depicts test scores by BMI percentiles; as observed for mile run/walk time, those who exceeded both the 85th and 95th percentiles for BMI-for-age scored significantly lower on the CAT6 and CST math, reading, and language tests than those in the recommended range for BMI. Tests for linear trend showed that increasing quintiles of BMI-for-age percentile scores scored progressively lower on both CAT6 math and reading (linear trend, $P_{math} = .007$; $P_{reading} = .028$) (Figure 1) as well as CST math and language tests (linear trend, $P_{math} = .013$; $P_{language} = .073$) (Figure 2).

Sequential hierarchical linear regression models (Stata xtreg procedure) were used to regress CAT6 and CST test score performance measures onto the following predictors: (1) null model; (2) age, ethnicity, sex, and eligibility for free/reduced price school lunches; (3) the foregoing covariates and BMI-for-age *z*-scores; and (4) the foregoing covariates/predictors and mile run/walk time. Mile run/walk time was a significant predictor of standardized CAT6 2002 math test score performance such that the math score dropped 1.9 points (of a possible 99) for every additional minute required to complete the 1-mile run/walk ($b = -1.94$, 95% CI $= -2.37, -1.53$) even when age, free or reduced-price lunch status, sex, and ethnicity and BMI-for-age were included as covariates. Adding the demographic covariates to the null model reduced the intraclass correlation from .09 to .01

Table III. Mean performance on standardized math, reading, and language tests by aerobic fitness classification and by BMI-for-age obesity status

	Underweight	Desirable weight	At risk for overweight	Overweight
CAT6 math 2002†	60.93 (56.57-65.28)	62.39 (59.69-65.08)	56.51 (52.67-60.35)*	52.09 (49.04-55.13)*
CAT6 read 2002	61.64 (56.09-67.19)	60.75 (58.30-63.20)	53.03 (49.64-56.42)*	50.52 (47.30-53.74)*
CAT6 math 2003	64.94 (58.33-71.56)	67.97 (64.41-71.53)	58.37 (52.84-63.89)*	53.56 (48.56-58.56)*
CAT6 read 2003	64.75 (52.90-76.60)	64.52 (60.77-68.27)	54.62 (49.59-59.56)*	50.94 (45.88-55.99)*
CST lang 2002	3.65 (3.34-3.98)	3.84 (3.71-3.96)	3.41 (3.16-3.67)*	3.21 (3.01-3.41)*
CST math 2002	3.56 (3.26-3.87)	3.75 (3.59-3.91)	3.42 (3.21-3.64)*	3.10 (2.92-3.28)*
CST lang 2003	3.79 (3.52-4.05)	3.82 (3.62-4.02)	3.50 (3.23-3.77)*	3.29 (3.11-3.46)*
CST math 2003	3.69 (3.33-4.05)	3.59 (3.33-3.85)	3.29 (2.99-3.58)*	3.02 (2.76-3.29)*

	Exceeded 2003 state aerobic fitness standards	Met 2003 state aerobic fitness standards	Did not meet 2003 state aerobic fitness standards
CAT6 math 2002†	65.72 (59.79-71.65)	64.92 (61.27-68.56)	56.52 (53.47-59.58)*
CAT6 read 2002	60.01 (54.10-65.92)	61.34 (57.74-64.93)	55.46 (52.12-58.79)*
CAT6 math 2003	71.23 (63.28-79.18)	69.81 (64.63-74.98)	58.38 (54.10-62.67)*
CAT6 read 2003	63.80 (55.23-72.36)	65.28 (59.88-70.69)	57.48 (52.75-62.21)*
CST lang 2002	3.83 (3.68-3.97)	3.85 (3.71-3.99)	3.52 (3.31-3.73)*
CST math 2002	3.82 (3.67-3.97)	3.81 (3.62-4.00)	3.44 (3.22-3.66)*
CST lang 2003	3.67 (3.17-4.17)	3.86 (3.58-4.13)	3.57 (3.31-3.82)*
CST math 2003	3.61 (3.11-4.11)	3.65 (3.27-4.03)	3.32 (3.02-3.61)*

