©Journal of Sports Science and Medicine (2016) **15,** 335-343
http://www.jssm.org

**Research article**

# Association of Quality Physical Education Teaching with Students' Physical Fitness

**Weiyun Chen [1] ✉, Steve Mason [1], Andrew Hypnar [2] and Austin Hammond-Bennett [3]**
[1] University of Michigan, School of Kinesiology, Ann Arbor, MI, USA; [2] Fairfax County Public Schools, Fairfax, VA, USA; [3] Livonia Public Schools, Livonia, MI, USA

**Abstract**
This study examined the extent to which four essential dimensions of quality physical education teaching (QPET) were associated with healthy levels of physical fitness in elementary school students. Participants were nine elementary PE teachers and 1, 201 fourth- and fifth-grade students who were enrolled in nine elementary schools. The students' physical fitness were assessed using four FITNESSGRAM tests. The PE teachers' levels of QPET were assessed using the Assessing Quality Teaching Rubrics (AQTR). The AQTR consisted of four essential dimensions including Task Design, Task Presentation, Class Management, and Instructional Guidance. Codes were confirmed through inter-rater reliability (82.4% and 84.5%). Data were analyzed through descriptive statistics, multiple R-squared regression models, and independent sample *t*-tests. The four essential teaching dimensions of QPET were significantly associated with the students' cardiovascular endurance, muscular strength and endurance, and flexibility. However, they accounted for relatively low percentage of the total variance in PACER test, followed by Curl-up test, while explaining very low portions of the total variance in Push-up and Trunk Lift tests. This study indicated that the students who had experienced high level of QPET were more physically fit than their peers who did not have this experience in PACER and Curl-up tests, but not in Push-up and Trunk lift tests. In addition, the significant contribution of the four essential teaching dimensions to physical fitness components was gender-specific. It was concluded that the four teaching dimensions of QPET were significantly associated with students' health-enhancing physical fitness.

**Key words:** Quality teaching, teaching assessment, fitness measures, Fitnessgram test, healthy fitness zone.

## Introduction

Health-related physical fitness provides physical foundations necessary for enjoyable and successful physical activity engagement in children and adolescents (Stodden et al., 2008). A healthy level of cardiorespiratory endurance is positively associated with a healthy cardiovascular profile in children and adolescents. It is inversely related to obesity, cardiovascular disease factors, and clustering of metabolic risk factors (Ortega et al., 2008). Improvements in muscular strength and endurance and flexibility have a positive effect on skeletal health (Ortega et al., 2008). Health-related physical fitness is a key indicator of health outcomes (Ortega et al., 2008).

Achieving and maintaining a health-enhancing level of fitness is one of the physical education content standards for school-aged children to meet (National Association for Sport and Physical Education [NASPE], 2014). Quality physical education (QPE), as a core component of the Comprehensive School-based Physical Activity Program (CSPAP), is a crucial vehicle for promoting physical fitness of school-aged children (Centers for Disease Control and Prevention [CDC], 2011). Meaningful content and appropriate instructions are essential parts of QPE (AAHPERD, 2013; Erwin et al., 2013). QPE incorporates fitness activities into a regular PE class and provides a variety of physical activities that are fitness-enhancing and developmentally appropriate for students. Appropriate instructional practices are promoted in QPE so that students have adequate opportunities to engage in moderate to vigorous physical activities for enhancing their physical fitness levels in a regular PE lesson. Implementation of QPE teaching (QPET) consists of four essential dimensions including Task Design, Task Presentation, Class Management, and Instructional Guidance (Ball and Forzani, 2009; Chen et al., 2011; 2014; Lampter and Graziani, 2009; Rink, 2006).

Task Design refers to how the teacher designs and organizes a series of learning tasks for their students to accomplish during a lesson (Rink, 2006). Learning tasks that are developmentally appropriate and maximally engaging are critical to ensure students have successful learning experiences, ample learning opportunities, and maximum participation (Chen et al., 2011; NASPE, 2010; Rink, 2006). Task Presentation refers to how the teacher presents learning tasks to students (Rink, 2006). Essential components of Task Presentation reflect how well the teacher precisely and accurately explains the learning task, demonstrates key features of the learning task, and relates information to students' life experiences (Ball and Forzani, 2009; Chen et al., 2011; Rink, 2006). Class Management implies how the teacher groups students, distributes physical learning materials/equipment, arranges physical layouts, locates students into working areas, and reinforces class norms and rules (Chen et al., 2011; Rink, 2006). Instructional Guidance refers to how the teacher observes and analyses students' task performances, adjusts the complexities of the task, directs students' focus to the task, and provides tailored instructional guidance (Ball and Forzani, 2009; Chen et al., 2011; Rink, 2006). In short, how well the teacher implements each of the four essential dimensions collectively contributes to the overall quality of instructional practices (Ball and Forzani, 2009; Chen et al., 2011; 2014; Rink, 2006).

To help students achieve and maintain health-enhancing physical fitness, there has been an increasing

Received: 18 November 2015 / Accepted: 08 April 2016 / Published (online): 23 May 2016

call for an implementation of QPET (CDC, 2011; NASPE, 2010). However, the degree to which QPET promotes students maintaining and enhancing healthy levels of physical fitness remains largely unexplored. Given the essential role of physical fitness in developing and building a physically active lifestyle throughout childhood, adolescence, and into adulthood, CDC (2011) recommends schools should conduct physical fitness testing in order to determine how well students are able to maintain health-enhancing physical fitness.

