

Promoting Elementary Physical Education: Results of a School-Based Evaluation Study

Trina Boyle-Holmes, MS

Lisa Grost, MHSA

Lisa Russell, PhD

B. A. Laris, MPH

Leah Robin, PhD

Elizabeth Haller, MS

Susan Potter, MS

Sarah Lee, PhD

Using a quasiexperimental design, the authors examine whether fourth- and fifth-grade students exposed to a developmental physical education (PE) curriculum, Michigan's Exemplary Physical Education Curriculum (EPEC), demonstrated stronger motor skill-specific self-efficacy and perceptions of physical activity competence, physical activity levels, motor skills, and physical fitness than did students exposed to existing PE curricula. The authors conducted a multilevel regression analysis with data from 1,464 students in the fourth and fifth grades. Data were collected using a student survey, an activity checklist, and motor and fitness assessments. Compared to students receiving standard PE, students exposed to EPEC showed significantly stronger results in motor skills but not fitness outcomes. The authors found significant positive intervention effects on indicators of motor skill self-efficacy and physical activity levels among the fourth-grade cohort. EPEC was more effective than standard PE curricula at improving motor skill performance (fourth- and fifth-grade cohorts) and at increasing self-reported motor skill-specific self-efficacy and physical activity (fourth-grade cohort).

Keywords: *physical education and training; motor skills; school-based program; program evaluation*

Trina Boyle-Holmes, Michigan Department of Education, Lansing. Lisa Grost, Michigan Department of Health, Lansing. Lisa Russell and B. A. Laris, ETR Associates, Scotts Valley, California. Leah Robin and Elizabeth Haller, Centers for Disease Control and Prevention, Atlanta, Georgia. Susan Potter, ETR Associates, Scotts Valley, California. Sarah Lee, Centers for Disease Control and Prevention, Atlanta, Georgia.

Address correspondence to Lisa Russell, ETR Associates, 4 Carbonero Way, Scotts Valley, CA 95066; phone: (831) 438-4060, ext. 184; e-mail: lisar@etr.org.

This study was supported by funding from the Centers for Disease Control and Prevention (Contract 200-2002-00800). We thank Kyle Guerrant of the Michigan Department of Education; Charlene Burgesen of the National Association for Sport and Physical Education; Glenna Dejong, Lee Kokinakakis, and Lorin Sheppard of the Michigan Fitness Foundation; Jackie Epping of the Centers for Disease Control and Prevention, Division of Nutrition and Physical Activity; Jill Glassman of ETR Associates; Krystin Martens and the team of Michigan-based data collectors; and the children, parents, and staff at the participating schools. 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.

Quality physical education (PE) classes are a key national strategy to promote the physical activity and health of youth. National guidelines promote the role of PE for students in Grades K-12 to help develop the knowledge, attitudes, skills, behaviors, and confidence to adopt and maintain physically active lifestyles (Centers for Disease Control and Prevention, 2006). However, within a PE curriculum the optimal balance between motor skill instruction and physical activity and fitness instruction that would enable youth to meet physical activity recommendations and obtain health-enhancing fitness levels has not been established (Trost, 2004).

Among youth, both motor skill level and perceived physical activity competence have been associated with increased levels of physical activity (Dishman et al., 2004; Okely, Booth, & Patterson, 2001). Recently, Wrotniak and colleagues examined motor proficiency and its association with self-perception of adequacy in and predilection for physical activity and physical activity levels among 8- to 10-year-old youth (Wrotniak, Epstein, Dorn, Jones & Kondilis, 2006). The authors concluded that motor proficiency was positively associated with physical activity among boys and girls and self-perception and predilection for physical activity were positively associated with motor proficiency among boys. One study that examined the relationship among overall physical activity, physical activity during PE, and performance on three motor skills tests concluded that motor skills test scores of the sixth- to eighth-grade students were not strongly associated with overall physical activity but that strong significant correlations did exist between physical activity during PE class and two of the motor skills (Reed, Metzker, & Phillips, 2004).

Relatively few studies have examined the impact of elementary school PE on the development of motor skills. One study that compared an enhanced, motor skills-focused curriculum with a regular, existing PE curriculum to determine the effect of self-testing activities on the development of fundamental movement skills in first-grade children was conducted. The enhanced motor skills curriculum was related to improvements in gross motor development of the students (Karabourniotis, Evaggelidou, Tzetzis, & Kourtessis, 2002). Similarly, another study found that motor performance (coordination) and fitness (endurance) significantly improved among first-grade children exposed to a health education and motor skills-focused PE curriculum (Graf et al., 2005). A third comparative study of a PE curriculum focused on improving fundamental movement skills demonstrated evidence that a PE program could improve motor skills without sacrificing time spent on moderate to vigorous physical activity among elementary school students (van Beurden et al., 2003). These studies of motor skills-enhanced PE curricula did not report on psychosocial correlates or mediators.