NOTE. Means adjusted by the covariates sex, free and reduced-price lunch eligibility, and ethnicity. BMI-for-age obesity status established using CDC sex- and age-specific percentiles (http://www.cdc.gov/nchs/nhanes/growthcharts/datafiles.htm), with underweight <5th percentile; 5th percentile > desirable weight < 85th percentile; 85th percentile ≤ overweight < 95th percentile, and obese ≥ 95th percentile. 2003 mile run/walk time standards are identical to 2008 California State standards for aerobic fitness, which may be found at: http://www.cde.ca.gov/ta/tg/pf/documents/healthfitzone08.pdf. Numbers are means; numbers in parentheses represent the 95% confidence intervals.
*Mean test scores differed significantly from those whose aerobic fitness or body composition met or exceeded state standards at the $P < .05$ level.
†CAT6 scores ranged from 1 (low) to 99 (high); CST scores were categorical, ranging from 1 ("far below basic) to 3 ("basic") to 5 ("advanced").

and explained 15.9% of the null model variance. Adding BMI-for-age to the demographic variables-augmented model decreased the null model variance only an additional 0.4%, although this change was still statistically significant (model difference likelihood ratio χ^2 (1) = 8.6, P = .003). The full model, including all of the foregoing covariates/ predictors (including BMI-for-age) but also including a measure of the student's performance on the 1-mile fitness test decreased the null model variance an additional 3.5% (model difference likelihood ratio χ^2 (1) = 81.6, $P < .0001$). With the inclusion of the fitness measure, the student's BMI-for-age was no longer a significant contributor to the CAT6 2002 math test score. All major ethnic groups differed significantly from whites in the full model, with Asian math scores higher than whites' scores (b_{Asian} = 4.6; 95% CI = 1.44, 7.85), and Hispanic and African American math scores lower than whites' scores, respectively ($b_{Hispanic}$ = −11.44; 95% CI = −13.74, −9.14; $b_{African\ American}$ = −16.33; 95% CI = −19.70, −12.96).

The standardized CAT6 2002 reading test score dropped 1.1 points for every additional minute required to complete the 1-mile run/walk (b = −1.13; 95% CI = −1.56, −0.70) even when age, sex, ethnicity, free and reduced-price lunch status, and BMI-for-age were included as covariates. Adding the demographic covariates to the null model reduced the intraclass correlation from .09 to .01 and explained 15.6% of the null model variance. Adding BMI-for-age to the demographic variables-augmented model decreased the null model variance only an additional 0.7%, although this change was still statistically significant (model difference likelihood ratio χ^2 (1) = 6.23, P = .013). The full model, including all of the foregoing covariates/predictors, including BMI-for-age, but also including a measure of the student's performance on the 1-mile fitness test decreased the null model variance an additional 1.3% (model difference likelihood ratio χ^2 (1) = 26.2, $P < .0001$). With the inclusion of the fitness measure, the student's BMI-for-age was no longer a significant contributor to the CAT6 2002 reading test score. Asian ethnicity generally explained no additional variance in CAT6 reading test scores relative to non-Hispanic whites, which was the referent ethnic group, but Hispanic and African American ethnicity did explain some additional variation (11.2-point drop for Hispanics, $b_{Hispanic}$ = −11.23; 95% CI = −13.57, −8.89; 14.5-point drop for African Americans, $b_{African\ American}$ = −14.53; 95% CI = −17.96, −11.12).

Similar findings were noted for CST math and language test performance (data not shown). These analyses also confirmed that the pattern of ethnic differences in standardized test scores (Asians and non-Hispanic whites > African Americans and Hispanics) was consistent with the pattern of ethnic differences in percent achieving recommended levels of BMI and aerobic fitness.

Discussion

The most impressive findings were the consistency of positive associations between aerobic fitness and standardized test score performance and the consistency of inverse associations between BMI-for-age and standardized test score performance. Even those children who were classified as overweight but not obese (ie, >85th percentile but <95th) scored significantly lower than did desirable weight children. Because

