Published Ahead of Print on February 28, 2008, as 10.2105/AJPH.2007.117176
The latest version is at http://www.ajph.org/cgi/doi/10.2105/AJPH.2007.117176

Physical Education and Academic Achievement in Elementary School: Data From the Early Childhood Longitudinal Study

Susan A. Carlson, MPH, Janet E. Fulton, PhD, Sarah M. Lee, PhD, L. Michele Maynard, PhD, David R. Brown, PhD, Harold W. Kohl III, PhD, and William H. Dietz, MD, PhD

Objectives. We examined the association between time spent in physical education and academic achievement in a longitudinal study of students in kindergarten through fifth grade.

Methods. We used data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998 to 1999, which employed a multistage probability design to select a nationally representative sample of students in kindergarten (analytic sample = 5316). Time spent in physical education (minutes per week) was collected from classroom teachers, and academic achievement (mathematics and reading) was scored on an item response theory scale.

Results. A small but significant benefit for academic achievement in mathematics and reading was observed for girls enrolled in higher amounts (70–300 minutes per week) of physical education (referent: 0–35 minutes per week). Higher amounts of physical education were not positively or negatively associated with academic achievement among boys.

Conclusions. Among girls, higher amounts of physical education may be associated with an academic benefit. Physical education did not appear to negatively affect academic achievement in elementary school students. Concerns about adverse effects on achievement may not be legitimate reasons to limit physical education programs. (*Am J Public Health.* 2008;98:XXX–XXX. doi:10.2105/AJPH.2007.117176)

Physically active youth may be less likely than physically inactive youth to experience chronic disease risk factors[1] and to become obese,[2] and they may be more likely to remain active throughout adolescence[3] and possibly into adulthood.[4] Physical activity also has beneficial influences on behavior and cognitive functioning that may result in improving students' academic achievement.[5–7] Direct indicators of academic achievement include grade-point averages, scores on standardized tests, and grades in specific courses; measures of concentration, memory, and classroom behavior provide indirect estimates.[1]

Several cross-sectional studies examined the association between physical activity and direct measures of academic achievement.[8–13] In addition, several intervention studies were conducted to examine the effect of introducing more physical activity and physical education programs during the school day on indirect estimates of behaviors related to academic achievement (e.g., concentration, memory, disruptive behavior) or on direct measures (e.g., standardized tests, academic record, teacher reports).[6,7,14–23] These studies had mixed results. Investigators observed either no association[6,8,13,14,16,18,23] or a modest-to-moderate positive association[6,7,9–12,15,17,19–22] between physical activity and academic achievement.

Physical education classes provide an opportunity for students to be physically active during the school day.[1] School-based physical education has many benefits, including increasing physical activity and improving physical fitness and muscular endurance.[24–28] Increasing physical activity through physical education is also a proposed public health strategy to reduce childhood obesity.[29] Although there has been no evidence to date to show that maintaining or increasing time in physical education class negatively affects academic achievement in other subjects, there is concern that physical education classes could take time away from those subjects.[1,28,30] More information is needed to address this concern and support public health objectives to maintain or expand physical education programs.[31]

We examined the influence of physical education in US elementary schools on direct measures of academic achievement in mathematics and reading from kindergarten through fifth grade. Our study was unique in at least 3 ways: first, the measurement of academic achievement was a standardized test administered at 5 time points. Second, we examined the association between physical education and academic achievement with a prospective cohort design. Finally, we examined participation in physical education as it existed in a representative sample of US students entering kindergarten in fall 1998 who were followed through spring 2004.

METHODS

Analytic Sample

Data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998 to 1999 (ECLS-K) were analyzed.[32] The ECLS-K employed a multistage probability sample design to select the nationally representative sample. In the baseline year, the 100 primary sampling units were geographic areas consisting of counties or groups of counties. The second-stage units were schools within the geographic areas (original selected sample = 1280 schools). The third-stage units were students within schools; the target number of students sampled at any 1 school was 24.[33] Further details of the study, including sampling procedures, are available elsewhere.[34]

The 5 time points for analyzing data were the fall of kindergarten, spring of kindergarten, spring of first grade, spring of third grade, and spring of fifth grade. At baseline, the overall school-level response rate (i.e., percentage of schools cooperating) was 69.4%

weighted (68.8% unweighted), and 19 173 eligible students (62.4% weighted, 61.8% unweighted) completed some part of the direct child assessment, which included the assessment of academic achievement.[33] Completion rates for the direct child assessment were as follows: spring of kindergarten, 88.0% weighted, 88.3% unweighted (n=19 967)[33]; first grade, 88.0% weighted, 91.8% unweighted (n=16 593)[35]; third grade, 80.8% weighted, 86.1% unweighted (n=14 349)[36]; and fifth grade, 84.7% weighted, 93.6% unweighted (n=11 260).[37] Students not eligible for follow-up (e.g., students who moved who were not subsampled or students who died or moved out of the country) were not included in the denominator.[33] Attrition of the sample was mainly attributable to students' moving, because only a subsample of movers were followed.[33,35–37]