This study, thus, aimed at examining the extent to which QPET were associated with achieving a healthy level of physical fitness in elementary school students. Two research hypotheses were: (a) four essential dimensions of QPET were significantly associated with physical fitness of elementary school students, and (b) the associations between the four essential dimensions of QPET and physical fitness were differed by gender. The significance of this study lies in providing empirical evidence for how QPET is linked to students achieving desired learning outcomes defined by the NASPE content standard 3 (NASPE, 2014). Also, this study adds to literature on QPET linking to physical fitness in elementary school students.

## Methods

### Participants and research settings

In this study, participants were nine elementary physical education teachers and 1,201 fourth-grade students (n = 573, 325 boys vs. 248 girls) and fifth-grade students (n = 627, 335 boys vs. 292 girls) with one student did not identify his grade level at nine elementary schools. All nine teachers were Caucasian. At the time of this study, five female and four male teachers' teaching experiences varied from 6 years to 26 years with ages ranging from 33-years old to 55-years old. The typical physical education class had 18 to 28 students.

Approval of conducting this study was granted by the university institutional review board for human subject research and the school district. The consent form was signed by each physical education teacher, indicating his/her voluntary participation in this study. Also, the consent form was signed by the student's parent /guardian to permit his/her child to participate in this study. In addition, the assent form was offered for the students to decide whether or not they wanted to participation in this study, although their parent/guardian granted permission for them to participation in this study.

### FITNESSGRAM tests

To determine the extent to which the students achieved the NASPE content standard 3 (achieves and maintains a health-enhancing level of physical fitness), the FITNESSGRAM test was used to measure levels of students' health-enhancing physical fitness. The FITNESSGRAM test is designed to assess five components of health-related fitness, including cardiovascular endurance, muscular strength and endurance, flexibility, and body composition through a variety of test items (Meredith and Welk, 2007).

To help PE teachers objectively conduct FITNESSGRAM tests with their students, one two-hour workshop was held to mainly train the PE teachers on administering the Progressive Aerobic Cardiovascular Endurance Run (PACER) test. Another two-hour workshop was conducted to train the PE teachers on administering the Curl-ups, Push-ups, and Trunk-Lift tests. During the two separate workshops, the PE teachers learned the testing directions, protocols, recording sheet, class organizations, and criteria for healthy and un-healthy fitness zone corresponding to a specific age and gender for each fitness test.

During the first two weeks of May, each trained PE teacher administered four FITNESSGRAM test items to their $4^{th}$- and $5^{th}$- grade students during their regular physical education lessons. The tests included: (a) 15-meter version of the PACER for cardiovascular endurance, (b) Curl-up test for abdominal muscular strength and endurance, (c) Push-up test for upper body strength and endurance, and (d) Trunk Lift for trunk extensor strength and flexibility.

The "FITNESSGRAM Standards for Healthy Fitness Zone for Boys" (Meredith and Welk, 2007, p. 61) and the "FITNESSGRAM Standards for Healthy Fitness Zone for Girls" (Meredith and Welk, 2007, p. 62) were used to determine whether or not a student's score on each test was placed into the Healthy Fitness Zone (HFZ). The HFZ uses age- and gender-specific criteria for each test item defined (Meredith and Welk, 2007). The 8.4 version of the FITNESSGRAM test software was used to record the testing results.

### Assessment of QPET

*Video-recording lessons:* The first investigator video-recorded 30 PE lessons taught by nine PE teachers to their students ($M_{\text{lessons / teacher}} = 3.33$, SD = 0.5) throughout one school year. Prior to the video-recording in the beginning of the school year, the investigator used the doodle meeting calendar to schedule the video-recording date and time blocks based on the teacher's preferences. To avoid interruption of a regular PE program, the investigator invited the teacher to decide what specific PE content he/she would plan to teach on the scheduled video-recording date and time. To video-record a lesson, the investigator placed a camcorder in an unobtrusive corner of the gymnasium to avoid interfering with the teaching. The teacher wore a wireless microphone and the voice transmitter was attached to the digital camcorder in order to capture the teacher's and the students' voices. The investigator adjusted the camcorder's angles and zoomed in and out if necessary to ensure the teacher and his/her students were in view. The lesson was digitally recorded when the teacher started his/her teaching and the recording was stopped when the teacher dismissed the class.

*Assessing quality teaching Rubrics (AQTR):* The AQTR was designed and validated to assess levels of teachers' QPET in a live PE lesson or a digitally recorded PE lesson (Chen et al., 2011; 2012). AQTR included four essential dimensions such as Task Design, Task Presentation, Class Management, and Instructional Guidance with 17 subsumed teaching components. The Task Design is

composed of three components: Developmental Appropriateness, Maximum Participation, and Progression. Task Presentation includes five components: Clarity and Accuracy, Linking Prior Knowledge, Demonstration, Learning Cues, and Checking for Understanding. Class Management is comprised of Gaining Attention, Equipment Distribution, Grouping Students, and Transition. Instructional Guidance contains four components: Monitoring, Adjusting the Task, Reflection, General Feedback, and Specific Feedback. The performance indicator of each teaching component was defined on a 3-point rating scale to identify a gradation of the quality of instructional practices. For example, a rating of "3" indicated that the teacher fully demonstrated the criteria of quality instructional practices in each teaching component. A rating of "2" indicated the teacher in some degree demonstrated the criteria of quality instructional practices. A rating of "1" indicated that the teacher did not demonstrate the criteria of quality instructional practices. Also an "n/a" indicated that the specific teaching component was not applicable to a given teaching episode.