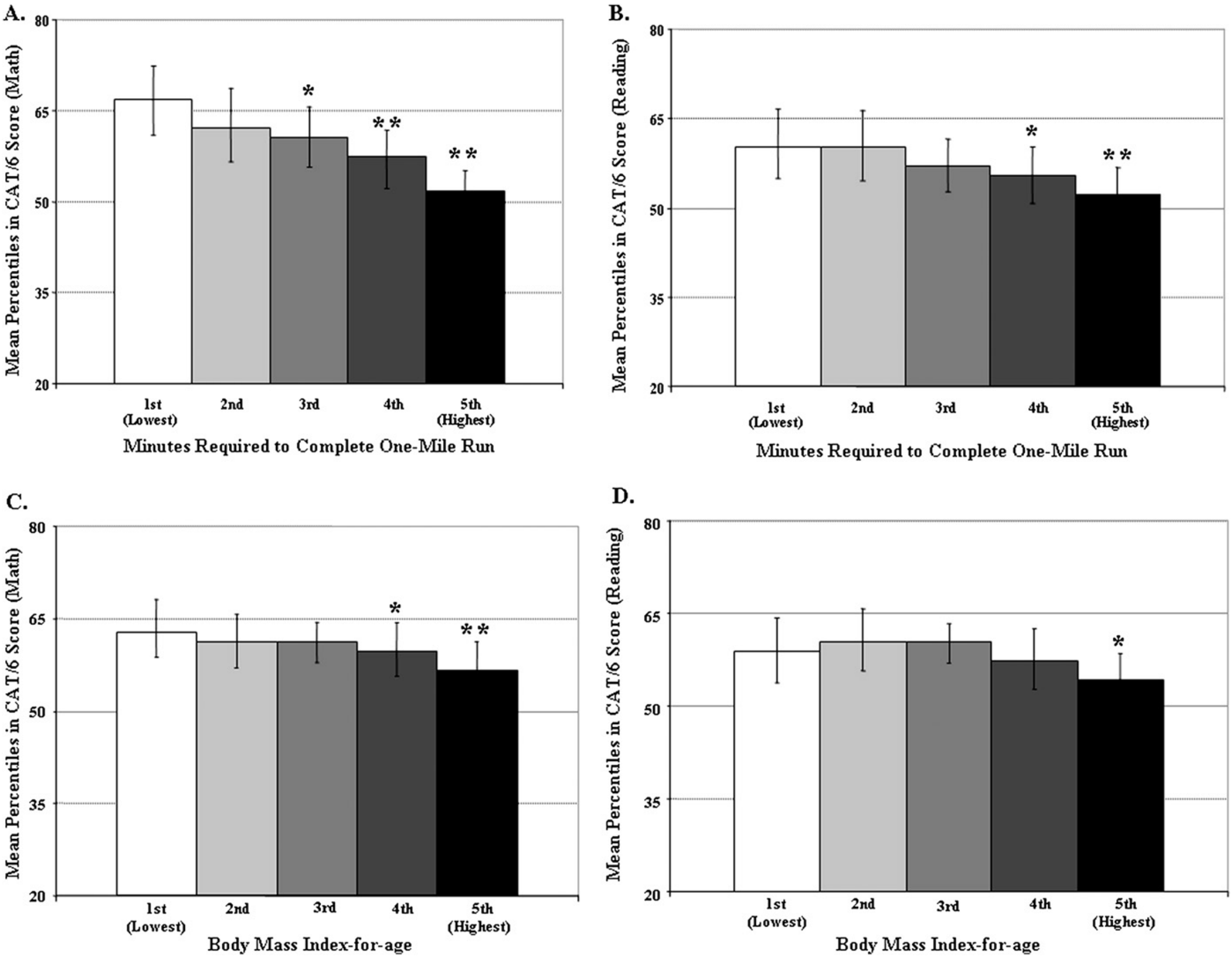

Figure 1. CAT6 mean scores, adjusted by the covariates sex, free and reduced-price lunch eligibility, and ethnicity. **A,** Percentiles in CAT6 score (math) by quintiles of minutes to complete mile run; *3rd quintile of mile time differed on 2002 CAT6 math score from 1st quintile, $z = -3.33$, $P = .001$; **4th quintile of mile time differed on 2002 CAT6 math score from 1st quintile, $z = -5.86$, $P < .0001$; **5th quintile of mile time differed on 2002 CAT6 math score from 1st quintile, $z = -8.41$, $P < .0001$. **B,** Percentile in CAT6 score (language) by quintiles of minutes to complete mile run. *4th quintile of mile time differed on 2002 CAT6 language score from 1st quintile, $z = -2.65$, $P = .008$; **5th quintile of mile time differed on 2002 CAT6 language score from 1st quintile, $z = -5.07$, $P < .0001$. **C,** Percentiles in CAT6 score (math) by BMI quintile; *4th quintile of BMI differed on 2002 CAT6 math score from 1st quintile, $z = -2.40$, $P = .016$; **5th quintile of BMI differed on 2002 CAT6 math score from 1st quintile, $z = -4.02$, $P < .0001$. **D,** Percentiles in CAT6 score (language) by BMI quintile. *5th quintile of BMI differed on 2002 CAT6 language score from 1st quintile, $z = -2.35$, $P = .019$. Error bars are 95% confidence intervals.