Data for this analysis were obtained from the ECLS-K *Longitudinal Kindergarten–Fifth Grade Public-Use Data File*.[32] There were 9796 students (51.1% of baseline) who completed some part of the direct child assessment at all 5 time points. Students who were not in their expected grade level at any time point or who were identified as having attended kindergarten previously were excluded (n=1344). Of the 8452 remaining students, 7729 had mathematics and reading scores on the item response theory (IRT) scale for each time point. Finally, students who were missing data on time in physical education for any time point (n=2352) or who were missing information for 1 of the covariates of interest (i.e., race/ethnicity, mother's education, family income, and kindergarten enrollment status; n=61) were excluded. The final sample used for this analysis included 5316 students.

Measures

Physical education. Classroom teachers reported the number of times during the week (never, <1, 1 or 2, 3 or 4, or daily) and minutes per day (do not participate, 1–15, 16–30, 31–60, or >60) that students participated in physical education. Minutes per week of physical education ware estimated by multiplying the median frequency by the median duration (except for >60 minutes per day, where 60 minutes was used). Minutes per week ware then categorized into increments that represented approximate tertiles of the combined distribution of physical education for kindergarten through fifth grade. The use of tertiles allowed us to best incorporate the natural breaks in the distribution of minutes per week. Groups were labeled low (0–35 minutes per week), medium (36–69 minutes per week), and high (70–300 minutes per week).

Academic achievement. Students were administered mathematics and reading tests.[37] The National Center for Education Statistics and contractor staff assembled specialists in school curricula, teachers, and academicians to consult on the design and development of the assessment instruments, and the direct child assessment was adapted from several copyrighted assessment batteries.[34] From these assessments, IRT scale scores were calculated for each child. IRT procedures use the pattern of responses to estimate the probability of correct responses for all assessment questions; IRT scale scores represent estimates of the number of items students would have answered correctly at each point in time if they had been given all test questions (maximum: 186 for reading, 153 for mathematics).[37]

Covariates. Data on covariates were collected during a telephone interview with a parent.[37] Four categories were constructed for family income (≤$25 000, $25 001–50 000, $50 001–75 000, and >$75 000), the child's race/ethnicity (White, non-Hispanic; Black, non-Hispanic; Hispanic; other), and mother's education (less than high school graduate, high school graduate, some college, college graduate). Information collected in the fifth grade was used to categorize students unless that variable was unknown, in which case values from the most recent round available were used. If the respondent answered "don't know" or refused to answer the questions on education or family income and there was no information from a previous round, we used values from children from the score data set with similar, but complete data.[37] Kindergarteners were categorized as full day or half day.

Data Analysis

All estimates were calculated from statistical weights provided with the ECLS-K data set. Longitudinal statistical weights were derived from the sample of 9796 students with some assessment data at each time point.[32]

Stratified analyses compared IRT scale scores for mathematics and reading for each grade level by physical education category (low, medium, and high) for boys and girls. To account for the complex sample design, we used SUDAAN version 9.0 (Research Triangle Institute, Research Triangle Park, NC) to obtain all estimates and 95% confidence intervals (CIs). Pairwise differences between physical education categories were assessed with simple linear contrasts. All *P* values reported were 2-sided and were considered significant at less than .05.

Multivariate linear regression models were estimated with SUDAAN to test the longitudinal association between physical education and IRT scale scores for mathematics and reading stratified by gender. SUDAAN's generalized estimating equation approach accounted for the multiple levels of clustering (i.e., students within a cluster and repeated measures for each student).[38–40] A model was estimated in which the dependent variable was the child's IRT scale score at each grade level. The independent variables in the model included physical education category, grade, baseline IRT (IRT scale score for the fall of kindergarten), time from baseline (centered for each grade level by subtracting the student's number of days from baseline assessment from the weighted sample's mean number of days from baseline to current assessment), mother's education, family income, and race/ethnicity.

Because the effects of demographic variables on academic achievement were not constant over time, we included them in the model with a year interaction term. Baseline scores with a year interaction term were also included in the model. Including baseline scores as an independent variable allowed us to adjust for initial differences in achievement. A variable that indicated whether kindergartners attended school for only half a day was included, because this was associated with academic achievement and physical education category. Other variables considered for the model but not included because they were either not associated with physical education category and academic achievement or were not significant in a baseline-only model

TABLE 1—Sample (n = 5316) Baseline Assessment Characteristics: Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999

Continuous Variables	Unweighted Mean	Weighted Mean (SE)
Age at kindergarten assessment,[a] mo	74.9	74.9 (0.1)
IRT scale score at kindergarten[b]		
Reading	30.9	30.2 (0.3)
Mathematics	25.2	24.3 (0.3)

Note. IRT = item response theory.
[a]Spring kindergarten assessment.
[b]Fall kindergarten assessment.

were indicators of degree of urbanization of the child's residence, whether it was a public or private school, and school size.