Few PE curricula have emphasized motor skills and physical activity. Michigan's Exemplary Physical Education Curriculum (EPEC), however, provides a unique opportunity to study a PE curriculum that includes extensive instruction on the developmental progression of motor skills along with physical activity and fitness instruction. This article describes a comparative evaluation of EPEC in elementary schools and allows further examination of the links among motor skills, physical activity, fitness, and psychosocial correlates.

METHOD

Study Design

To assess the impact of the EPEC curriculum, this evaluation study used a quasi-experimental design involving 16 schools, comparing cohorts of fourth- and fifth-grade

students receiving EPEC to those receiving a variety of other PE curricula during the 2003-2004 and 2004-2005 school years. Baseline measures were obtained in fall 2003, and three follow-up measures were obtained in spring 2004, fall 2004, and spring 2005. All fourth and fifth graders in participating schools received PE as part of their regular school curriculum.

Intervention Arm. Eight public elementary schools with PE teachers who were actively implementing EPEC among fourth and fifth graders participated in the intervention condition. All EPEC teachers were trained in using EPEC before the start of the study. EPEC is a PE curriculum that focuses on developing knowledge, attitudes, skills, and behaviors that are associated with lifelong physical activity through teaching and motor skills learning progressions. EPEC lessons address all of the content standards of the National Association for Sport and Physical Education (NASPE, 2004) and the Michigan Physical Education Content Standards (Michigan Department of Education, 1998; Michigan's Exemplary Physical Education Curriculum Project, 2001). In EPEC, teaching progressions lead students through small cumulative and connected learning steps (learning progressions) that are postulated to improve student confidence from novice through mastery stages of movement skills. Assessment and feedback are intrinsic to matching teaching progressions to students' learning progressions. EPEC contains 51 lessons per grade, designed to be taught 2 days per week for 30 minutes throughout the school year. On average, intervention teachers reported teaching 44 lessons per year (about 75% of the curriculum), with lessons focusing predominantly on the following three standards: motor skill and movement, values of physical activity for health and enjoyment, and regular physical activity.

Comparison Arm. Eight public elementary schools taught their existing PE curricula. Through implementation surveys and logs, comparison condition teachers described their PE curricula most typically as an eclectic collection of materials from various sources, some of which were developed at the school or district level and some of which were obtained from other sources by the PE staff. Teachers did not report using EPEC as a primary resource. All content was reported to be linked to the Michigan Physical Education Content Standards for PE. Implementation logs were completed by 94% of the teachers in the comparison condition. Comparison group teachers were not asked about underlying instructional frameworks or philosophies that may have guided how they taught the content standards. On average, comparison teachers reported teaching 48 lessons per year, with lessons focusing predominantly on the following three standards: motor skills and movement, movement concepts, and responsible personal and social behavior.

Curricula Comparison Across Arms. PE classes were taught with similar frequency and duration (i.e., an average of 40 minutes per class, 2 to 3 days per week) across study arms. All of the instructors were Michigan certified to teach PE in Grades K-12. Comparisons of implementation log data indicate that all NASPE content standards were addressed by the PE curriculum in each study arm during the study period. In both conditions, similar amounts of class time were spent on lesson components (explaining, demonstrating, practicing, reviewing skills, and classroom management). However, intervention teachers reported using a written lesson plan for more than four fifths of their PE lessons, whereas comparison teachers reported using a written lesson plan less than half of the time.

Sample

To track fourth- and fifth-grade cohorts for 2 years, only the schools that included a sixth grade were eligible for participation. In all, 16 schools were selected to represent the diversity of Michigan. The principals and the PE teachers at all of the 16 invited schools agreed to participate in data collection activities. All of the enrolled schools completed the study. The schools were in rural, semiurban, and urban locales spread across five geographic regions of the state. Participating schools reported a wide range of racial (see the results section) and economic diversity (19% to 86% economically disadvantaged) and academic achievement (5% to 81% passing state standardized tests) in their student populations (Michigan Department of Education, 2003). The comparison schools were matched to the intervention schools based on the proximity of the school districts, town population size and urbanicity, school size, grade levels taught, percentage of the students who were eligible for a free lunch, proportion of economically disadvantaged students, standardized test scores, teacher to student ratio, number of hours of PE required, and district expenditure per student. All of the fourth and fifth graders in the participating schools were invited to join the study.