decreased socioeconomic status has been consistently associated with decreased standardized test scores,[12] it is an obvious potential confounder of any association between obesity and standardized test performance. Controlling for age, socioeconomic status, sex, and ethnicity did attenuate the significance of the relationships between aerobic fitness and standardized test scores and between BMI-for-age percentiles and standardized test scores. Nevertheless, associations remained significant. It appears that both BMI-for-age and performance on the 1-mile run/walk predict standardized test score performance, above and beyond the large amount of variance predicted by sex, ethnicity, and socioeconomic status, but the remaining variance explained is small.

The findings presented here confirm and extend previous findings that aerobic fitness is associated with enhanced performance on standardized achievement tests. The extensions include generalization to ethnically and socioeconomically diverse students varying across primary and secondary school grades, using objective measures of both aerobic and standardized test score performance. Additionally, the data suggest that the association of obesity status and test score performance may be mediated by fitness.

Our data indicate that the fitness level of nearly two thirds of the students surveyed did not fall in the healthy fitness zone. These data are in agreement with the 2004 California Fitness Test data, in which only 27% of more than 1.3 million students tested in grades 5, 7, and 9 met fitness standards in all 6 examined variables. Decades of declining participation in physical education[13-15] have resulted in decades of declining student fitness levels.[16] Sixty-one percent of children ages 9 to 13 years do not participate in any organized physical activity during their nonschool hours.[17] The result is that one third of adolescents fail to achieve a recommended minimum of 30 minutes of moderate to vigorous physical activity 3 times per week.[18,19]

We also noted ethnic disparities in aerobic fitness, with both Hispanic and African American youth reporting slower mile run/walk performances than whites and Asians. This finding is consistent with that of Beets et al,[20] who reported a similar trend in all California students in 2002. Participation in vigorous activity is higher in whites (67%) compared with blacks (54%) and Hispanics (60%) and decreases with advancing grade. The higher BMI-for-age (**Table II**) presenting in African Americans and Hispanics may also contribute to their lower aerobic fitness levels.

We noted approximately 30% prevalence of overweight/obesity in the current study, with higher prevalence in Hispanic (48%) and African American (41%) students. Our data corroborate prior studies such as the California Children Healthy Eating and Exercise Practices Survey (calCHEEPS), in which 32% of 4th and 5th graders were overweight or obese.[21] When comparing the prevalence of overweight/obesity, our data for non-Hispanic whites and Asians are similar to those reported for all California students in 2002, and our data for Hispanics and African Americans indicated higher prevalence rates than corresponding estimates for the state.[20] This is of interest, given the high socioeconomic status of the cohort as a whole, and agrees with prior data that even in higher socioeconomic strata, obesity risk is increasing.

The mechanism(s) by which students with higher aerobic fitness and/or lower BMI-for-age might perform better on standardized academic achievement tests is unknown. Some studies have suggested that cognitive function may be impaired by obesity,[22] low fitness,[23,24] and metabolic syndrome.[25] Lifestyle behaviors such as everyday physical activity and food choices can affect both aerobic fitness and body weight. There may be links between these lifestyle behaviors and learning and objective academic performance. When adding a daily physical activity program to existing primary school curricula, there was no evidence of any loss of academic performance as measured by arithmetic and reading tests despite a 45- to 60-minute loss of formal teaching time each day.[26-28] Mechanistic studies of cognitive function suggest a positive effect of physical activity on intellectual performance,[29] although the relevance to children is unknown because most studies of mechanisms that might explain how physical activity affects brain function have been performed in adults to date.