RESULTS

The analytic sample was 52.1% girls, and the majority of the sample was non-Hispanic White (69.2%; Table 1). Weighted and unweighted percentages were similar for the demographic variables presented (differences in weighted and unweighted percentages reflected patterns in nonresponse and oversampling of Asians and Pacific Islanders; Table 2).[37] Demographic characteristics were similar for boys and girls, with only small significant differences in age in months (boys: 75.1; girls: 74.7) and baseline reading IRT scale score (boys: 30.8; girls: 29.6). There were significant differences by gender in the percentage with a family income of $25 000 or less (boys: 15.9%; girls: 20.0%) and in the percentage that was Hispanic (boys: 10.6%; girls: 15.8%).

Teachers most commonly reported that students had physical education 1 to 2 times per week (Table 3). The most commonly reported duration was 16 to 30 minutes for kindergarten and first grade and 31 to 60 minutes for third and fifth grade. From kindergarten through fifth grade, 12.6% (95% CI = 9.8, 16.0) of students (kindergarten: 14.5% [95% CI = 11.3, 18.4]; first grade: 12.8% [95% CI = 9.6, 17.0]; third grade: 11.5% [95% CI = 8.4, 15.6]; fifth grade: 11.5% [95% CI = 8.2, 16.0]) met the national Healthy People 2010 objective of participating in physical education daily (i.e., 5 times per wk).[31] Combining frequency and duration, the average minutes spent in physical education per week increased slightly as students progressed from kindergarten through fifth grade (kindergarten: 55.8 min [95% CI = 51.7, 60.0]; first grade: 65.7 min [95% CI = 61.9, 69.5]; third grade: 69.3 min [95% CI = 64.9, 73.8]; fifth grade: 76.3 min [95% CI = 70.4, 82.3]).

When grouped into 3 physical education categories, the percentage of students in the high category remained relatively constant for each of the grade levels. The percentage of students in the low category was highest in kindergarten and decreased as grade increased; the opposite was observed for the medium category (Table 3). The distribution in the 3 categories did not differ significantly by gender. In kindergarten, this distribution differed greatly by kindergarten enrollment status. The prevalence in the high category for half-day kindergartners was 9.5% (95% CI = 6.6, 13.6); for full-day kindergarteners the prevalence was 34.3% (95% CI = 27.7, 41.7).

The cross-sectional associations between academic achievement and physical education category for each grade level for boys and girls are presented in Table 4. Girls in all grades who were in the low physical education category had the lowest IRT scale scores for mathematics and reading, although only in kindergarten and first grade were these differences significant for reading and mathematics (in kindergarten, only the difference between the low and medium category was significant). In fifth grade differences were significant for reading only. For boys, there were no significant differences in IRT scale scores in reading or mathematics between any of the 3 physical education categories for any of the grades.

Similar results were observed in the longitudinal association between physical education and academic achievement. No association between physical education category and academic achievement in reading or mathematics was observed among boys (Table 5). After adjusting for grade-level gains, baseline scores, and kindergarten program, girls in the medium and high physical education categories had a small but significant benefit in their IRT scale score for reading compared

TABLE 2—Sample Sociodemographic Characteristics: Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999

Categorical Variables	No. (n = 5316)	Unweighted %	Weighted % (SE)
Girls	2769	52.1	51.8 (1.4)
Half-day kindergarten enrollment	2402	45.2	44.9 (3.0)
Race/ethnicity			
Non-Hispanic White	3680	69.2	68.4 (2.3)
Non-Hispanic Black	462	8.7	11.4 (1.5)
Hispanic	592	11.1	13.3 (1.6)
Other[a]	582	10.9	7.0 (1.0)
Mother's education			
Less than high school	269	5.1	6.9 (0.8)
High school graduate	1322	24.9	26.0 (1.1)
Some college	1962	36.9	37.1 (1.2)
College graduate	1763	33.2	30.0 (1.5)
Family income, $			
≤25 000	791	14.9	18.0 (1.2)
25 001–50 000	1491	28.0	29.4 (1.2)
50 001–75 000	1126	21.2	20.4 (1.0)
>75 000	1908	35.9	32.2 (1.3)

Note. Percentages may not add to 100 because of rounding.
[a]Includes Asian, Pacific Islander, American Indian, Alaska Native, and more than 1 race (non-Hispanic).

