Large-Scale Youth Physical Fitness Testing in the United States: A 25-Year Retrospective Review

Russell R. Pate
University of South Carolina

Gregory J. Welk
Iowa State University

Kerry L. McIver
University of South Carolina

Introduction

Twenty-five years ago, one of the authors of this article (Pate) published in this journal a “point of view” that was aimed at making the case for large-scale physical fitness testing in American youth (17). While the author expressed caveats and conditions, the basic tenor of the article was positive. The article argued that carefully designed physical fitness testing protocols can have educational value for students, should be included in public health surveillance systems, and ought to be updated regularly in light of new knowledge. This article reflects on developments and advances that have played out over the intervening two and a half decades. Specifically, we will examine a) developments in the methodology for measuring physical fitness in youth, b) contemporary uses of fitness testing in educational settings, c) application of youth fitness testing in public health surveillance, and d) policy applications related to youth fitness. Fundamentally, we will address two questions: How has youth fitness testing evolved over the past 25 years? Is it playing a more productive role in education and public health today than it was in 1989? Because the original article limited its focus to fitness testing in the United States, this article will focus on changes that have occurred in the U.S. The authors recognize, however, that interesting developments have taken place in other countries and regions, including Canada (27), Australia (25), and Europe (26).
Where Were We 25 Years Ago?

When the article was published in 1989, the field of youth fitness testing was nearing the end of a period of tumultuous change. For the preceding three decades, the dominant approach to measuring youth fitness was embodied in the American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD) Youth Fitness Test, which served as the basis for the Presidential Fitness Award offered by the President’s Council on Physical Fitness and Sports. This test, often labeled a measure of motor fitness, emphasized measures of muscular strength, muscular endurance, and anaerobic power. Recipients of the Presidential Fitness Award were expected to score at or above the 85th percentile on all test items. Not surprisingly, most of the legitimate recipients of the award were athletically talented and motivated youth.

Beginning in the late 1970s and continuing through the 1980s, a number of professional and scholarly groups challenged the primacy of the motor fitness approach to evaluation of physical fitness in youth. An alternative to motor fitness, usually referred to as health-related physical fitness, gained traction, and by 1989, tests of health-related fitness were beginning to supplant motor fitness batteries. In 1980, AAHPERD released its health-related physical fitness test, but the organization also continued to sponsor its traditional youth fitness test. The latter remained the basis for the Presidential Fitness Award. This confusing state of affairs prompted the Cooper Institute to develop Fitnessgram, a youth fitness testing initiative that was fully invested in the health-related fitness philosophy, in 1982 (19). The Fitnessgram test items were linked to health-based outcomes, and the associated recognition system encouraged youth to attain a “healthy fitness zone.”

This schizophrenic condition, which promoted both motor fitness and health-related physical fitness testing methods, was reflected in both public health surveillance and school-based fitness testing procedures. From the 1950s through the 1970s, the President’s Council on Physical Fitness and Sports sponsored national surveys of youth fitness. The items in these surveys mimicked the composition of the AAHPERD Youth Fitness Test; hence, the products were large-scale surveys of motor fitness in American children and youth. In the 1980s, the President’s Council on Physical Fitness and Sports continued this tradition and again conducted a national survey that emphasized motor fitness. However, also in the 1980s, the Office of Disease Prevention and Health Promotion of the U.S. Department of Health and Human Services conducted a national survey, the National Children and Youth Fitness Studies (NCYFS; Study 1 on 10- to 18-year-olds; Study 2 on 6- to 9-year-olds). The protocols for the NCYFS surveys clearly were based on the health-related physical fitness philosophy. What we didn’t know in 1989 was that the practice of regularly surveying fitness in large national samples of U.S. children and youth had ended. No large-scale national surveys have been conducted since the 1980s.

How Far Have We Advanced?

