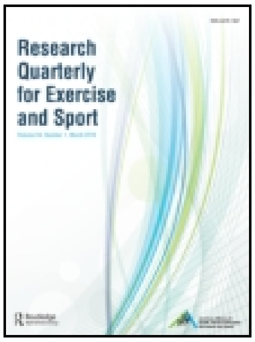

Routledge
Taylor & Francis Group

Research Quarterly for Exercise and Sport

ISSN: 0270-1367 (Print) 2168-3824 (Online) Journal homepage: http://www.tandfonline.com/loi/urqe20

Top 10 Research Questions Related to Musculoskeletal Physical Fitness Testing in Children and Adolescents

Sharon Ann Plowman

To cite this article: Sharon Ann Plowman (2014) Top 10 Research Questions Related to Musculoskeletal Physical Fitness Testing in Children and Adolescents, Research Quarterly for Exercise and Sport, 85:2, 174-187, DOI: 10.1080/02701367.2014.899857

To link to this article: http://dx.doi.org/10.1080/02701367.2014.899857

Research Quarterly for Exercise and Sport, *85*, 174–187, 2014
Copyright © SHAPE America
ISSN 0270-1367 print/ISSN 2168-3824 online
DOI: 10.1080/02701367.2014.899857

Top 10 Research Questions Related to Musculoskeletal Physical Fitness Testing in Children and Adolescents

Sharon Ann Plowman
Northern Illinois University

The purpose of this article is to bring attention to the 10 most pressing questions relevant to musculoskeletal physical fitness testing in children and adolescents. The goal is to stimulate research to answer these questions. The most pressing needs include establishing definitive links between valid, reliable, and feasible field test measures of muscular strength, endurance, power, and flexibility and health risk factors/markers in children and adolescents; determining the effects of exercise training on these relationships; and documenting the tracking of these relationships. The role of flexibility in health-related physical fitness (HRPF) needs to be carefully and specifically examined. Although body weight/composition is a separate component of health-related fitness, it is also a factor that can influence the performance of musculoskeletal test items. The role of body weight, body fat, and central adiposity and the possibility of adjustment of tests results are important research questions. Several questions relate to which field tests are best for use in schools. Finally, actual health-related criterion-referenced cutoff values need to be developed. In conclusion, more quality research is needed to firmly establish the musculoskeletal area for HRPF in youth.

Keywords: health-related physical fitness, neuromuscular tests, youth

When the philosophy behind physical fitness testing for children and adolescents shifted from performance-based norm-referenced evaluation to health-related criterion-referenced evaluation (~1975–1987), great debate occurred over what fitness components and which actual test items should be included (Plowman et al., 2006). Even more debate surrounded what criterion standards should be used for the selected test items for boys and girls at each age to represent adequate health-related physical fitness (HRPF). Ultimately, three fitness components were selected: cardiovascular endurance/aerobic capacity, body composition/body mass index (BMI), and muscular strength/endurance and flexibility, often labeled together as musculoskeletal fitness. Even then, data were available to solidly link the aerobic capacity and body composition components to health risk factors and/or mortality/morbidity in adults and, albeit less strongly, to children and adolescents (Cureton, Plowman, & Mahar, 2014; Going, Lohman, & Eisenmann, 2014). However, almost no scientific information linked musculoskeletal fitness with health in either adults or youngsters. As late as 2001, Warburton, Gledhill, and Quinney (2001b) stated: “For years, exercise scientists, fitness professionals, and physicians have intuitively expounded the health virtues of high levels of musculoskeletal fitness. However, little experimental data exist regarding the impact of high levels of musculoskeletal fitness on indicators of health status” (p. 218). Thus, intuitively, the HRPF tests (Table 1) kept musculoskeletal test items in these batteries. In the ensuing years, a great deal of research has reinforced the link between aerobic capacity and body composition and health, and standards for those test items have been directly linked to health risk (Morrow, Going, & Welk, 2011). Progress has also been made in linking musculoskeletal fitness with health, but weak links remain (Plowman, 2014). Now that the goal of one unified physical fitness test has been

---

Correspondence should be addressed to Sharon Ann Plowman, Emeritus, 2201 Lake Terrace Drive, Harrisonburg, VA 22802. E-mail: splowman@niu.edu

TABLE 1
Musculoskeletal Items in Selected Past and Present Nonmilitary Physical Fitness Batteries for Youth