Human Participant Protection

Consistent with Michigan Department of Education guidelines, both passive parent permission and student assent were required for a student to participate in the data collection activities. Students with a preexisting and diagnosed physical or medical condition that might have caused health problems during the motor skills and fitness tests (e.g., uncontrolled asthma) were identified by school personnel or parents and were excluded from the study. Study protocols and informed consent procedures were approved by the institutional review boards of ETR Associates and the Centers for Disease Control and Prevention.

Measures and Instruments

At each of the four data collection points, the participating students were asked to complete a survey and physical activity checklist and to be observed for three motor skill and four physical fitness assessments. The physical activity checklist was administered to reflect physical activity on the previous weekday. A pool of 25 data collectors was trained and supervised on a highly detailed data collection protocol; the majority of all of the assessments were completed by a core team of 11 data collectors, with an average of 23 years of experience in PE, fitness testing, and coaching (range = 2 to 40 years). Published measures with acceptable psychometric characteristics were used whenever possible. When such measures did not exist, new measures were developed for the study.

Demographic and Psychosocial Measures. Written student surveys were used to collect data on student demographic and psychosocial characteristics. Demographic measures included race, ethnicity, age, and sex. Psychosocial measures included a nine-item motor skill-specific self-efficacy scale (developed for this study; Cronbach's $\alpha = .72$), the six-item Perceived Physical Activity Competence Scale (adapted from Harter, 1982; Cronbach's $\alpha = .69$), a two-item parent involvement in PE measure (developed for this study; $r = .70$), a two-item peer support for physical activity measure (Sallis, n.d.; $r = .63$), and a two-item environmental support for physical activity measure (Sallis, n.d.; $r = .71$).

Physical Activity Level Measures. The Self-Administered Physical Activity Checklist (SAPAC; Sallis et al., 1996) was used to assess the students' self-reported physical activity level for the prior school day. From the SAPAC, we calculated the total minutes of physical activity, minutes of moderate to vigorous physical activity, the physical activity metabolic equivalent task (MET) score, and the weighted physical activity MET score (an MET score weighted by the student's self-reported activity intensity). The physical activity MET score represents the summed products of the frequency in minutes and MET value for each activity listed. The Compendium of Physical Activities (Ainsworth et al., 1993) was used to code physical activity intensity levels according to their respective METs. Ainsworth and colleagues (1993) define an MET as the ratio of work metabolic rate to a standard resting metabolic rate.

Motor Skills Measures. An observation rubric was developed with content experts to assess students' locomotor, posture, and object control motor skills, skill categories that were defined by the Michigan Physical Education Content Standards. Respectively, forehand strike, lift and carry, and leap were observed as indicators of each of these motor skills and were chosen based on their relationships to the NASPE content standards and for their ability to capture the developmental range of movement skills expected for youth aged 8 to 12. In addition, the students who possessed fundamental motor skills in each of the movement categories would be expected to demonstrate competence in the measured motor skills, even if they had not been taught that particular movement. To account for the possibility that students in the comparison arm may not have been taught the specific movements that are required by the motor skills measures, all of the students in both arms received training in the specific movements that were required before being observed and rated. Student training consisted of verbal explanations, demonstrations, and student practice. The rubric procedures were informed by the performance measures that were developed by van Beurden and colleagues (2003) for rating motor skill development among elementary school children.

The motor skills rubric guides observers to systematically analyze the form and function elements of each observed skill. Form elements included the specific physical movements that were required to complete the skill (e.g., strike ball at waist level for forehand strike). Function elements rated the achievement of an outcome (e.g., the number of hurdles cleared during leap). Each successfully demonstrated element received 1 point and was summed to create a composite score across form and function. Trained rubric content experts who had achieved acceptable interrater reliability (>.80 agreement with a referent expert rating) during practice observations used the rubric to assess motor skill performance for students in the study.

Fitness Level Measures. Four fitness tests from the Fitnessgram (Welk, Morrow, & Falls, 2001) battery were used to assess the students' fitness levels. The Progressive Aerobic Cardiovascular Endurance Run (PACER) was used to measure VO_2 max as an indicator of aerobic capacity, curl-ups were used to measure abdominal strength, push-ups were used to measure arm strength, and sit and reach was used to measure lower body flexibility. The raw scores of the PACER indicate the number of course laps that were completed in time with a recorded cadence; the raw scores were converted into VO_2 max scores through a computerized algorithm. Curl-ups and push-ups were measured as the number completed to a cadence with proper form. Sit and reach was measured in inches twice, one measure for each extended leg with the other leg folded.