*Coding the video-recorded lessons:* Firstly, four investigators spent at minimum 15 hours studying AQTR in terms of the performance indicators of each teaching component on each of the 3-point rating scales and the coding protocols. Then, they used the coding protocols to practice assessment of four video-recorded lessons which were randomly selected from the pool of the video-recorded lessons. The coding protocols included (a) for each task cycle, observing what specific task the teacher presented to his/her students, how the teacher organized the class for performing the task, and how the teacher interacted with the students during their task engagement each teaching episode; (b) coding the teacher's instructional practices of each teaching component under the Task Design, Task Presentation, Class Management, and Instructional Guidance using a 3-point rating scale; and (c) writing comments down in the Open Comment portion of the assessment; and (d) evaluators could rewind the tape whenever necessary.

Next, four investigators were paired-up and began to code two randomly selected video-recorded lessons to check the inter-rater reliability (IR). While watching each video-recorded lesson together, each pair independently coded each lesson with the AQTR Assessment Sheet. The inter-rater reliability of the coded lessons was examined by checking each investigator's coding results using the formula: % IR = [numbers of agreement ÷ (numbers of agreement + numbers of disagreement)] * 100 (van der Mars, 1989). The inter-rater reliability of the two coded lessons was 82.4% and 84.5%.

After meeting the satisfied inter-rater reliability, the four investigators began to officially code the 30 video-recorded lessons with the AQTR assessment sheet using the coding protocols. The two investigators watched each taped lesson together, but each pair independently coded each taped lesson. The alpha reliability coefficients values of the four dimensions and the total scale of AQTR were .87, .88, .83, .89, and .91, indicating a high degree of measurement reliability.

### Data analysis

To examine levels of the students' physical fitness, descriptive statistics such as mean, standard deviation, and percentage of meeting the HZF for each test were calculated. To examine the extent to which the four essential teaching dimensions of QPET were associated with students' four physical fitness tests, four multiple R-Squared linear regression models were conducted using Weighted Least Squares Regression-Weighted by school method. To classify the 30 coding lessons into two levels of quality teaching, a mean score of overall QPET was calculated to classify the 30 coded lessons into two levels of overall QPET. Last, to examine if there was a significant mean difference on each physical fitness test between the two levels of overall QPET, an independent sample *t*-test was conducted with the data of each fitness test. A significant level of $p < 0.05$ was set for all statistical methods. All statistical analyses were conducted by means of IBM SPSS statistics version 22 for windows.

## Results

### Descriptive statistics of physical fitness tests

Table 1 presents the descriptive statistics of each physical fitness test by grade and gender. To meet the Healthy

**Table 1. Descriptive statistics of FITNESSGRAM tests by grade and gender.**

| Fitness Test | | | M | SD | Skewness | Kurtosis | % HFZ |
|---|---|---|---|---|---|---|---|
| **PACER** | 4th grade | Boys (325) | 32.18 | 17.08 | .526 | -.715 | 50.8 |
| | | Girls (248) | 24.59 | 13.58 | 1.349 | 1.502 | 97.2 |
| | 5th grade | Boys (325) | 35.90 | 18.07 | .359 | -1.004 | 57.7 |
| | | Girls (248) | 30.35 | 17.61 | .762 | -.708 | 67.5 |
| **Curl-up** | 4th grade | Boys (325) | 28.41 | 21.12 | .858 | -.351 | 75.4 |
| | | Girls (248) | 28.35 | 19.99 | .964 | -.079 | 82.7 |
| | 5th grade | Boys (325) | 35.44 | 22.53 | .562 | -.845 | 82.4 |
| | | Girls (248) | 32.94 | 21.23 | .536 | -.751 | 75.5 |
| **Push-up** | 4th grade | Boys (325) | 13.63 | 10.14 | 2.606 | 12.252 | 78.5 |
| | | Girls (248) | 11.94 | 9.17 | 2.295 | 9.834 | 71.0 |
| | 5th grade | Boys (325) | 14.44 | 10.76 | 2.126 | 6.686 | 78.3 |
| | | Girls (248) | 11.78 | 9.73 | 2.391 | 8.927 | 69.5 |
| **Trunk lift** | 4th grade | Boys (325) | 10.68 | 2.04 | -1.906 | 3.650 | 86.5 |
| | | Girls (248) | 10.79 | 1.95 | -1.967 | 4.354 | 87.5 |
| | 5th grade | Boys (325) | 11.00 | 1.74 | -2.001 | 4.315 | 89.0 |
| | | Girls (248) | 11.24 | 1.69 | -2.2895 | 9.525 | 91.4 |

**Table 2. A multiple regression model of four teaching dimensions predicting PACER test.**

| | R | $R^2$ | *F* | *df* | *Sig.* | *Beta* | *t* | *Sig* |
|---|---|---|---|---|---|---|---|---|
| **Total** | **Model** | .122 | 41.69 | 4, 1196 | .000 | | | |
| | **Task Design** | | | | | .13 | 2.85 | .005 |
| | **Task Presentation** | | | | | .37 | 10.85 | .000 |
| | **Class Management** | | | | | 1.175 | -4.06 | .000 |
| | **Instructional Guidance** | | | | | 1.057 | -1.308 | .000 |
| **Boys** | **Model** | .097 | 17.50 | 4, 655 | .000 | | | |
| | **Task Design** | | | | | .12 | 1.89 | .059 |
| | **Task Presentation** | | | | | .333 | 7.15 | .000 |
| | **Class Management** | | | | | -.106 | -1.76 | .080 |
| | **Instructional Guidance** | | | | | -.100 | -1.59 | .111 |
| **Girls** | **Model** | .173 | 27.95 | 4, 535 | .000 | | | |
| | **Task Design** | | | | | .144 | 2.23 | .026 |
| | **Task Presentation** | | | | | .418 | 8.63 | .000 |
| | **Class Management** | | | | | -.256 | 14.20 | .000 |
| | **Instructional Guidance** | | | | | -.007 | -.12 | .904 |
| **4th grade** | **Model** | .127 | 16.52 | 5, 567 | .000 | | | |
| | **Task Design** | | | | | .212 | 3.15 | .000 |
| | **Task Presentation** | | | | | .271 | 5.77 | .000 |
| | **Class Management** | | | | | -.119 | -1.83 | .067 |
| | **Instructional Guidance** | | | | | -.222 | -3.65 | .000 |
| **5th grade** | **Model** | .215 | 34.06 | 5, 621 | .000 | | | |
| | **Task Design** | | | | | .036 | .60 | .550 |
| | **Task Presentation** | | | | | .426 | 9.34 | .000 |
| | **Class Management** | | | | | -.171 | -3.09 | .002 |
| | **Instructional Guidance** | | | | | .062 | 1.03 | .305 |