Test Battery	*Kraus-Weber, 1954 (Kraus & Hirschland, 1954)*	*AAHPER(D) 1958, 1965, 1976*	*CAHPER, 1966*	*AAHPERD HRPFT, 1980*	*FITNESSGRAM® 1987, 1992 (Meredith & Welk, 2010)*
Upper body: Strength, Endurance, Power					
FAH		X	X		X
Pull-ups		X			X
Pushups					X
Dips					
Handgrip					
Throws		X			
Core: Strength, Endurance					
Sit-ups	X	X	X	X	
Curl-ups	X				X
Trunk extension					
Trunk lift	X				X
Leg lift	X (P&S)				
Lower body: Strength, Endurance, Power					
SL(broad)J		X	X		
Vertical jump					
Flexibility					
S&R	X			X	X
Shoulder					X

Test battery	*EUROFIT (Council of Europe, 1988)*	*PCPFS, 1987 PCFSN, 2010*	*YMCA, 1989 (Franks, 1989)*	*AFEA (Australian Council for Health, Physical Education and Recreation, 1996)*	*Japan (Shingo & Takeo, 2002)*
Upper body: Strength, Endurance, Power					
FAH	X	X			
Pull-ups		X	X		X
Pushups		X			
Handgrip	X				X
Throws				X	X
Core: Strength, Endurance					
Sit-ups	X				
Curl-ups	X	X	X	X	
Trunk lift					X
Leg lift					
Lower body: Strength, Endurance, Power					
SL(broad)J	X				X
Vertical jump					X
Flexibility					
S&R	X	X	X	X	X
Shoulder					X

(continued)

TABLE 1 – (*Continued*)

Test Battery	*CPAFLA (Canadian Society for Exercise Physiology, 2003)*	*ALPHA, 2011*	*IOM Recommendations, 2012*
Upper body: Strength, Endurance, Power			
FAH			
Pull-ups			
Pushups	X		
Dips			
Handgrip	X	X	X
Throws			
Core: Strength, Endurance			
Sit-ups			
Curl-ups	X		
Trunk extension	X		
Trunk lift			
Lower body: Strength, Endurance, Power			
SL(broad)J	X	X	X
Vertical jump			
Flexibility			
S&R	X		
Shoulder			

Note. FAH = flexed-arm hang; P&S = prone and supine; SLJ = standing long jump; S&R = sit-and-reach; AAHPER = American Alliance for Health, Physical Education and Recreation; CAHPER = Canadian Association for Health, Physical Education, and Recreation; AAHPERD = American Alliance for Health, Physical Education, Recreation, and Dance; HRPFT = Health-Related Physical Fitness Test; EUROFIT = European Test of Physical Fitness; PCPFS = President's Council on Physical Fitness and Sports; PCFSN = President's Council on Fitness, Sports and Nutrition; YMCA = Young Men's Christian Association; AFEA = Australian Fitness Education Award; CPAFLA = Canadian Physical Activity, Fitness, and Lifestyle Approach; ALPHA = Assessing Levels of Physical Activity and Fitness; IOM = Institute of Medicine. Execution of the test items varies widely among the test batteries and, in some cases, within the same test battery at different times. See references for specifics.

achieved with the adoption of the FITNESSGRAM® as the assessment tool in the Presidential Youth Fitness Program (http://www.presidentialyouthfitnessprogram.org) and with the publication of the Institute of Medicine's report *Fitness Measures and Health Outcomes in Youth* (2012), which recommends different musculoskeletal test items than currently used in the Fitnessgram, the musculoskeletal component is front and center. The purpose of this article is to bring attention to 10 of the most-pressing questions relevant to musculoskeletal fitness testing in children and adolescents. The goal is to stimulate research to answer these questions. The first and second questions are the basis of the entire inclusion of musculoskeletal items into any truly HRPF test battery; after the third question, the order is not particularly important. References within the context of each question are not intended to constitute a complete review of the literature in that area, but they are representative examples supporting the issue.

1. WHAT IS THE RELATIONSHIP BETWEEN HEALTH AND MUSCULOSKELETAL FITNESS (MUSCULAR STRENGTH, POWER, ENDURANCE, AND/OR FLEXIBILITY) IN CHILDREN AND ADOLESCENTS? WHAT MUSCULOSKELETAL FITNESS TEST ITEMS (THAT ARE VALID, RELIABLE, AND FEASIBLE MEASURES OF THESE CONSTRUCTS) CAN BE UNEQUIVOCALLY LINKED TO HEALTH RISK FACTORS/MARKERS IN CHILDREN AND ADOLESCENTS?