Sample Size and Attrition

At the 16 schools, 93% of the fourth- and fifth-grade students with passive positive parent permission agreed to participate in the study. At baseline, 1,464 students completed the surveys, motor skills, and assessments (760 students in the intervention condition and 704 in the comparison condition).

At the last follow-up, 1,195 students completed the surveys and fitness and motor skills (600 students in the intervention condition and 595 students in the comparison condition). Of the original sample, 18% was lost to follow-up. Of the students missing at follow-up, 99% had moved away from their schools, as reported by school administrators.

Statistical Method

Outcome measures were compared between the two conditions using multilevel regression analyses that controlled for baseline differences between the conditions and the clustering effects within the schools. Three-level models were fit for the final analysis, where Level 1 was the data collection time point (t_2 to t_4), Level 2 was the student, and Level 3 was the school. Separate models were fit for each of the outcome measures. These models allowed for the estimation of the overall average intervention effects over time as well as individual follow-up time point effects.

For each outcome, the following model was analyzed: a multilevel regression model in which follow-up outcome variables (i.e., from t_2 to t_4) were the dependent variables, with the baseline outcome variable (t_1) as a covariate. Additional covariates were included in the outcome models if the following conditions were met: the covariate's distributions differed significantly across conditions at baseline ($p < .15$), the covariate was related to the outcome at baseline ($p < .15$), and the covariate's regression coefficient was significant at $p < .15$ in the final model, using the Wald test (Harrell, Lee, & Mark, 1996). Student race, ethnicity, and gender, parent involvement in PE, peer support for physical activity, and environmental support for physical activity were screened and kept as covariates, except student gender, which did not meet inclusion criteria. Linear regression models were used for all of the normally distributed continuous outcomes. Negative binomial models were used for skewed and count data. For each outcome, separate models were run for the fourth- and fifth-grade cohorts. Univariate descriptive and bivariate screening analyses were conducted using SPSS Version 14 (SPSS, Inc., 2005). MLwiN Version 2.02 (Rasbash, Browne, Healy, Cameron, & Charlton, 2005) was used for the multilevel analyses.

RESULTS

Baseline Data

The ethnic distribution of the students was 71.6% White, 13.5% African American, 6.2% American Indian, 5.4% Multiracial, and 3.3% Other (e.g., Asian, Pacific Islander, and Latino). The students ranged in age from 8 to 12 years ($M = 9.8$). Of the students, 49% were female. When looking at student demographic characteristics by grade across treatment condition, only race was significantly different by condition; proportionately more White students were in the comparison condition. For the fourth grade, there were 72.9% White students in the comparison condition versus 64.9% in the intervention

condition ($p = .07$, $\chi^2 = 5.48$, $df = 2$); for the fifth grade, there were 78.1% White students in the comparison condition versus 70.3% in the intervention condition ($p = .002$, $\chi^2 = 12.20$, $df = 2$).

Attrition Analysis

Multilevel analyses were run to determine whether significant differential attrition occurred between study arms across data collection time points by grade level for selected behavioral outcomes. Attrition analyses included the following baseline characteristics: gender, race/ethnicity, physical activity, and motor skills scores. There was significant differential attrition by race and baseline outcome scores for some measures at some time points.

In the treatment arm, African American students were more likely to be missing the t_4 data point for strike and lift and carry (fourth and fifth grade) and leap (fifth grade only) than were White students. In the comparison arm, African American students (fifth grade only) were more likely to be missing measures on strike (t_2) and lift and carry (t_3).

Fourth-grade students in the treatment arm who had reported fewer total minutes of physical activity at baseline were more likely to be missing this measure at t_2 . Fourth-grade students in the treatment arm who had reported lower physical activity MET scores at baseline were more likely to be missing this measure at t_2 and t_4 . Likewise, fifth-grade students in the comparison arm who had lower scores on lift and carry at baseline were more likely to be missing their follow-up scores at t_3 .

The source and consequences of these attrition patterns are unclear. African American students may have been more at risk for missing follow-up measures on lift and carry due to reactions to their lower baseline scores. Because race at baseline was unevenly distributed across treatment and comparison arms, a covariate for race at baseline was entered into each outcome multivariate model but was retained as significant only for a handful of models. The loss of students with lower physical activity performance scores at baseline can result in lower variation of that measure over time and a false rise in the aggregate mean score of the study arm in question. However, none of the attrition patterns affected only one treatment arm or extended across all significant follow-up time points of the multilevel models, thereby reducing the likelihood that the observed program impact was exclusively due to differential attrition.