TABLE 3—Time in Physical Education, by Grade Level: Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999

	Kindergarten		First Grade		Third Grade		Fifth Grade	
	No.	% (SE)	No.	% (SE)	No.	% (SE)	No.	% (SE)
Physical education, times/wk by min/d[a]								
Never	401	7.5 (1.3)	102	2.0 (0.4)	132	2.4[b] (0.8)	155	3.0 (0.6)
<1 Time/wk, min/d								
1–15	80	1.4[b] (0.4)	18	0.5[b] (0.2)	7	0.1[b] (0.05)	23	0.5[b] (0.3)
16–30	108	3.0 (0.7)	51	1.0 (0.2)	43	1.3[b] (0.4)	56	1.3[b] (0.4)
31–60	33	0.4[b] (0.2)	85	1.6[b] (0.5)	79	1.3[b] (0.4)	94	2.3 (0.7)
>60	0	0	0	0	1	0.01[b] (0.01)	17	0.2[b] (0.2)
1–2 Times/wk, min/d								
1–15	215	4.2 (0.7)	191	3.9 (0.8)	140	3.0 (0.7)	129	2.3 (0.5)
16–30	2238	40.9 (2.4)	1862	35.1 (2.3)	1630	28.9 (2.7)	1029	18.5 (2.1)
31–60	861	16.3 (2.1)	1543	26.4 (2.1)	1987	37.2 (3.1)	2435	44.5 (2.5)
>60	0	0	6	0.1[b] (0.04)	2	0.03[b] (0.03)	5	0.1[b] (0.1)
3–4 Times/wk, min/d								
1–15	90	1.6 (0.5)	18	0.4[b] (0.2)	6	0.1[b] (0.1)	13	0.3[b] (0.1)
16–30	375	7.1 (1.0)	609	11.5 (1.7)	397	8.3 (1.6)	393	8.5 (1.9)
31–60	168	3.1[b] (1.1)	232	4.6 (1.2)	324	5.9 (1.3)	416	6.7 (1.4)
>60	0	0	0	0	0	0	2	0.1[b] (0.1)
5 Times/wk, min/d								
1–15	88	1.6 (0.5)	16	0.4[b] (0.2)	8	0.2[b] (0.1)	8	0.2 (0.2)
16–30	532	10.3 (1.6)	471	9.4 (1.8)	417	7.9 (1.6)	278	4.7 (1.0)
31–60	125	2.6 (0.6)	112	3.0 (0.6)	142	3.5 (0.6)	255	6.5 (1.5)
>60	2	0.02[b] (0.02)	0	0	1	0.01[b] (0.01)	8	0.1[b] (0.1)
Physical education category[c]								
Low	3165	59.0 (2.3)	2327	44.5 (2.1)	2038	37.1 (2.9)	1516	28.5 (2.2)
Medium	949	17.8 (2.0)	1559	26.9 (2.1)	1995	37.4 (3.1)	2443	44.8 (2.4)
High	1202	23.2 (2.1)	1430	28.6 (2.4)	1283	25.6 (2.6)	1357	26.8 (2.7)

Note. Percentages are weighted and may not add to 100 because of rounding.
[a]Frequency and duration of physical education were reported by teachers, who used the categories provided in the table for frequency and duration.
[b]Estimates may be unstable because the coefficient of variation was more than 30%.
[c]Minutes per week of physical education was estimated by multiplying the median frequency by the median duration and categorized by minutes per week as low (0–35), medium (36–69), or high (70–300).

with girls in the low category (Table 5). After we further controlled for selected demographic variables, the association with reading IRT scale scores was attenuated, and only those in the high physical education category (vs the low category) showed a small but significant benefit.

For mathematics, only the difference for girls between the high and low categories was significant in the adjusted models, and this association remained after we controlled for selected demographics. From kindergarten through fifth grade, girls with the highest exposure to physical education scored on average 2.4 points higher on the IRT reading scale and on average 1.5 points higher on the IRT mathematics scale than did those in the low physical education category. There was no significant interaction for this association between grade level and physical education category.

DISCUSSION

In this study on the association between time spent in physical education and academic achievement, we observed that girls with the highest exposure to physical education (70–300 minutes per week) versus the lowest exposure (0–35 minutes per week) exhibited a small academic benefit for mathematics and reading; we observed no association for boys. Our study supports findings from previous studies in which investigators concluded that time spent in physical education did not harm academic achievement[6,14] and that it may have a modest favorable effect on achievement.[7,19]

Our study was unique in that we examined physical education in a nationally representative sample of kindergarteners. We found that the overall exposure to physical education in this representative sample was much lower than the national health objective of daily physical education, with an average of only 12.6% meeting the objective.[31] An experimental design would have allowed us to manipulate the physical education exposure to higher levels, but even with exposure to physical education as low as 70 to 300 minutes per week, a small benefit was observed among girls. One experimental study that implemented an intervention of 1 hour per day of physical education versus 40 minutes per week for controls observed a positive effect on academic scores for boys and girls in primary school.[19] Expanding physical education programs may increase the benefit in academic achievement as well as enhance other potential benefits, such as increasing physical activity levels, improving physical fitness, increasing knowledge about physical activity, and improving psychological health (e.g., raising self-esteem and reducing stress and anxiety).[24,25,28,41]