Balanced healthy integrative functioning of the musculoskeletal system requires that muscles be able to exert force or torque (measured as strength), exert force quickly (measured as power), resist fatigue (measured as muscular endurance), and move freely through a full range of motion (measured as flexibility). There is now increasing evidence for adults that these attributes of a healthy functioning musculoskeletal system are associated with specific and overall health status, independent living, and a reduction of risk for chronic disease, disability, and mortality (Plowman, 2014). Because most chronic diseases, disability, and mortality do not appear until well into adulthood, health risk factors or markers are commonly substituted as end points for children and adolescents. However, the literature linking specific musculoskeletal components and/or fitness test items with body composition/obesity, bone/joint health, cardiovascular disease risk factors, prediabetes/diabetes, low back pain (LBP), mental/cognitive function, and metabolic risk factors in children and adolescents is lagging behind. The Institute of Medicine report (2012) puts it this way:

"… There is an insufficient body of high-quality literature to support a strong link between performance on any specific musculoskeletal fitness test by youth … across all ages and stages of development and any health outcomes or markers" (p. 176).

There is a need for randomized, controlled trials, as well as longitudinal studies, to firmly establish the relationships between health and musculoskeletal fitness in children and adolescents and to determine the possible effects of modifying factors such as age, sex, and body composition on the relationships. In identifying specific fitness tests, these tests need to measure a range from limiting dysfunction to high capacity and need to accurately reflect the individual's fitness status (American Alliance for Health, Physical Education, Recreation, and Dance [AAHPERD, now SHAPE America – The Society of Health and Physical Educators], 1980). Furthermore, there is a need to utilize as many varied field tests in these trials as possible. Only those tests that are administered can be related, so the tendency to further investigate only the few items that have to date shown marginal evidence of a linkage to health must be avoided. Different health conditions may relate to different musculoskeletal fitness tests.

One of the difficulties in identifying valid specific fitness tests is the lack of agreement on criterion measures for muscular strength, endurance, power, and flexibility. If at all possible, however, field fitness test items should be linked directly to both the health risk/marker and to the laboratory tests of that construct.

If/when relationships between laboratory and field tests of musculoskeletal fitness are solidified, the question "What are the mechanisms responsible for the relationships?" must be answered.

2. WHAT ARE THE EFFECTS OF PHYSICAL ACTIVITY/EXERCISE TRAINING ON TESTS OF MUSCULOSKELETAL FITNESS AND THE RELATIONSHIP BETWEEN HEALTH AND FITNESS IN CHILDREN AND ADOLESCENTS?

Once specific musculoskeletal fitness test items have been identified that link performance with health markers, there is a need for systematic longitudinal and high-quality randomized controlled studies to determine the influence of appropriate physical activity/exercise training on these relationships. Do musculoskeletal fitness scores on the identified tests change positively with exercise training (and negatively with detraining) in children and adolescents independent of growth and maturation? To be an acceptable field test, both the health marker/risk factor and the musculoskeletal test items should respond positively to appropriate exercise. Do changes in musculoskeletal fitness scores accurately reflect changes in health markers/risk factors in children and adolescents?

There is a reasonable body of work that indicates that children and adolescents can significantly increase their muscular strength in response to resistance training and that resistance training is a safe activity although the long-term effects of structured resistance training on youth have not been documented (British Association of Exercise and Sport Sciences [BASES], 2004; Faigenbaum et al., 2009; Faigenbaum & Myer, 2010). There is a moderate body of work that indicates that health markers in youth respond favorably to resistance training (Benson, Torode, & Singh, 2008b; BASES, 2004; Faigenbaum et al., 2009). However, there is limited evidence from randomized controlled trials on which to base resistance-training prescription for obtaining specific health outcomes (Benson, Torode, & Singh, 2008a). Such evidence is needed.

3. HOW DOES THE RELATIONSHIP BETWEEN HEALTH AND MUSCULOSKELETAL FITNESS TRACK FROM CHILDHOOD TO ADOLESCENCE TO ADULTHOOD? ARE MEASURES OF MUSCULOSKELETAL FITNESS IN YOUTH USEFUL PREDICTORS OF ADULT HEALTH?

Once specific musculoskeletal fitness test items have been identified that link performance with health markers, there is a need for systematic longitudinal studies to determine whether the benefits of musculoskeletal fitness track from childhood to adolescence to the various stages of adulthood. Tracking refers to the stability of a characteristic or maintenance of relative position in a group over time. Do musculoskeletal fitness results and health risk factors/markers track and parallel each other as an individual ages? Do childhood and/or adolescent musculoskeletal fitness results predict adult health?