Outcomes

For all of the outcomes, the overall average intervention effects over time are reported. Individual time point effects are reported only when they were present and the overall effects were not. Adjusted overall means are derived from the multilevel regression coefficient estimates of overall average group (intervention vs. control) differences across all three time points. They are provided to illustrate the magnitude of group differences on the outcome variable's scale rather than as precise point estimates.

Self-Efficacy and Competence (see Table 1). Fourth-grade cohort students exposed to EPEC reported greater levels of motor skill-specific self-efficacy compared to the students receiving standard PE curricula. A covariate for perceived environmental support for physical activity (i.e., having adequate and safe access to equipment and space) remained in the model for the fourth-grade cohort. No difference was observed across study arms for the Perceived Physical Activity Competence Scale score.

Table 1. Multilevel Regression Analyses of Adjusted Overall Average Intervention Effects Over Time for Measures of Psychosocial Constructs and Levels of Physical Activity (PA)

Outcome	Fourth-Grade Cohort				Fifth-Grade Cohort			
	Adjusted Overall Mean ^a	<i>p</i> Value	<i>n</i>	Effect Size ^b	Adjusted Overall Mean	<i>p</i> value	<i>n</i>	Effect Size
Motor skill-specific self-efficacy								
Intervention	3.50	.01	684	.13	3.57	.64	710	.02
Comparison	3.41				3.56			
Perceptions of PA competence								
Intervention	0.80	.60	644	.03	0.83	.39	675	-.04
Comparison	0.80				0.84			
Total minutes of PA								
Intervention	169.22	.04	636	.10	158.94	.64	643	-.02
Comparison	150.87				162.79			
Minutes of moderate to vigorous PA								
Intervention	100.23	.12	518	.09	89.33	.47	546	-.04
Comparison	87.30				94.81			
PA MET score								
Intervention	926.26	.01	636	.12	857.41	.59	643	-.03
Comparison	801.84				883.42			
Weighted PA MET score								
Intervention	955.95	.02	517	.14	842.75	.30	545	-.06
Comparison	804.11				903.57			

Note: MET = metabolic equivalent task.

a. Adjusted overall group mean estimates are *not* simple raw sample means. Rather, they are based on estimates from the analysis equations, thus allowing for the control of covariates. These estimates serve to illustrate the magnitude of the difference between treatment and control groups.

b. Effect size (Cohen, 1988; Rosenthal & Rosnow, 2008) represents the difference in the standardized mean for comparison versus intervention groups at follow-up relative to baseline, averaged over all three follow-up time points. Cohen (1988) defines small, medium, and large effect sizes for behavioral studies to be .2, .5, and .8, respectively.

Physical Activity (see Table 1). Compared to the students receiving standard PE, fourth-grade cohort students exposed to EPEC reported significantly greater total minutes of physical activity and energy expended during physical activity (as measured by physical activity MET scores). A covariate for peer support for physical activity remained in the model for the fourth-grade cohort. No significant intervention effects for physical activity were observed among the fifth-grade cohort.

Motor Skills (see Table 2). The students who were exposed to EPEC in both cohorts demonstrated statistically significantly ($p < .05$) higher skill levels over the 2-year study period for two of three motor skills, forehand strike and lift and carry, compared to the students receiving alternate PE curricula. For leap, intervention effects were statistically significant at $p < .05$ only for the fifth-grade cohort. A covariate for race remained in the model for forehand strike and lift and carry in the fourth-grade cohort, and a covariate for peer support for physical activity remained in the model for forehand strike in the fifth-grade cohort.

Table 2. Multilevel Regression Analyses of Adjusted Overall Average Intervention Effects Over Time for Measures of Motor Skills and Fitness Outcomes