The difference in our findings between boys and girls accords with results for other researchers who have examined effects of physical education and school-day physical activity programs on various factors, such as academic scores, cardiorespiratory fitness, and body mass index. Shephard et al. showed, in a controlled study, that girls gained a larger advantage than did boys in academic scores from the addition of 5 hours per week of physical education.[19] Other studies that have examined the association of participating in physical education[24,42] or participating in a structured school-day physical activity program emphasizing fitness[43] with measures

TABLE 4—Cross-Sectional Mean IRT Scale Scores in Mathematics and Reading, by Gender, Grade Level, and Physical Education Category: Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999

	Physical Education Category[a]			
	Low, Mean (95% CI)	Medium, Mean (95% CI)	High, Mean (95% CI)	Overall, Mean (95% CI)
	Boys			
Reading				
Kindergarten	41.4 (40.2, 42.7)	42.6 (40.7, 44.5)	42.8 (40.4, 45.1)	42.0 (41.0, 42.9)
First grade	75.6 (72.9, 78.4)	76.4 (72.8, 80.0)	74.3 (72.0, 76.6)	75.5 (73.6, 77.4)
Third grade	121.4 (118.1, 124.7)	121.1 (118.0, 124.1)	124.1 (120.7, 127.4)	122.0 (119.9, 124.0)
Fifth grade	143.0 (139.1, 146.8)	144.1 (141.8, 146.3)	143.2 (140.4, 146.1)	143.5 (141.7, 145.4)
Mathematics				
Kindergarten	35.9 (34.7, 37.2)	37.2 (35.2, 39.2)	35.5 (33.4, 37.5)	36.0 (35.0, 37.1)
First grade	61.4 (59.7, 63.1)	62.4 (60.1, 64.7)	63.8 (61.8, 65.9)	62.4 (61.1, 63.7)
Third grade	99.1 (96.5, 101.8)	99.4 (96.8, 102.0)	102.1 (98.9, 105.4)	100.0 (98.2, 101.9)
Fifth grade	119.5 (116.1, 123.0)	120.3 (118.1, 122.5)	119.7 (116.7, 122.7)	119.9 (118.2, 121.6)
	Girls			
Reading				
Kindergarten[b,c]	41.9 (40.5, 43.4)	44.7 (42.9, 46.6)	44.6 (42.8, 46.4)	43.0 (42.0, 44.1)
First grade[b,c]	75.0 (72.5, 77.5)	80.2 (77.0, 83.3)	78.9 (76.7, 81.2)	77.5 (75.8, 79.1)
Third grade	123.7 (120.5, 126.9)	126.3 (123.9, 128.6)	126.3 (123.3, 129.3)	125.3 (123.5, 127.1)
Fifth grade[b,c]	141.9 (138.8, 145.1)	147.1 (144.3, 149.8)	145.8 (143.9, 147.7)	145.4 (143.7, 147.1)
Mathematics				
Kindergarten[b]	34.0 (32.9, 35.2)	36.2 (34.5, 37.8)	35.2 (33.7, 36.7)	34.7 (33.7, 35.6)
First grade[b,c]	58.0 (56.3, 59.6)	61.9 (60.0, 63.7)	61.2 (59.2, 63.2)	59.9 (58.8, 61.0)
Third grade	92.1 (89.2, 94.9)	94.3 (91.9, 96.8)	93.7 (89.1, 98.3)	93.3 (91.4, 95.3)
Fifth grade	112.6 (109.3, 115.8)	116.0 (113.3, 118.7)	114.9 (111.8, 117.9)	114.8 (112.8, 116.7)

Note. IRT = item response theory; CI = confidence interval. All assessments carried out in the spring.
[a]Minutes per week of physical education was estimated by multiplying the median frequency by the median duration of time reported by teachers for physical education and then categorized by minutes per week as low (0–35), medium (36–69), or high (70–300).
[b]Significant difference between low and medium category ($P < .05$).
[c]Significant difference between low and high category ($P < .05$).

of cardiorespiratory fitness[24,43] or body mass index[42] also observed differences in effectiveness by gender, with girls having larger effects in measures of body mass index and cardiorespiratory fitness.

When examining the association between physical activity and cardiorespiratory fitness, researchers have suggested that gender differences in the effects of physical activity on fitness may be attributable to lower levels of fitness in girls at baseline.[24,43] Boys are generally more fit than girls,[44] and thus the stimulus achieved during physical education may not be sufficient in boys to produce the same physiological effect experienced by girls. This difference may help explain why we observed a benefit of physical education on academic achievement in girls but not in boys; in addition to its physiological effects, physical education can influence other developmental domains, such as the social and cognitive domains,[41] and these effects may differ for girls and boys. Even so, we cannot say with any certainty that influences on these (or other) key developmental factors may explain our findings. Future investigations should seek to clarify how physical activity during physical education relates to measures of academic achievement and, in particular, how factors related to physical, social, and psychological growth and development may modify or mediate this relation differently in boys and girls.