Advances in Fitness Testing Practice

Over the past 25 years, fitness testing has become a standard, if not defining, component of most school-based physical education programs. For the most part, schools have had the option of choosing one of two nationally promoted fitness-testing
batteries: the President’s Challenge or Fitnessgram. The two programs followed their historical roots in motor fitness and health-related fitness, respectively, in the selection of items included in the test batteries. However, the programs also differed philosophically about how fitness should be promoted in schools. The President’s Challenge offered a highly visible award program that rewarded youth with a Presidential Youth Fitness Award if they achieved the 85th percentile on all five tests. Studies have determined that only a small percentage of youth achieved the award, and the system has been criticized to some extent on motivational grounds (5,6). The Fitnessgram program did not use an award system, but emphasized providing personalized reports to children and parents.

Both programs evolved over time to keep pace with changing technology and shifting priorities. However, in academic circles, considerable debate took place regarding the inherent value and utility of coordinated fitness testing in schools. The concerns focused primarily on pragmatic issues, such as whether fitness testing takes too much time away from curriculum, whether it focuses too much on outcomes, and whether it contributes to dislike of physical education in youth. A prominent editorial by Tom Rowland, former editor-in-chief of *Pediatric Exercise Science*, titled “The Horse is Dead; Let’s Dismount,” brought the issue to light (21). The editorial was followed by responses or new perspectives that either supported fitness testing (8,10) or added new critiques or approaches (1).

The debate, while interesting, did little to change or influence the popularity and use of fitness testing in schools. However, it may have led to an increased focus on the importance of providing schools and teachers with appropriate guidelines and strategies for using fitness testing effectively. The Fitnessgram Scientific Advisory Board provided one key set of guidelines, a position statement titled “Appropriate and Inappropriate Uses of Physical Activity and Physical Fitness Tests in Schools.” The board shared the guidelines with Fitnessgram users on the web and disseminated them in publication form and in the Fitnessgram Reference Guide (7). The appropriate uses included personal fitness self-testing, parental reporting, and personal tracking, while the inappropriate uses included using fitness scores for grading, evaluating teacher effectiveness, or determining exemptions (Table 1). The journal *Measurement in Physical Education and Exercise Science* also emphasized recommendations for effective use of fitness testing in a journal supplement devoted to the subject. The supplement included papers on guidelines for valid and reliable fitness testing (12), guidelines for incorporating testing into the curriculum (23), strategies for using fitness testing to enhance motivation (31), and the advantages of combined activity and fitness assessments (30).

**Surveillance**

As defined by the World Health Organization, public health surveillance is “the continuous, systematic collection, analysis, and interpretation of health-related data needed for planning, implementation, and evaluation of public health practice.” (33) By this definition, efforts to maintain a system for surveillance of physical fitness in children and youth have largely gone backward over the past 25 years. As noted above, no large-scale national surveys of youth fitness have been conducted in the U.S. since the 1980s. One possible explanation for this change is that public health agencies shifted to a focus on physical activity and away from physical
fitness. And it’s clear that measurement of physical activity has expanded in U.S. surveillance systems. For example, the Youth Risk Behavior Survey includes items to assess physical activity behavior, and this protocol is administered to state-based samples of high school students on alternate years (2). Also, over the past decade, the National Health and Nutrition Examination Survey (NHANES) has measured physical activity by both self-report and accelerometry in youth (28).

Despite the negative trend in large-scale surveys of fitness in American youth, some positive changes are taking place—most of them recently. At the national level, NHANES recently conducted the National Youth Fitness Survey (NYFS; 3). This survey assessed fitness in a small but nationally representative sample of children between the ages of 3 and 15 years; however, the results of the survey had not been released at the time of this publication. Although the National Center for Health Statistics (NCHS), which oversees the ongoing NHANES surveys, is managing the National Youth Fitness Survey, it is a stand-alone study that has not been incorporated into the overall NHANES protocol. Further, NCHS has not indicated that it will administer the NYFS at regular intervals. At this point, the current study is a singular effort.