Outcome	Fourth-Grade Cohort				Fifth-Grade Cohort			
	Adjusted Overall Mean ^a	<i>p</i> Value	<i>n</i>	Effect Size ^b	Adjusted Overall Mean	<i>p</i> Value	<i>n</i>	Effect Size
Forehand strike (rubric score)								
Intervention	4.02	<.001	635	.36	4.23	<.001	659	.40
Comparison	3.37				3.51			
Lift and carry (% correct) ^c								
Intervention	0.90	.02	636	.11	0.87	<.001	658	.17
Comparison	0.84				0.80			
Leap (% correct) ^c								
Intervention	0.93	.12	635	.08	0.95	.02	680	.11
Comparison	0.91				0.94			
Progressive Aerobic Cardiovascular								
Endurance Run (VO ₂ max)								
Intervention	43.27	.19	644	.06	42.07	.75	687	-.01
Comparison	43.00				42.19			
Curl-ups (count)								
Intervention	9.12	.85	651	.009	16.45	.94	682	.003
Comparison	8.85				16.28			
Push-ups (count)								
Intervention	3.97	.67	647	.02	5.10	.97	686	.002
Comparison	3.78				5.05			
Sit and reach, left leg (inches)								
Intervention	8.38	.94	646	-.004	8.56	.50	687	-.03
Comparison	8.40				8.70			
Sit and reach, right leg (inches)								
Intervention	8.53	.95	645	.003	8.60	.46	687	-.03
Comparison	8.51				8.75			

a. Adjusted overall group mean estimates are *not* simple raw sample means. Rather, they are based on estimates from the analysis equations, thus allowing for the control of covariates. These estimates serve to illustrate the magnitude of the difference between treatment and control groups.

b. Effect size (Cohen, 1988; Rosenthal & Rosnow, 2008) represents the difference in the standardized mean for comparison versus intervention groups at follow-up relative to baseline, averaged over all three follow-up time points. Cohen (1988) defines small, medium, and large effect sizes for behavioral studies to be .2, .5, and .8, respectively.

c. This motor skill had different numbers of elements by grade. For compatibility across cohorts and time, scores presented above are the percentage of elements correctly demonstrated (i.e., 0 to 1.0).

Fitness Outcomes (see Table 2). Overall, no significant intervention effect was observed for aerobic capacity (VO₂ max) as measured with the PACER (students in both of the study arms showed an overall decrease in VO₂ max over time). At one measurement time, the fourth-grade students exposed to EPEC showed a significantly greater aerobic capacity, but this was not sustained over the 2 years and was not observed in the fifth-grade cohort students. No significant intervention effects were observed in either cohort for either abdominal strength or arm strength measures (curl-ups and push-ups, respectively) or for lower body flexibility (sit and reach). The

students in both the intervention schools and the comparison schools showed an increase in the number of push-ups that were completed over the course of the study; results for the other fitness measures were mixed by grade and intervention arm. A covariate for peer support remained in the model for push-ups in the fourth-grade cohort, and a covariate for race remained in the model for sit and reach in the fifth-grade cohort.

DISCUSSION

The results suggest that EPEC was modestly more effective than standard PE curricula at increasing motor skill-specific self-efficacy, improving two of three indicators of motor skill performance and increasing self-reported levels of physical activity for the fourth-grade cohort. Only the motor skill outcome findings were significant for the fifth-grade cohort. The study was not of sufficient duration to determine whether EPEC students would maintain these improvements during adolescence and adulthood. These findings support and add to previous research that has indicated an association between motor skill development and overall physical activity levels of young people (Dishman et al., 2004; Wrotniak et al., 2006).

In addition, the EPEC curriculum appears to have had a similar impact on motor skill development through enhanced and motor skill-focused PE as those seen in the studies by Graf et al. (2005) and Karabourniotis et al. (2002). As with the van Beurden et al. (2003) study, this study demonstrated that a school-based PE curriculum can improve motor skill proficiency without loss of physical activity, indeed, in some cases, with improvement in self-reported physical activity levels. Also, this study provides evidence that a school-based PE curriculum can improve motor skill-specific efficacy and proficiency without loss of fitness levels compared to standard PE curricula requiring similar amounts of class time.

EPEC students did not become more physically fit than the other students during the 2 years of the study; this finding was expected given that the EPEC elementary curriculum focuses on the theoretical precursors to fitness, such as motor skill development and physical activity levels. For example, the EPEC teachers more frequently described the development of specific motor skills as *primary* lesson objectives than did teachers using the standard PE curricula represented in the comparison condition. It is unclear why the self-reported measures of physical activity levels and motor skill-specific self-efficacy were not significantly improved by exposure to EPEC among the fifth-grade cohort. One possible explanation is that the developmental changes affected self-reporting motivation or quality. Another explanation is that, compared to the fourth-grade cohort, the fifth-grade cohort was exposed to slightly fewer EPEC lessons that were optimally matched to their developmental stage during the sixth grade, potentially reducing EPEC's effectiveness.