There were several limitations in this study. One limitation is the mile run/walk test; the validity of the results assumes that all students exerted the maximum effort possible when completing this assessment. The validity of the results also assumes that the 1-mile run/walk test is a reliable surrogate of gold standard measures of aerobic fitness, a point that is disputed by some.[30,31] Additionally, given that excess adiposity may affect mile time, we cannot unequivocally state that the effect of fitness is independent of obesity status. Strict adherence to test administration and data collection procedures could not be confirmed and could have affected the reliability of the data; however, mile time and BMI assessment are measures that a state-certified physical education instructor would be qualified to perform. BMI is only a surrogate measure of body composition and cannot be used to differentiate between changes in lean and adipose tissue. For example, resistance training can elicit muscle hypertrophy, resulting in greater lean body mass for height. Future research might consider using waist circumference or body composition testing rather than BMI-for-age as a better predictor of obesity-related conditions. Results were limited to grades studied; however, there is reason to believe that what is true of 5th, 7th, and 9th graders would also be true of children in other grades. The cross-sectional nature of the data militates against drawing causal inferences from the observed relationship between variations in fitness and variations in standardized test performance. Longitudinal research is needed to confirm whether changes in body composition or physical fitness over time explain variations in test scores and to shed light on the temporal mechanisms that may explain how physical fitness and body composition influence school children's performance on standardized tests.

The current study suggests that even in a higher socioeconomic status cohort, the prevalence of low fitness and obesity are common. If future studies confirm a causal role for the influence of fitness on academic performance, schools will have to reverse their recent disinvestment in physical education ostensibly for the purpose of boosting student achievement. ■

We are grateful to the Director, Information Services, of the anonymous school district that participated in this study, for his work in transmitting to the investigators the data used in this report.

Submitted for publication Jun 4, 2009; last revision received Aug 31, 2009; accepted Nov 11, 2009.

Reprint requests: Dr William J. McCarthy, University of California, Los Angeles, Division of Cancer Prevention and Control Research, Los Angeles, CA 90095-6900. E-mail: wmccarth@ucla.edu.

References

1. Bogden JF. Fit, healthy, and ready to learn: a school health policy guide. Accessed Nov 16, 2007, at http://www.nasbe.org/HealthySchools/fithealthy.html. Alexandria (VA): NASBE; 2000.
2. Hillman CH, Erickson KI, Kramer AF. Be smart, exercise your heart: exercise effects on brain and cognition. Nat Rev Neurosci 2008;9:58-65.
3. Datar A, Sturm R, Magnabosco JL. Childhood overweight and academic performance: national study of kindergartners and first-graders. Obes Res 2004;12:58-68.
4. Cooper Institute. FitnessGram Test Administration Manual. 2nd ed. Champaign, IL: Human Kinetics; 1999.