Fitness Zone (HFZ) on PACER test, a 10-year-old boy should run 30-80 laps and 11-year-old boy should run 30-94 laps. In contrast, a 10-year-old girl should run 9-54 laps and 11-year-old girl should run 19-54 laps (Meredith and Welk, 2007). As seen in table 1, both fourth-grade (mean age = 10-year-old) and fifth-grade (mean age = 11-year-old) boys and girls on average met the HFZ for cardiovascular endurance (Meredith and Welk, 2007). Of the students who completed the PACER test, 50.8% of fourth-grade boys and 57.5% of fifth-grade boys reached the HFZ, while 97.2% of fourth-grade girls and 67.5% of fifth-grade girls met the HFZ.

The HFZ on Curl-up test for boys and girls at age of 10 should perform 12-24 and 12-26 curl-ups, respectively, while for boys and girls at age of 11 should perform 15-28 curl-ups and 15-29 curl-ups, respectively (Meredith and Welk, 2007). As shown in table 1, both fourth-grade and fifth-grade boys' and girls' average numbers of curl-ups were beyond the high end of the HFZ. In addition, 75.4% of fourth-grade boys and 82.4% of fifth-grade boys achieved the HFZ, 82.7% of fourth-grade girls and 75.5% of fifth-grade girls reached the HFZ for curl-up test.

To reach the HFZ for push-up, boys at age of 10 should perform 7-20 push-ups and at age of 11 should perform 8-20 push-ups, while girls at age of 10 and 11 should perform 7-15 push-ups (Meredith and Welk, 2007). As presented in table 1, both fourth-grade and fifth-grade boys and girls on average met the HFZ in Push-up test. 78.5% of fourth-grade boys and 78.3% fifth grade boys met the HFZ. In contrast, 71% of fourth-grade girls and 69.5% of fifth-grade girls achieved the HFZ for push-up test.

For Trunk Lift test, lifting the upper body 9-12 inches off the floor from the prone position is to meeting the HFZ for both boys and girls at the age of 10 and 11 years old. As shown in table 1, both fourth-grade and fifth-grade boys and girls on average met the HFZ in Trunk Lift test. 86.5% of fourth-grade boys and 89% of fifth-grade boys were in HFZ. Similarly, 87.5% of fourth-grade girls and 91.4% of fifth-grade girls were in the HFZ for Trunk Lift test.

## Association of QPET with physical fitness

The regression model consisting of four independent variables including Task Design, Task Presentation, Class Management, and Instructional Guidance and one dependent variable, the number of laps completed on PACER test was conducted for the total sample, for boys, for girls, for fourth-grade, and for fifth-grade students at a time. As seen in Table 2, the four essential teaching dimensions were significantly associated with PACER test for the total sample ($F = 41.691$, $p < 0.01$), for boys ($F = 17.495$, $p < 0.01$), for girls ($F = 27.949$, $p < 0.01$), for fourth-grade students ($F = 16.517$, $p < 0.01$), and for fifth-grade students ($F = 34.063$, $p < 0.01$). The four essential dimensions accounted for 12.2% of the variance in the PACER test for the total sample, 9.7% for boys, 17.3% for girls, 12.7% for fourth-grade students, and 21.5% for fifth-grade students. Furthermore, Task Presentation obtained the largest β weight ($\beta = 0.37$, $p < 0.01$) for the total sample, ($\beta = 0.33$, $p < 0.01$) for boys, ($\beta = 0.42$, $p < 0.01$) for girls, ($\beta = 0.27$, $p < 0.01$) for fourth-grade students, and ($\beta = 0.43$, $p < 0.01$) for fifth-grade students. The results indicated that Task Presentation made the largest contribution to prediction of the PACER test, while holding all other three teaching dimensions constant. In addition, the results of β weight revealed that all other three teaching dimensions were statistically significant contributors to the students' PACER test for the total sample

In this regression model, the four teaching dimen-

sions were specified as independent variables and the number of Curl-ups performed as a dependent variable (see Table 3). The regression model was run separately for the total sample, for boys, for girls, for fourth-grade, and for fifth-grade students at a time. The results indicated that the four essential teaching dimensions significantly predicted Curl-ups for the total sample ($F = 19.394$, $p < 0.01$), for boys ($F = 16.249$, $p < 0.01$), for girls ($F = 5.655$, $p < 0.01$), for fourth-grade students ($F = 5.603$, $p < 0.01$), and for fifth-grade students ($F = 12.611$, $p < 0.01$). The four essential teaching dimensions accounted for 6.1% of the total variance in Curl-up test for the total sample, 9% for boys, 4.1% for girls, 4.7% for fourth-grade students, and 9.2% for fifth-grade students. Class Management had the largest β weight for the total sample ($\beta = 0.32$, $p < 0.01$), for boys ($\beta = 0.22$, $p < 0.01$), for girls ($\beta = 0.21$, $p < 0.01$), for fourth-grade students ($\beta = 0.26$, $p < 0.01$), and for fifth-grade students ($\beta = 0.39$, $p < 0.01$), while holding the other three teaching dimensions constant. The results of β weight indicated that Class Management made the most significant contribution to prediction of the students' curl-up test.