EPEC provides a unique opportunity to examine the balance among instruction in motor skills, physical activity, and fitness. As noted earlier, the optimal balance among these content areas has not been previously established for either short-term outcomes (e.g., motor skill confidence, motor skill level, and physical activity level) or long-term outcomes (e.g., physical activity and fitness levels over the lifespan). EPEC's weighting of instructional material, in terms of primary learning objectives for each lesson, across the Michigan and NASPE content standards may partially explain its success with certain outcomes in this evaluation. Another of EPEC's strengths may be its focus on teaching and learning progressions that strongly incorporate assessment and feedback.

These design factors may be more important for improvement in motor skills and physical activity levels than for fitness.

Study Limitations and Strengths

Because we used a quasiexperimental design, we were able to control only for measured differences between the study conditions; other uncontrolled differences that were related to outcomes may have existed. Some study data, such as data from the SAPAC, are self-reported and could not be verified. The limitations of such self-reports have been described elsewhere (Ainsworth, Montoye, & Leon, 1994) and have included social desirability bias (Warnecke et al., 1997), high cognitive demands (Baranowski, 1988), and context specificity (Sallis & Saelens, 2000). Investigators using the SAPAC and similar measures have noted that young students tend to significantly overestimate their physical activity levels, especially vigorous activity, through self-report data compared to data from heart rate monitors and physical activity monitors (e.g., pedometers and accelerometers; McMurray et al., 2004; Sallis & Saelens, 2000). If this error is stable, then such measures may still be useful for the purposes of assessing change. Our curriculum coverage data are based solely on teacher self-report and not on classroom observations, so content coverage descriptions and comparisons across the arms may be imprecise. Also, we did not examine copies of the written curricula from the comparison group teachers, although written lessons were used less than one half of the time. It is possible that comparison curricula were fairly similar to EPEC, given that all instructors in the study described their PE curricula as meeting state PE curriculum standards.

Our choice of observed motor skills may have favored students in the intervention condition over those in the comparison condition because of the greater familiarity with those particular motor skills among EPEC students. In anticipation of this possibility, we selected skill indicators that all of the fourth-grade students instructed with standards-based curricula should have been able to competently perform with minimal introduction and practice.

Although the motor skills rubrics were developed based on current accepted standards for motor skills performance, we were limited in our ability to ascertain validity and reliability. We could not find any published criterion measures that measured both form and function of the three selected motor skills in developmentally appropriate ways for fourth- and fifth-grade students with which to compare our rubrics. As such, the rubrics do not have established psychometric properties.

Although interrater reliability was high, because of resource limitations student motor skill performance was rated only “live” and not via videotaping. The inability of the data collectors to review performance demanded vigilance and attention to detail over the entire testing period; fatigue and variability may have affected scoring. However, data collectors were blinded to the study arm, so this limitation would equally affect intervention and comparison schools.

We argue that a strength of this study is its naturalistic design and suggest that these results could be obtained in similar real-world settings. We invited existing PE teachers to join the study and asked them to document how and what they taught. A balance appeared between correlates of curriculum fidelity and reported curriculum modifications and adaptations for teachers in both study arms. Several factors that are often associated with higher levels of implementation fidelity were reported by the EPEC teachers, such as having a well-organized written curriculum guide, teacher training, administrator support, adequate educational supplies, and adequate teaching environments.

Implications for Practice

EPEC was more effective than standard PE curricula at improving motor skill performance (fourth- and fifth-grade cohorts) and at increasing self-reported motor skill-specific self-efficacy and levels of physical activity (fourth-grade cohort only). These findings contribute to the limited research on the impact of PE on the development of motor skills. It is unclear whether EPEC students will maintain these improvements during adolescence and adulthood. Some have argued that, historically, PE has focused on motor skill acquisition largely in the context of competitive team sports but that the most common adult physical activities are individual based (e.g., walking, bicycling, golf), not team endeavors (Trost, 2004). EPEC appears to represent a new generation of PE curricula in which motor skill acquisition and performance occur in the context of learning both individual and team physical activities, skills that may ultimately enable lifelong fitness.