Strengths and Limitations

Our study had limitations as well as strengths. We used a large database that was representative of the nation's kindergartners (class of 1998–1999), thus allowing us to examine physical education as it existed in US elementary schools. The ECLS-K data were weighted to be nationally representative, and they were adjusted for the effects of nonresponse.[32] The statistical weights, however, did not adjust for the largest source of missing data, which was time in physical education. The absence of such data may have biased our sample, because after weighting, students with complete data were more likely than those with missing data to be White and less likely to be Black, and they were more likely to have a family income of more than $75 000 and less likely to have a family income of $25 000 of below. For academic achievement, missing-data status was significantly associated with the reading IRT scale score only in the fifth grade.

In addition, our measures had strengths and limitations. Academic achievement was assessed through a standardized test, thereby removing any potential rater bias that might have been introduced if academic achievement had been based on ratings by classroom teachers. The reading and mathematics IRT scale scores exhibited a high level of reliability (reliability of Θ: 0.90–0.95 for reading, 0.89–0.94 for mathematics).[45] The standardized scores, however, did not allow us to examine the role of physical education on more indirect indicators of academic achievement, such as concentration, memory, and classroom behavior.[1]

We had no reliability or validity information on the questions used to assess time spent in physical education. The teacher's assessment of time in physical education may have been less accurate than more direct observational assessment methods, and we did not know the actual time that the students spent being active during physical education classes, nor did we have information on the quality of the physical education classes, such as whether the class was taught by a certified physical education specialist or specific information about the class curriculum. If we had measurements of quality and actual active time during physical education, we might

TABLE 5—Longitudinal Analysis of IRT Scale Scores for Mathematics and Reading, by Gender and Physical Education Category: Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999

Gender, by Physical Education Category[a]	Unadjusted,[b] B (95% CI)	Adjusted for Baseline and Kindergarten Program,[c] B (95% CI)	Adjusted for Baseline, Kindergarten Program, and Race/Ethnicity,[d] B (95% CI)	Adjusted for Baseline, Kindergarten Program, Race/Ethnicity, Family Income, and Mother's Education,[e] B (95% CI)
Reading				
Boys				
Low (Ref)	…	…	…	…
Medium	0.6 (−1.9, 3.1)	0.5 (−1.3, 2.3)	0.0 (−1.6, 1.7)	0.2 (−1.6, 1.9)
High	0.7 (−1.8, 3.2)	−0.4 (−2.5, 1.6)	−0.3 (−2.2, 1.7)	−0.3 (−2.3, 1.7)
Girls				
Low (Ref)	…	…	…	…
Medium	3.9 (1.7, 6.1)	2.1 (0.6, 3.6)	1.7 (0.3, 3.0)	1.1 (−0.2, 2.4)
High	3.2 (0.8, 5.6)	2.8 (1.2, 4.4)	2.4 (1.0, 3.9)	2.4 (1.0, 3.9)
Mathematics				
Boys				
Low (Ref)	…	…	…	…
Medium	0.7 (−1.4, 2.8)	0.2 (−1.2, 1.6)	−0.2 (−1.6, 1.1)	−0.2 (−1.5, 1.1)
High	1.3 (−0.8, 3.4)	0.3 (−1.2, 1.7)	0.3 (−1.1, 1.8)	0.4 (−1.0, 1.8)
Girls				
Low (Ref)	…	…	…	…
Medium	2.9 (0.8, 5.0)	1.2 (−0.1, 2.5)	1.0 (−0.2, 2.1)	0.7 (−0.4, 1.8)
High	2.1 (−0.3, 4.5)	1.7 (0.3, 3.1)	1.5 (0.1, 2.8)	1.5 (0.4, 2.7)

Note. IRT = item response theory; CI = confidence interval.
[a]Minutes per week of physical education was estimated by multiplying the median frequency by the median duration of time reported by teachers for physical education and then categorized by minutes per week as low (0–35), medium (36–69), or high (70–300).
[b]Model contained 2 independent variables: physical education category and grade level.
[c]Model added a baseline IRT scale score (with a grade-level interaction term), an indicator for time from baseline assessment, and a variable to indicate whether a student attended school for half a day (i.e., half-day kindergarteners).
[d]Model added race/ethnicity with a grade-level interaction term.
[e]Model added family income and mother's education with a grade-level interaction term.

have obtained more-striking results. On the other hand, because information about time in physical education was collected from classroom teachers at each of the time points, the potential for recall bias may have been low. In addition, we chose to use empirically developed categories to examine time in physical education. This allowed us to compare categories of physical education on the basis of the distribution of a nationally representative sample. We did not use criterion-referenced categories, although our high group included students who met the Healthy People physical education objectives when we included frequency and duration (assuming 20 minutes per day).[31]