As seen in Table 4, this regression model containing four teaching dimensions as predictors and the number of Push-ups performed as dependent variable was conducted separately for the total sample, for boys, for girls, for fourth-grade, and for fifth-grade students at a time. The results revealed that the four teaching dimensions significantly predicted Push-ups for the total sample ($F = 9.537$, $p < 0.01$), for boys ($F = 8.837$, $p < 0.01$), for girls ($F = 4.568$, $p < 0.01$), for fourth-grade students ($F = 3.197$, $p < 0.01$), and for fifth-grade students ($F = 7.388$, $p < 0.01$). The four teaching dimensions explained 3.1% of the total variance in Push-up test for the total sample, 5.1% for boys, 3.3% for girls, 2.2% for fourth-grade, and 4.5% for fifth-grade students. Further, the results of β weight indicated that Task Design made a significant contribution to the multiple regression model for the total sample ($\beta = 0.14$, $p < 0.01$), for boys ($\beta = 0.18$, $p < 0.01$), and for fourth-grade students ($\beta = 0.18$, $p < 0.01$), but not for girls and fifth-grade students, while holding all other three dimensions constant. In addition, Instructional Guidance obtained the largest β weight with a negative sign for the total sample ($\beta = -0.28$, $p < 0.01$), for boys ($\beta = -0.34$, $p < 0.01$), for girls ($\beta = -0.23$, $p < 0.01$), for fourth-grade students ($\beta = -0.21$, $p < 0.01$), and for fifth-grade students ($\beta = -0.35$, $p < 0.01$). The other two teaching dimensions did not have significant β weight.

In Table 5, the regression model specifying four teaching dimensions as independent variables and Trunk Lift test as dependent variable was conducted separately the total sample, for boys, for girls, for fourth-grade, and for fifth-grade students at a time. The multiple regression model indicated that the four essential teaching dimensions significantly predicted Trunk Lift for the total sample ($F = 4.657$, $p < 0.01$), for boys ($F = 7.287$, $p < 0.01$), for girls ($F = 3.483$, $p < 0.01$), for fourth-grade students ($F = 2.869$, $p < 0.05$), and for fifth-grade students ($F = 4.550$, $p < 0.01$). The four teaching dimensions explained 1.5 % of the total variance in Trunk Lift test for the total sample, 4.3% for boys, 2.5% for girls, 2.0% for fourth-grade, and 2.8% for fifth-grade students. The results of β weight revealed that Instructional Guidance made the significant contribution to prediction of the total samples' trunk lift test ($\beta = 0.17$, $p < 0.01$), boys' trunk lift test ($\beta = 0.34$, $p < 0.01$), fourth-grade students' trunk lift test ($\beta = 0.20$, $p < 0.01$), and fifth-grade students' trunk lift test ($\beta = 0.14$, $p < 0.05$) but the other three teaching dimensions did not.

**Table 3. A multiple regression model of four teaching dimensions predicting CURL-UP test.**

| | R | $R^2$ | *F* | *df* | *Sig.* | *Beta* | *t* | *Sig* |
|---|---|---|---|---|---|---|---|---|
| **Total** | **Model** | .061 | 19.39 | 4, 1196 | .000 | | | |
| | **Task Design** | | | | | -.106 | -2.24 | .025 |
| | **Task Presentation** | | | | | .060 | 1.71 | .088 |
| | **Class Management** | | | | | .319 | 7.16 | .000 |
| | **Instructional Guidance** | | | | | -.050 | -1.10 | .273 |
| **Boys** | **Model** | .090 | 16.25 | 4, 655 | .000 | | | |
| | **Task Design** | | | | | -.073 | -1.13 | .260 |
| | **Task Presentation** | | | | | -.037 | -.80 | .425 |
| | **Class Management** | | | | | .416 | 6.87 | .000 |
| | **Instructional Guidance** | | | | | -.142 | -2.26 | .024 |
| **Girls** | **Model** | .041 | 5.66 | 4, 535 | .000 | | | |
| | **Task Design** | | | | | -.136 | -1.960 | .051 |
| | **Task Presentation** | | | | | .085 | 1.63 | .104 |
| | **Class Management** | | | | | .209 | 3.18 | .002 |
| | **Instructional Guidance** | | | | | .043 | .66 | .510 |
| **4th grade** | **Model** | .047 | 5.60 | 5, 567 | .000 | | | |
| | **Task Design** | | | | | -.130 | -1.85 | .066 |
| | **Task Presentation** | | | | | .092 | 1.87 | .062 |
| | **Class Management** | | | | | .263 | 3.86 | .000 |
| | **Instructional Guidance** | | | | | -.024 | 1.37 | .712 |
| **5th grade** | **Model** | .092 | 12.611 | 5, 621 | .000 | | | |
| | **Task Design** | | | | | -.079 | -1.23 | .221 |
| | **Task Presentation** | | | | | .170 | .33 | .740 |
| | **Class Management** | | | | | .389 | 6.53 | .000 |
| | **Instructional Guidance** | | | | | -.077 | 1.19 | .234 |