5. Cureton KJ, Sloniger MA, O'Bannon JP, Black DM, McCormack WP. A generalized equation for prediction of VO2 peak from 1-mile run/walk performance. Med Sci Sports Exerc 1995;27:445-51.
6. California Department of Education. Fitnessgram Standards for Healthy Fitness Zone. Accessed Apr 28, 2005, from: http://www.cde.ca.gov/ta/tg/pf/documents/healthfitzones.pdf.
7. Centers for Disease Control and Prevention. Percentile Data Files with LMS Values. Accessed Nov 16, 2007 at: <http://www.cdc.gov/nchs/about/major/nhanes/growthcharts/datafiles.htm>.
8. Barlow SE, Dietz WH. Obesity evaluation and treatment: Expert Committee recommendations. The Maternal and Child Health Bureau, Health Resources and Services Administration and the Department of Health and Human Services. Pediatrics 1998;102:E29.
9. National Heart, Lung, and Blood Institute. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. Report accessed on Nov 16, 2007, at: http://www.nhlbi.nih.gov/guidelines/obesity/ob_model.htm: National Heart, Lung and Blood Institute; 1998.
10. Krebs NF, Jacobson MS. Prevention of pediatric overweight and obesity. Pediatrics 2003;112:424-30.
11. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research. Principles and Quantitative Methods. New York: van Nostrand Reinhold; 1982.
12. Jackson AP. Mothers' employment and poor and near-poor African-American children's development: a longitudinal study. Soc Serv Rev 2003;77:93-109.
13. Kann L, Kinchen SA, Williams BI, Ross JG, Lowry R, Grunbaum JA, et al. Youth risk behavior surveillance: United States, 1999. MMWR CDC Surveill Summ 2000;49:1-32.
14. Kann L, Kinchen SA, Williams BI, Ross JG, Lowry R, Hill CV, et al. Youth risk behavior surveillance: United States, 1997. MMWR CDC Surveill Summ 1998;47:1-89.
15. Pratt M, Macera CA, Blanton C. Levels of physical activity and inactivity in children and adults in the United States: current evidence and research issues. Med Sci Sports Exerc 1999;31:S526-33.
16. US Department of Health and Human Services. 1985 President's Council on Physical Fitness and Sports. Youth Fitness Survey. Washington, DC: US Government Printing Office; 1986.
17. Physical activity levels among children aged 9-13 years: United States, 2002. MMWR Morb Mortal Wkly Rep 2003;52:785-8.
18. Gordon-Larsen P, McMurray RG, Popkin BM. Determinants of adolescent physical activity and inactivity patterns. Pediatrics 2000;105:E83.
19. Grunbaum JA, Kann L, Kinchen S, Ross J, Hawkins J, Lowry R, et al. Youth risk behavior surveillance: United States, 2003. MMWR Surveill Summ 2004;53:1-96.
20. Beets MW, Pitetti KH. One-mile run/walk and body mass index of an ethnically diverse sample of youth. Med Sci Sports Exerc 2004;36:1796-803.
21. Public Health Institute. A Special Report on Policy Implications from the 1999 California Children's Healthy Eating and Exercise Practices Survey (CalCHEEPS). Sacramento, CA: Public Health Institute; May 2001.
22. Li Y, Dai Q, Jackson JC, Zhang J. Overweight is associated with decreased cognitive functioning among school-age children and adolescents. Obesity (Silver Spring) 2008;16:1809-15.
23. Laurin D, Verreault R, Lindsay J, MacPherson K, Rockwood K. Physical activity and risk of cognitive impairment and dementia in elderly persons. Arch Neurol 2001;58:498-504.
24. Hillman CH, Castelli DM, Buck SM. Aerobic fitness and neurocognitive function in healthy preadolescent children. Med Sci Sports Exerc 2005;37:1967-74.
25. Kalmijn S, Foley D, White L, Burchfiel CM, Curb JD, Petrovitch H, et al. Metabolic cardiovascular syndrome and risk of dementia in Japanese-American elderly men: the Honolulu-Asia aging study. Arterioscler Thromb Vasc Biol 2000;20:2255-60.
26. Dwyer T, Coonan WE, Leitch DR, Hetzel BS, Baghurst RA. An investigation of the effects of daily physical activity on the health of primary school students in South Australia. Int J Epidemiol 1983;12:308-13.
27. Sallis JF, McKenzie TL, Kolody B, Lewis M, Marshall S, Rosengard P. Effects of health-related physical education on academic achievement: project SPARK. Res Q Exerc Sport 1999;70:127-34.
28. Shephard R. Curricular physical activity and academic performance. Pediatr Exerc Sci 1997;9:113-26.
29. Cotman CW, Berchtold NC. Exercise: a behavioral intervention to enhance brain health and plasticity. Trends Neurosci 2002;25:295-301.
30. Rowland T, Kline G, Goff D, Martel L, Ferrone L. One-mile run performance and cardiovascular fitness in children. Arch Pediatr Adolesc Med 1999;153:845-9.
31. Plowman SA, Liu NY- S. Norm-referenced and criterion-referenced validity of the one-mile run and PACER in college age individuals. Meas Phys Ed Exerc Sci 1999;3:63-84.