Another positive development was the release in 2012 of an Institute of Medicine (IOM) report titled *Fitness Measures and Health Outcomes in Youth* (9). This report was produced by a panel of experts that assessed the relationship between youth fitness test items and health outcomes and recommended test items for use

---

Table 1  **Appropriate and Inappropriate Uses of Fitness Assessment as Presented by the Fitnessgram Scientific Advisory Board**

<table>
<thead>
<tr>
<th>Appropriate Uses of Fitness Assessments in Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal testing to help students evaluate their level of health-related fitness.</td>
</tr>
<tr>
<td>Institutional testing to allow teachers to view group data (for curriculum development).</td>
</tr>
<tr>
<td>Personal-best testing to allow individual students to privately determine performance levels.</td>
</tr>
<tr>
<td>Teaching students about criterion referenced health standards and what types of activity are needed to reach them.</td>
</tr>
<tr>
<td>Helping students track fitness results over time (in portfolios, for example).</td>
</tr>
<tr>
<td>Documenting that Fitnessgram is being administered in schools and that Fitnessgram student self-assessments are being tracked over time.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inappropriate Uses of Fitness Assessments in Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluating individual students in physical education (e.g., grading or state standards testing).</td>
</tr>
<tr>
<td>Grading students on their fitness performance or posting results for other students to see.</td>
</tr>
<tr>
<td>Evaluating teacher effectiveness.</td>
</tr>
<tr>
<td>Evaluating overall physical education quality.</td>
</tr>
</tbody>
</table>

Source: Fitnessgram/Activitygram Reference Guide (24); used with permission
in a national survey of health-related fitness in children and adolescents (9). The panel’s recommendations were conceptualized as applying to a large-scale, field-based survey similar to the NCYFS surveys of the 1980s. The panel was charged with selecting test items which had been shown, in previous research, to be associated with health outcomes such cardio-metabolic risk status in youth. The recommendations of the panel are presented in Table 2. Note that the IOM panel did not recommend specific protocols for the skinfold thickness and shuttle run items.

Although surveillance of youth fitness at the national level has taken as many steps back as forward over the past three decades, encouraging developments have occurred at the state level. Several states have mandated regular assessment of fitness in students (14), and some have aggregated data from school districts to create statewide estimates of fitness. California was the first to establish a state mandate for fitness testing in grades 5, 7, and 9 in 1995, and such testing now has been institutionalized within the state. Texas launched perhaps the largest state-based fitness testing initiative with the passage of Senate Bill 530, which called for all public schools to report fitness results from all students K-12 to the state (4). The state selected the Fitnessgram program as the test battery and created customized data aggregation utilities to facilitate submission of data to the state. Summary data were reported by the Texas Education Association. These initiatives in California and Texas may portend growing interest in state-level surveillance of youth fitness.

## Policy Applications

While all 50 states have standards for physical education and 26 require student assessment, only 14 states require fitness assessments. The initiatives in California and Texas, however, not only mandate assessment of fitness but also provide avenues for reporting and utilizing the data to encourage further development of policies related to physical activity and fitness in the schools.

In California, reporting of fitness survey results lead to an important policy change—funding to pay for licensed physical education teachers in California elementary schools. This was an important accomplishment because, previously,

### Table 2  Institute of Medicine Recommendations for Measures of Health-Related Fitness in National Surveys

<table>
<thead>
<tr>
<th>Measures of Body Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body mass index</td>
</tr>
<tr>
<td>Waist circumference</td>
</tr>
<tr>
<td>Skinfold thickness</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cardiorespiratory Endurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressive shuttle run</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Musculoskeletal Fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handgrip strength</td>
</tr>
<tr>
<td>Standing long jump</td>
</tr>
</tbody>
</table>

Source: Fitness Measures and Health Outcomes in Youth (9)
physical education had been taught by classroom teachers at the elementary level. In addition, the California Department of Education (CDE) provided data to examine the relationship between PE policy compliance and fitness and found that students in policy-compliant districts were more likely to be fit than students in noncompliant districts (22).

In Texas, the Physical Fitness Assessment Initiative (PFAI) allowed researchers to examine the relationships between fitness and school performance outcomes, including academic achievement and attendance (13). Currently, fitness test data can be searched, downloaded, and evaluated by district, school, and grade year for all schools in Texas (http://kinney.tea.state.tx.us/Pfi/ReportGenerator.aspx).

What Does the Future Hold?