**Table 4.** A multiple regression model of four teaching dimensions predicting PUSH-UP test.

| | R | $R^2$ | F | df | Sig. | Beta | t | Sig |
|---|---|---|---|---|---|---|---|---|
| **Total** | **Model** | .031 | 9.54 | 4, 1196 | .000 | | | |
| | **Task Design** | | | | | .138 | 2.85 | .004 |
| | **Task Presentation** | | | | | .031 | .874 | .383 |
| | **Class Management** | | | | | .069 | 1.53 | .125 |
| | **Instructional Guidance** | | | | | -.279 | -5.04 | .000 |
| **Boys** | **Model** | .051 | 8.84 | 4, 655 | .000 | | | |
| | **Task Design** | | | | | .182 | 2.74 | .006 |
| | **Task Presentation** | | | | | -.029 | -.61 | .545 |
| | **Class Management** | | | | | .173 | 2.80 | .005 |
| | **Instructional Guidance** | | | | | -.336 | -5.20 | .000 |
| **Girls** | **Model** | .033 | 4.57 | 4, 535 | .001 | | | |
| | **Task Design** | | | | | .096 | 1.37 | .171 |
| | **Task Presentation** | | | | | .108 | 2.06 | .040 |
| | **Class Management** | | | | | -.056 | -.84 | .399 |
| | **Instructional Guidance** | | | | | -.232 | -3.54 | .000 |
| **4th grade** | **Model** | .022 | 3.197 | 4, 568 | .013 | | | |
| | **Task Design** | | | | | .177 | 2.48 | .013 |
| | **Task Presentation** | | | | | -.001 | -.019 | .985 |
| | **Class Management** | | | | | .021 | .306 | .760 |
| | **Instructional Guidance** | | | | | -.206 | -3.19 | .002 |
| **5th grade** | **Model** | .045 | 7.39 | 4, 623 | .000 | | | |
| | **Task Design** | | | | | .105 | 1.58 | .116 |
| | **Task Presentation** | | | | | .075 | 1.47 | .143 |
| | **Class Management** | | | | | .112 | 1.84 | .066 |
| | **Instructional Guidance** | | | | | -.206 | -3.19 | .002 |

**Table 5.** A multiple regression model of four teaching dimensions predicting TRUNK LIFT test.

| | R | $R^2$ | F | df | Sig. | Beta | t | Sig |
|---|---|---|---|---|---|---|---|---|
| **Total** | **Model** | .015 | 4.66 | 4, 1196 | .001 | | | |
| | **Task Design** | | | | | -.014 | -.30 | .768 |
| | **Task Presentation** | | | | | -.034 | -.96 | .383 |
| | **Class Management** | | | | | -.044 | -.97 | .331 |
| | **Instructional Guidance** | | | | | .174 | 3.74 | .000 |
| **Boys** | **Model** | .043 | 7,29 | 4, 655 | .000 | | | |
| | **Task Design** | | | | | -.116 | -1.73 | .084 |
| | **Task Presentation** | | | | | -.028 | -.59 | .558 |
| | **Class Management** | | | | | -.131 | -2.11 | .036 |
| | **Instructional Guidance** | | | | | .338 | 5.22 | .000 |
| **Girls** | **Model** | .025 | 3.48 | 4, 535 | .008 | | | |
| | **Task Design** | | | | | .133 | 1.89 | .059 |
| | **Task Presentation** | | | | | -.052 | -.98 | .327 |
| | **Class Management** | | | | | -.131 | -2.11 | .036 |
| | **Instructional Guidance** | | | | | .338 | 5.22 | .000 |
| **4th grade** | **Model** | .020 | 2.87 | 4, 568 | .023 | | | |
| | **Task Design** | | | | | -.012 | -.17 | .862 |
| | **Task Presentation** | | | | | -.001 | -.02 | .987 |
| | **Class Management** | | | | | -.145 | -2.10 | .036 |
| | **Instructional Guidance** | | | | | .195 | 3.03 | .003 |
| **5th grade** | **Model** | .028 | 4.55 | 4, 623 | .001 | | | |
| | **Task Design** | | | | | .000 | .001 | .999 |
| | **Task Presentation** | | | | | -.069 | -1.32 | .187 |
| | **Class Management** | | | | | .077 | 1.26 | .210 |
| | **Instructional Guidance** | | | | | .139 | 2.07 | .039 |

**Table 6.** Descriptive statistics of physical fitness tests between two levels of QPET

| | Above average Group | | | Below average Group | | |
|---|---|---|---|---|---|---|
| | n | M | SD | n | M | SD |
| **PACER** | 751 | 33.20 | 18.37 | 450 | 27.85 | 14.83 |
| **Curl-up** | 751 | 34.90 | 11.19 | 450 | 25.77 | 19.05 |
| **Push-up** | 751 | 12.77 | 9.49 | 450 | 13.52 | 10.98 |
| **Trunk lift** | 751 | 11.11 | 1.68 | 450 | 11.18 | 3.91 |

## Mean differences of physical fitness between two QPET groups

The mean score (2.67) of the overall QPET was computed to classify the 30 video-recorded lessons into two groups: above-average QPET group and below-average QPET group. Table 6 presents descriptive statistics of each physical fitness test between the two levels of QPET groups. To examine if there was a significant mean difference in each physical fitness test between the two levels of QPET groups, an independent sample t-test was conducted. The results of t-tests revealed that the students in the above-average QPET group significantly outperformed their

counterparts in the below-average QPET group in PACER test ($t_{1096.27} = 5.536$, $p < 0.01$) and Curl-up test ($t_{1057.45} = 7.553$, $p < 0.01$), but not in Push-up and Trunk Lift tests.