Finally, the longitudinal design enabled us to control for baseline scores and grade-level gains in academic achievement as we followed students from kindergarten through fifth grade. Although students may have had some exposure to physical education before the baseline academic achievement assessment, because they may have been attending school for a few weeks before that assessment, our inclusion of baseline scores in the model should have enabled us to effectively control for baseline academic achievement. Unfortunately, we did not have yearly time points (data were not collected in second and fourth grades). Our method of data analysis allowed us to account for the repeated measures in the analysis, as well as the clustering in the complex study design, with robust variance estimators.[38] If used on data outside the context of a complex sample survey, random-effects analysis of longitudinal data might have given us a more powerful analysis, but by using SUDAAN we were able to appropriately apply statistical weights and account for the multiple levels of clustering in our sample design.

Conclusions

We found that girls obtained a small benefit from having the highest level (70–300 minutes per week) of exposure to physical education (vs the referent), but no association was observed among boys. More time in physical education, therefore, may help students perform better academically. Physical education should be promoted for its many benefits, and fear of negatively affecting academic achievement does not seem to be a legitimate reason for reducing or eliminating programs in physical education. Schools should strive to meet the national health objective of daily physical education[35] and offer students a balanced academic program that includes opportunities for physical activity. ■

About the Authors

The authors are with the Centers for Disease Control and Prevention, Atlanta, Ga.

Requests for reprints should be sent to Susan A. Carlson, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 4770 Buford Highway, NE, Mailstop K-46, Atlanta, GA 30345 (e-mail: scarlson1@cdc.gov).

This article was accepted August 28, 2007.

Acknowledgments

The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Contributors

S.A. Carlson completed the analyses, drafted the article, and incorporated revisions from the other coauthors. J.E. Fulton originated the study and assisted with the analyses and writing of the article. S.M. Lee, L.M. Maynard, D.R. Brown, H.W. Kohl III, and W.H. Dietz assisted with the interpretation of the data and contributed to critical revisions of the article. In addition, W.H. Dietz introduced the data source to the team.

Human Participant Protection

No protocol approval was needed for this study.

References

1. Strong WB, Malina RM, Blimkie CJ, et al. Evidence

based physical activity for school-age youth. *J Pediatr.* 2005;146:732–737.

2. Moore LL, Nguyen US, Rothman KJ, Cupples LA, Ellison RC. Preschool physical activity level and change in body fatness in young children. The Framingham Children's Study. *Am J Epidemiol.* 1995;142:982–988.

3. Malina RM. Tracking of physical activity and physical fitness across the lifespan. *Res Q Exerc Sport.* 1996;67(suppl 3):S48–S57.

4. *Physical Activity and Health: A Report of the Surgeon General.* Atlanta, Ga: Centers for Disease Control and Prevention; 1996.

5. Tomporowski PD. Cognitive and behavioral responses to acute exercise in youths: a review. *Pediatr Exerc Sci.* 2003;15:348–359.

6. Coe DP, Pivarnik JM, Womack CJ, Reeves MJ, Malina RM. Effect of physical education and activity levels on academic achievement in children. *Med Sci Sports Exerc.* 2006;38:1515–1519.

7. 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;70:127–134.

8. Daley AJ, Ryan J. Academic performance and participation in physical activity by secondary school adolescents. *Percept Mot Skills.* 2000;91:531–534.

9. Dwyer T, Sallis JF, Blizzard L, Lazarus R. Dean K. Relationship of academic performance to physical activity and fitness in children. *Pediatr Exerc Sci.* 2001; 13:225–237.

10. Field T, Diego M, Sanders CE. Exercise is positively related to adolescents' relationships and academics. *Adolescence.* 2001;36:105–110.

11. Hanson TL, Austin G. *Student Health Risks, Resilience and Academic Performance in California: Year 2 Report, Longitudinal Analysis.* Los Alamitos, Calif: West Ed; 2002.

12. Sanders CE, Field TM, Diego M, Kaplan M. Moderate involvement in sports is related to lower depression levels among adolescents. *Adolescence.* 2000;35:793–797.

13. Fisher M, Juszczak L, Friedman SB. Sports participation in an urban high school: academic and psychologic correlates. *J Adolesc Health.* 1996;18:329–334.

14. 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: 308–313.

15. Klein SA, Deffenbacher JL. Relaxation and exercise for hyperactive impulsive children. *Percept Mot Skills.* 1977;45:1159–1162.

16. MacMahon JR, Gross RT. Physical and psychological effects of aerobic exercise in boys with learning disabilities. *J Dev Behav Pediatr.* 1987;8:274–277.

17. McNaughten D, Gabbard C. Physical exertion and immediate mental performance of sixth-grade children. *Percept Mot Skills.* 1993;77:1155–1159.

18. Raviv S, Low M. Influence of physical activity on concentration among junior high-school students. *Percept Mot Skills.* 1990;70:67–74.