Practice Developments

An important recent development is the establishment of the new Presidential Youth Fitness Program (PYFP). The PYFP is ushering in a new era of youth fitness programming for the United States (www.presidentialyouthfitnessprogram.org) by creating a unified national program. The new program is a collaborative effort designed to create a more unified approach to enhance assessment practices, professional development of teachers and recognition of students and schools. These three areas are viewed as pillars of the program (see Photo 1) and reflect the integrated contributions provided by three of the partner organizations: Cooper Institute and Fitnessgram (assessment), AAHPERD (professional development), and the President’s Council and AAU (recognition).

The three pillars will impact the way in which fitness assessment is used by teachers and schools. The involvement of the CDC in the PYFP collaboration also will help to facilitate appropriate evaluation practices and possible surveillance efforts through the program. An important characteristic of the PYFP is that basic support will be available for free, but schools can opt to use supplemental software,
training, and awards to enhance the effectiveness of the program. The unification of services and programming will help to facilitate more effective translation of guidelines into practice.

**Surveillance**

The mixed history of the last 25 years leaves us uncertain of the future for surveillance of physical fitness in the population of U.S. children and adolescents. Inclusion of young people down to age 12 in the comprehensive NHANES protocol in the 1999–2002 rounds of data collection is encouraging (18), and this may indicate some level of interest in monitoring fitness in American children and youth on an ongoing basis. Likewise, the recent National Youth Fitness Survey, also undertaken by the NHANES program, seems to indicate a growing interest in this area. However, at this point, there is no evidence that a large-scale, comprehensive, field-based survey of youth fitness, analogous to the National Children and Youth Fitness Studies of the 1980s, is likely to occur in the foreseeable future. However, if such a study is undertaken, the recent IOM report on Fitness Measures is available to inform the process (9).

**Policy Developments**

Two broad and important policy-related issues seem likely to impact youth fitness programming in the future. Driven by the dramatic increases in population levels of overweight and obesity in U.S. youth, those who are concerned about military preparedness have begun to focus on fitness in the adolescent population. Likewise, evidence that higher levels of physical activity and fitness may be associated with better student academic performance is prompting education leaders to consider greater investments in promotion of activity and fitness.

**Fitness and Military Preparedness.** Current military enlistment requirements do not include a fitness standard, but each division of the armed forces employs its own established fitness assessments during basic training. Weight-related limitations are among the highest reasons for recruits failing the fitness standards, and among current male recruits, 25% are disqualified for having a BMI above the threshold. Military experts suggest that, in addition to general improvements in military skills performance, basic training injury and attrition rates could be significantly reduced if recruits came into military service with higher baseline fitness and greater exposure to fitness interventions (15). To affect the fitness of military recruits, fitness interventions must begin in the school years.

**Fitness and Academic Achievement.** Multiple research studies have illustrated the positive relationship between cardiorespiratory fitness and academic achievement as assessed by both standardized tests and executive function and cognition assessments (11,20,29,32). Research also shows that the relationship between fitness and academic performance may already be established before the transition to middle school (16). In a longitudinal study, improvements in fitness by low-performers did not help them achieve the same academic level as consistently high fitness performers (32). While the relationship between fitness and academic achievement is strongest in middle and high school students (29), research suggests that assessment of fitness in younger age groups (young elementary and below) may be beneficial based on the potential need for intervention in younger ages (11).
Summary and Conclusions

The 25-year period since 1989 has been a fascinating one for large-scale fitness testing of youth. As a component of the school physical education curriculum, fitness testing seems to have become solidly established. The recent development of the Presidential Youth Fitness Program, based on a collaboration of public and private entities, seems likely to encourage even broader application of fitness testing in U.S. schools. However, monitoring of fitness in the population of U.S. children and youth lost momentum over the last 25 years, and surveillance of fitness has only recently received some significant attention by government officials. The policy-relevance of youth fitness has ebbed and flowed over the past decades. At present, the interest of policy makers in youth fitness seems to be rising, but only time will tell if that interest will result in greater investments in youth fitness programs.

Reference

9. IOM Committee on Fitness Measures and Health Outcomes in Youth. Fitness Measures and Health Outcomes in Youth. Washington, DC: Institute of Medicine, 2012.