## Discussion

This study was central to examining the extent to which four essential teaching dimensions were associated with students' health-enhancing physical fitness. In line with the essence of QPE (NASPE, 2009), the results of the multiple regression models indicated that all four essential teaching dimensions were significantly associated with students' cardiovascular endurance, muscular strength and endurance, and flexibility for different samples including the total sample, the boys and the girls, and the fourth-grade and the fifth-grade students. Further, each teaching dimension showed different $\beta$ weights to prediction of each fitness test. In this study, Task Presentation made the most significant contribution to prediction of the students' cardiovascular endurance. Class Management contributed most to prediction of the students' curl-up test. Task Design and Instructional Guidance made significant contributions to prediction of the students' push-up test. Instructional Guidance contributed significantly to prediction of students' trunk lift test.

Consistent with the studies (Chen, 2005; 2012), this study empirically confirmed that to better enhance and maintain a healthy level of physical fitness among students, PE teachers need to provide students with sequentially progressive, developmentally appropriate, and maximally and actively engaging learning tasks; PE teachers also need to present learning tasks precisely along with demonstration of the task performance. To maximize students' learning time, PE teachers need to use effective class management strategies and routines to group students, to distribute learning materials/equipment, to arrange formations, and to locate students into working areas. During the students' practice, PE teachers need to provide tailored instructional guidance and flexibly adjusted learning tasks based on students' ongoing leaning responses. In short, PE teachers should enact high quality of all four essential dimensions in a PE lesson.

However, it is noted that the four teaching dimensions accounted for the total variance in PACER test (12.2%), followed by curl-up test (6.1%), and very low percentage of the variance in push-up test and Trunk Lift test for the total sample. Similarly, for the fourth-grade and the fifth-grade students, the highest percentage of the total variance explained by the four teaching dimensions was PACER test and followed by curl-up test, while the very small percentages of the total variance explained by the four teaching dimensions were push-up and trunk list tests. In addition, the students who had experienced a high level of QPET were more physically fit in PACER and curl-up tests, compared to their peers who did not have this experience. But, no significant mean differences in push-up and trunk lift tests between the two groups of QPET were found. The results showed that the four essential teaching dimensions played moderately significant roles in promoting students' cardiovascular endurance and abdominal muscular strength and endurance, while making relatively weak contribution to enhancing upper-body muscular strength/ endurance and back extensor muscular endurance and flexibility. Also, the results revealed that the four teaching dimensions were significantly linked to the students' physical fitness, but they were not the main factors that contributed to enhancing and maintaining a healthy level of physical fitness. Given the limitation of this study, that is, examining the extent to which the four essential teaching dimensions of QPET were associated with students' physical fitness, future studies may explore how other components of CSPAP (CDC, 2013), such as physically active classroom, physically active recess, and physical activity-based afterschool programs along with QPE, contributed to students' health-enhancing physical fitness.

One unique result of this study indicated that the significant contribution of the four essential teaching dimensions to students' physical fitness was gender-specific. The four essential teaching dimensions explained a higher percentage of total variance in PACER test for girls (17.3%) than for boys (9.7%). In other words, this study indicated that the QPET played a more significant role in enhancing girls' cardiovascular endurance, compared to boys'. Researchers found that children with healthy cardiovascular endurance were more likely to participate in both competitive and noncompetitive sports and physical activities (Castelli and Valley, 2007; Erwin and Castelli, 2008). In contrast, this study found that the four teaching dimensions accounted for a higher percentage of total variance in Curl-up, Push-up, and Trunk Lift tests for boys than for girls. This study indicated that the improvement of boys' abdominal, upper-body, and back extensor muscular strength and endurance as well as flexibility depended largely on the teachers' providing QPET in physical education lessons. It is well documented that an individual's having strong muscles in the core, shoulders, and back is critical to developing his/her good body postures and is beneficial to increasing metabolism, healthy lean body mass, and bone mass (CDC, 2011; Ortega et al., 2008).

Given the paramount role of each fitness component in health and the significant contributor of the QPET to girls' cardiovascular endurance and to boys' muscular strength and endurance as well as flexibility, this study suggests that PE teachers should use a balanced approach to teaching a variety of object control skills, small-sided and modified games, team building activities, sports-related physical activities, and fitness-enhancing games. PE teachers need to engage their students in maximum participation in skill practices, game play, and health-related physical activities through providing developmentally appropriate learning experiences. They also need to reduce class management time and to increase the time spent in MVPA engagement during a PE lesson.

Also, this study suggests that while incorporating health-related fitness activities into a regular PE lesson, PE teachers need to provide targeted activities and instructions for girls to improve their upper body muscular strength/endurance and for boys to improve their flexibility of specific major muscle groups. For example, during

a regular PE lesson, a PE teacher may allocate a small amount of lesson time to having students engage needs-based fitness activities. PE teachers may use age-appropriate upper body strength building activities as routinized warm-up to improve girls' muscular strength and endurance. PE teachers may use different types of stretching exercises and kids-yoga as routinized warm-up and cool-down activities to particularly help boys improve their flexibility.