References

- Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs Jr., D. R., Montoye, H. J., Sallis, J. F., et al. (1993). Compendium of physical activities: Classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*, *25*, 71-80.
- Ainsworth, B. E., Montoye, H. J., & Leon, A. S. (1994). Methods of assessing physical activity during leisure and work. In C. Bouchard, R. J. Shephard, & T. Stephens (Eds.), *Physical activity, fitness, and health: International proceedings and consensus statement* (pp. 146-159). Champaign, IL: Human Kinetics.
- Baranowski, T. (1988). Validity and reliability of self-report of physical activity: An information processing perspective. *Research Quarterly for Exercise and Sport*, *59*, 314-327.
- Centers for Disease Control and Prevention. (2006). *Promoting better health for young people through physical activity and sports: A report to the president*. Atlanta, GA: Author.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Dishman, R. K., Motl, R. W., Saunders, R., Felton, G., Ward, D. S., Dowda, M., et al. (2004). Self-efficacy partially mediates the effect of a school-based physical-activity intervention among adolescent girls. *Preventive Medicine*, *38*, 628-636.
- Graf, C., Koch, B., Falkowski, G., Jouck, S., Christ, H., Stauenmaier, K., et al. (2005). Effects of a school-based intervention on BMI and motor abilities in childhood. *Journal of Sports Science and Medicine*, *4*, 291-299.
- Harrell, F. E., Jr., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics Medicine*, *15*, 361-387.
- Harter, S. (1982). The perceived competence scale for children. *Child Development*, *53*, 87-97.
- Karabourniotis, D., Evagelinou, C., Tzetzis, G., & Kourtessis, T. (2002). Curriculum enrichment with self-testing activities in development of fundamental movement skills of first-grade children in Greece. *Perceptual and Motor Skills*, *94*, 1259-1270.
- McMurray, R. G., Ring, K. B., Treuth, M. S., Welk, G. J., Pate, R. R., Schmitz, K. H., et al. (2004). Comparison of two approaches to structured physical activity surveys for adolescents. *Medicine and Science in Sports and Exercise*, *36*, 2135-2143.
- Michigan Department of Education. (1998). *Physical education content standards and benchmarks*. Lansing: Author.
- Michigan Department of Education. (2003). *Data on school demographics and achievement by district*. Lansing: Author.

- Michigan's Exemplary Physical Education Curriculum Project. (2001). *EPEC lessons—Grades K, 1, 2, 3, 4, 5, user's manual and teaching/learning progressions*. Lansing: Michigan Fitness Foundation.
- National Association for Sport and Physical Education. (2004). *Moving into the future: National standards for physical education*. Reston, VA: Author.
- Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of physical activity to fundamental movement skills among adolescents. *Medicine and Science in Sports and Exercise*, 33, 1899-1904.
- Rasbash, J., Browne, W., Healy, M., Cameron, B., & Charlton, C. (2005). *MLwiN version 2.02, multilevel models project*. London: University of London, Institute of Education.
- Reed, J. A., Metzker, A., & Phillips, D. A. (2004). Relationships between physical activity and motor skills in middle school children. *Perceptual and Motor Skills*, 99, 483-494.
- Rosenthal, R., & Rosnow, R. (2008). *Essentials of behavioral research: Methods and data analysis* (3rd ed.). Boston: McGraw-Hill.
- Sallis, J. F. (n.d.). *Amherst Health and Activity Student Survey*. Retrieved April 1, 2003, from <http://www.drjamesallis.sdsu.edu/measures.html>
- Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: Status, limitations, and future directions. *Research Quarterly for Exercise and Sport*, 71, S1-S14.
- Sallis, J. F., Strikmiller, P. K., Harsha, D. W., Feldman, H. A., Ehlinger, S., Stone, E. J., et al. (1996). Validation of Interviewer- and Self-Administered Physical Activity Checklists for fifth grade students. *Medicine and Science in Sports and Exercise*, 28, 840-851.
- SPSS, Inc. (2005). *SPSS for Windows*. Chicago: Author.
- Trost, S. G. (2004). School physical education in the post-report era: An analysis from public health. *Journal of Teaching in Physical Education*, 23, 318-337.
- van Beurden, E., Barnett, L. M., Zask, A., Dietrich, U. C., Brooks, L. O., & Beard, J. (2003). Can we skill and activate children through primary school physical education lessons? "Move it Groove it"—A collaborative health promotion intervention. *Preventive Medicine*, 36, 493-501.
- Warnecke, R. B., Johnson, T. P., Chavez, N., Sudman, S., O'Rourke, D. P., Lacey, L., et al. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology*, 7, 334-342.
- Welk, G. J., Morrow, J. R., Jr., & Falls, H. B. (2001). *Fitnessgram reference guide*. Dallas, TX: Cooper Institute.
- Wrotniak, B. H., Epstein, L. H., Dorn, J. M., Jones, K. E., & Kondilis, V. A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics*, 118, 1758-1765.