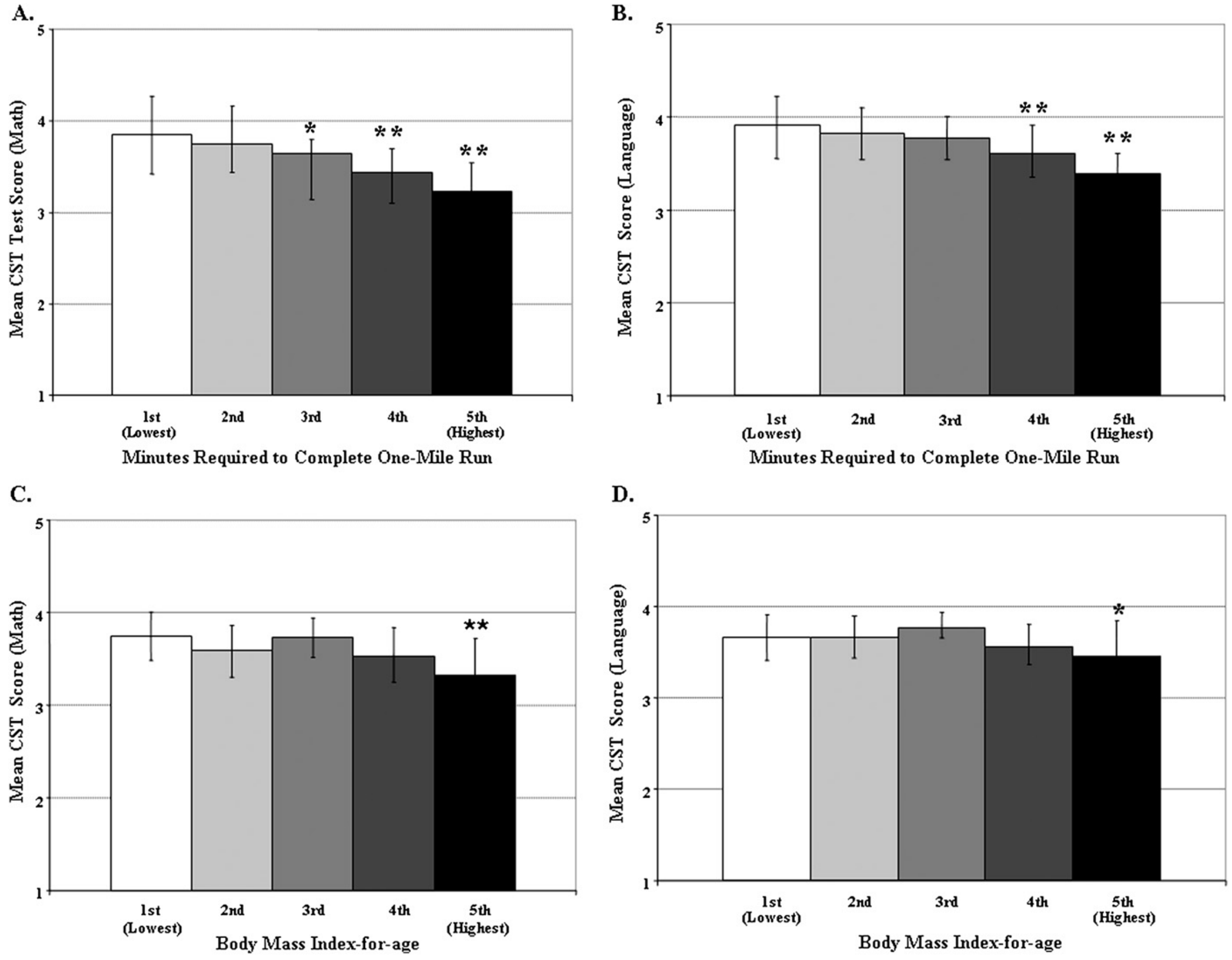

Figure 2. CST mean scores, adjusted by the covariates sex, free and reduced-price lunch eligibility, and ethnicity. **A,** Percentiles in CST score (math) by quintiles of minutes to complete mile run; *3rd quintile of mile time differed on 2002 CST math score from 1st quintile, $z = -2.27$, $P = .023$; **4th quintile of mile time differed on 2002 CST math score from 1st quintile, $z = -5.22$, $P < .0001$; **5th quintile of mile time differed on 2002 CST math score from 1st quintile, $z = -8.11$, $P < .0001$. **B,** Percentile in CST score (language) by quintiles of minutes to complete mile run; **4th quintile of mile time differed on 2002 CST language score from 1st quintile, $z = -3.41$, $P < .0001$; **5th quintile of mile time differed on 2002 CST language score from 1st quintile, $z = -6.46$, $P < .0001$. **C,** Percentiles in CST score (math) by BMI quintile; **5th quintile of BMI differed on 2002 CST math score from 1st quintile, $z = -5.14$, $P < .0001$. **D,** Percentiles in CST score (language) by BMI quintile; *5th quintile of BMI differed on 2002 CST language score from 1st quintile, $z = -2.51$, $P = .01$. Error bars are 95% confidence intervals.