## Conclusion

In conclusion, the four essential teaching dimensions were significantly associated with the students' health-enhancing physical fitness for total sample, the boys and the girls, and the fourth-grade and the fifth-grade students. However, the QPET accounted for a relatively low percentage of the total variance in PACER test, followed by Curl-up test, and explained a very low percentage of the total variance in Push-up and Trunk Lift tests. The students in the high level of QPET group significantly outperformed in PACER test and Curl-up test than their counterparts in the low level of QPET group. However, no significant mean differences in Push-up and Trunk Lift tests between the two groups were found. The significant contribution of the four essential teaching dimensions to physical fitness components was gender-specific. The four essential teaching dimensions played more significant roles in contributing to girls' cardiovascular endurance and to boys' muscular strength and endurance as well as flexibility. This study suggests PE teachers need to implement high quality features of all four essential teaching dimensions in a regular PE lesson.

## References


American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD). (2013) *Comprehensive school physical activity programs: Helping all students achieve 60 minutes of physical activity each day* (position statement). AAHPERD, Reston, VA. Available form URL: http://www.aahperd.org/naspe/standards/uploads/CSPAP-final-7-22-13-2.pdf.

Ball, D.L. and Forzani, F.M. (2009) The work of teaching and the challenge for teacher education. *Journal of Teacher Education* **60,** 497-511.

Castelli, D.M. and Valley, J.A. (2007) The relationship of physical fitness and motor competency to physical activity. *Journal of Teaching in Physical Education* **26,** 358-374.

Centers for Disease Control and Prevention (CDC). (2011) School health guidelines to promote healthy eating and physical activity. *Morbidity and Mortality Weekly Report* **2011,** 60(5).

Atlanta, GA: U.S. Department of Health and Human Services.Centers for Disease Control and Prevention (CDC). (2013) *Comprehensive school physical activity programs: A guide for schools*. U. S. Department of Health and Human Services: Atlanta, GA, 2013.

Chen, W. (2005) Examination of curricula, teaching practices and assessment through the national standards. *Physical Education and Sport Pedagogy* **2,** 159-180.

Chen, W., Hendricks, K. and Archibald, K. (2011) Assessing Preservice Teachers' Quality Teaching practices. *Educational Research and Evaluation* **17,** 13-32.

Chen, W., Hypnar, A., Mason, S., Hammond-Bennett, A. and Zalmout, S. (2014) Elementary School Students' Daily Physical Activity Behaviors: A Contributing Role of Quality Physical Education Teaching in Comprehensive School-Based Physical Activity Program. *Journal of Teaching in Physical Education* **33,** 592-610.

Chen, W., Mason, S., Staniszewski, C., Upton, A. and Valley, M. (2012) Assessing the Quality of Teachers' Teaching Practices. *Educational Assessment, Evaluation and Accountability* **24,** 25-41.

Chen, W., Rovegno, I., Cone, T.P. and Cone, S.L. (2012) An accomplished teacher's use of scaffolding during a second grade unit on designing games. *Research Quarterly for Exercise and Sport* **83,** 221-234.

Ewrin, H.E. and Castelli, D.M. (2008) National physical education standards: A summary of student performance and its correlates. *Research Quarterly for Exercise and Sport* **79,** 495-505.

Erwin, H., Beighle, A., Carson, R. and Castelli, D. M. (2013) Comprehensive school-based physical activity promotion: A review. *Quest* **65,** 412-428.

Lampert, M. and Graziani, F. (2009) Instructional activities as a tool for teachers and teacher educators' learning. *The Elementary School Journal* **109,** 492-509.

Meredith, M.D. and Welk, G.J. (2007) *FITNESSGRAM and activitygram test administration manual*. 4th edition. Champaign, IL: Human Kinetics.

National Association for Sport and Physical Education (NASPE). (2014) *Moving into the future: National standards for physical education.* 3nd edition. Reston, VA: Author.

National Association for Sport and Physical Education (NASPE). (2009) *Appropriate Instructional practice guidelines for elementary school physical education.* 2nd edition. Reston, VA: NASPE Publication.

Ortega, F.B., Ruiz, J.R., Castillo, M.J. and Sjöström, M. (2008) Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity* **31,** 1-11.

Rink, J.E. (2006) *Teaching physical education for learning.* 5th edition. Boston, MA: McGraw Hill.

Stodden, D.F., Goodway, J.D., Langendorfer, S.J., Roberton, M.A., Rudisill, M.E., Garcia, C. and Garcia, L.E. (2008) A developmental perspective on the role of motor skill competence in physical activity: an emergent relationship. *Quest* **60,** 290-306.


## Key points

- Although Task Design, Task Presentation, Class Management, and Instructional Guidance has its unique and critical teaching components, each essential teaching dimensions is intertwined and immersed in teaching practices.
- Four essential teaching dimensions all significantly contributed to students' health-enhancing physical fitness.
- Implementation of QPET in a lesson plays more significant role in contributing to improving girls' cardiovascular endurance.
- Implementation of QPET in a lesson contributed significantly to improving boy's abdominal, upper-body, and back extensor muscular strength and endurance as well as flexibility

## AUTHOR BIOGRAPHY


**Weiyun CHEN**

**Employment**

School of Kinesiology, University of Michigan

**Degree**

Ph.D.

**Research interests**

School-based physical activity promotion

**E-mail:** chenwy@umich.edu

**Steve MASON**
**Employment**
School of Kinesiology, University of Michigan
**Degree**
Bachelor degree
**Research interests**
Physical education teaching practices

**Andrew HYPNAR**
**Employment**
Fairfax County Public Schools, Fairfax County, VA
**Degree**
Bachelor degree
**Research interests**
Physical education teaching and physical activity promotion

**Austin HAMMOND-BENNETT**
**Employment**
Livonia Public Schools, Livonia, MI
**Degree**
Bachelor degree
**Research interests**
Physical education teaching and physical activity promotion

✉ **Weiyun Chen**
University of Michigan, School of Kinesiology, Ann Arbor, MI 48168-2013, USA