19. Shephard RJ, Volle M, Lavallee H, LaBarre R, Jequier JC, Rajic M. Required physical activity and academic grades: a controlled study. In: Ilmarinen J, Valimaki I, eds. *Children and Sport.* Berlin, Germany: Springer-Verlag; 1984:58–63.

20. Shephard RJ. Habitual physical activity and academic performance. *Nutr Rev.* 1996;54(4 part 2): S32–S36.

21. Shephard RJ, Lavallee H, Volle M, LaBarre R, Beaucage C. Academic skills and required physical education: the Trois Rivieres experience. *CAPHER J Res Suppl.* 1994;1:1–12.

22. Mahar MT, Murphy SK, Rowe DA, Golden J, Shields AT, Raedeke TD. Effects of a classroom-based program on physical activity and on-task behavior. *Med Sci Sports Exerc.* 2006;38:2086–2094.

23. Ahamed Y, Macdonald H, Reed K, Naylor PJ, Liu-Ambrose T, McKay H. School-based physical activity does not compromise children's academic performance. *Med Sci Sports Exerc.* 2007;39:371–376.

24. Sallis JF, McKenzie TL, Alcaraz JE, Kolody B, Faucette N, Hovell MF. The effects of a 2-year physical education program (SPARK) on physical activity and fitness in elementary school students. Sports, Play and Active Recreation for Kids. *Am J Public Health.* 1997; 87:1328–1334.

25. Trudeau F, Shephard RJ. Contribution of school programmes to physical activity levels and attitudes in children and adults. *Sports Med.* 2005;35:89–105.

26. Centers for Disease Control and Prevention. Guidelines for school and community programs to promote lifelong physical activity among young people. *Morb Mortal Wkly Rep.* 2006;46(RR-6):1–36.

27. Fairclough S, Stratton G. "Physical education makes you fit and healthy." Physical education's contribution to young people's physical activity levels. *Health Educ Res.* 2005;20:14–23.

28. Task Force on Community Preventive Services. Recommendations to increase physical activity in communities. *Am J Prev Med.* 2002;22(4S):67–72.

29. Committee on Prevention of Obesity in Children and Youth, Koplan JP, Liverman CT, Kraak VI. *Preventing Childhood Obesity: Health in the Balance: Executive Summary.* Washington, DC: National Academies Press; 2005.

30. Kahn EB, Ramsey LT, Brownson RC, et al. The effectiveness of interventions to increase physical activity. A systematic review. *Am J Prev Med.* 2002;22 (suppl 4):73–107.

31. *Healthy People 2010.* Conference ed, 2 vols. Washington, DC: US Department of Health and Human Services; 2000.

32. National Center for Education Statistics. *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): Longitudinal Kindergarten–Fifth Grade Public-Use Data Files.* Washington, DC: US Dept of Education; 2006. NCES 2006-035.

33. National Center for Education Statistics. *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): User's Manual for the Base Year Public-Use Data Files and Electronic Codebook.* Washington, DC: US Dept of Education; 2001. NCES 2001-029.

34. US Department of Education. Early Childhood Longitudinal Study. Available at: http://nces.ed.gov/ecls. Accessed January 9, 2007.

35. National Center for Education Statistics. *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Codebook.* Washington, DC: US Dept of Education; 2002. NCES 2002-135.

36. National Center for Education Statistics. *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): User's Manual for the ECLS-K Third Grade Public-Use Data File and Electronic Codebook.* Washington, DC: US Dept of Education; 2004. NCES 2004-001.

37. National Center for Education Statistics. *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): User's Manual for the ECLS-K Fifth Grade Public-Use Data File and Electronic Codebook.* Washington, DC: US Dept of Education; 2006. NCES 2006-032.

38. *SUDAAN Language Manual, Release 9.0.* Research Triangle Park, NC: Research Triangle Institute; 2004.

39. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics.* 1988;44:1049–1060.

40. Horton NJ, Lipsitz SR. Review of software to fit generalized estimating equation regression models. *Am Stat.* 1999;53:160–169.

41. Bailey R. Physical education and sport in schools: a review of benefits and outcomes. *J Sch Health.* 2006; 76:397–401.

42. Datar A, Sturm R. Physical education in elementary school and body mass index: evidence from the early childhood longitudinal study. *Am J Public Health.* 2004;94:1501–1506.

43. Vandongen R, Jenner DA, Thompson C, et al. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. *Prev Med.* 1995;24:9–22.

44. Pate RR, Wang CY, Dowda M, Farrell SW, O'Neill JR. Cardiorespiratory fitness levels among US youth 12 to 19 years of age: findings from the 1999–2002 National Health and Nutrition Examination Survey. *Arch Pediatr Adolesc Med.* 2006;160:1005–1012.

45. National Center for Education Statistics. *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): Psychometric Report for the Fifth Grade.* Washington, DC: US Dept of Education; 2005. NCES 2006-036.