There is evidence that tracking of strength, power, and flexibility occurs from childhood to adolescence and adolescence to young adulthood in both boys and girls (Beunen et al., 1997; Maia et al., 2001; Malina, 1997; Matton et al., 2006), and adolescent physical fitness results have been shown to predict adult fitness (Mikkelsson et al., 2006; Taeymans, Clarys, Abidi, Hebbelinck, & Duquet, 2009). There is evidence of moderate tracking of health markers (IOM, 2012). However, it is important to track both the musculoskeletal fitness results and the health markers/risk factors in the same individuals more systematically over time.

Ruiz et al. (2009) recently reviewed the predictive validity of health-related fitness in youth by analyzing the results of longitudinal studies in children and adolescents on the relationship between physical fitness and the risk for developing an unhealthy cardiovascular or musculoskeletal profile later in life. Musculoskeletal fitness was assessed in only 8 of the 42 studies identified. The researchers concluded that there was strong evidence indicating that muscular strength changes from childhood to adolescence are negatively associated with changes in overall adiposity, but there was only moderate evidence for central adiposity. Evidence that muscular strength changes are associated with changes in other cardiovascular disease risk factors was inconclusive as was the evidence that flexibility or muscular strength in childhood and adolescence was predictive of LBP later in life. Few tracking studies have been extended into any of the stages of adulthood. Obviously a great deal of work remains to be done in this area. Longitudinal studies should include the impact of the time and tempo of the growth spurt and maturation (early, average, and late).

4. SHOULD FLEXIBILITY BE PART OF A HEALTH-RELATED FITNESS ASSESSMENT?

The stand-and-reach or sit-and-reach test has been around since at least the early 1950s as part of the Kraus-Weber test and all health-related test batteries from 1980 to 2003 (Table 1). However, neither the new ALPHA (**A**ssessing **L**evels of **PH**ysical **A**ctivity and fitness) test battery for children and adolescents (Ruiz et al., 2011) nor the IOM (2012) recommendations for a U.S. national fitness survey contain any flexibility item although the "committee suggests that in schools… flexibility test items may be included to educate youth and their parents about flexibility as a component of overall musculoskeletal fitness, function, and performance" (IOM, 2012, p. 201). The reasoning for the omission is that there is currently a dearth of information definitively linking flexibility to health when health is defined as cardiovascular disease risk factors (blood lipids, blood pressure), cancer, osteoporosis, obesity, diabetes, metabolic syndrome, or mental health conditions (anxiety, depression). Even for those areas where flexibility might logically and mechanistically be linked such as LBP, posture, and injury prevention, a relationship has not been consistently supported by research data, especially in children and adolescents (Plowman, 1992b, 2014; Ruiz et al., 2009). However, high levels of flexibility are associated with improved ability to complete activities of daily living, increased functional independence, unrestricted mobility, and a reduction in falls as individuals age (Kell, Bell, & Quinney, 2001; Warburton, Gledhill, & Quinney, 2001a, 2001b). Furthermore, it has been shown that along with the increase in obesity and decrease in aerobic capacity scores, significantly lower flexibility scores have been observed for both boys and girls during the last 25 to 30 years (Shingo & Takeo, 2002; Tomkinson, Léger, Olds, & Cazoria, 2003; Tremblay & Lloyd, 2010). Is it to be assumed that there is no reason to be concerned with these longitudinal flexibility changes? Or will this decrease in flexibility mean increased problems with age? Shingo and Takeo (2002) "… propose that the trend of… drastically reduced body flexibility (trunk extension and standing flexion) are issues confronting school physical education and deserve international focus" (p. 159). Mikkelsson and colleagues (2006) determined that for men, but not women, the adolescent physical fitness factors explaining the variance in an adult physical fitness index were distance running and flexibility. In addition, a link, albeit weak, has been found between flexibility and fundamental movement skills in children and adolescents. In turn, the link between fundamental movement skill competency and physical activity is strong (Lubans, Morgan, Cliff, Barnett, & Okely, 2010). Does flexibility influence activity level? Finally, flexibility appears to be a discrete construct of musculoskeletal fitness. A recent principal component analysis (Dumith, Van Dusen, & Kohl, 2012) of a battery of eight physical fitness abilities/skills (sit-and-reach, standing long jump, curl-up, pull-up, medicine-ball throw, 9-min run, 20-m run, and 4-m shuttle run) revealed that all of these items except flexibility were strongly associated with each other. What does this uniqueness say about the inclusion/exclusion of flexibility in an HRPF test battery?

The major question, of course, is: Is this lack of experimental connections between flexibility and health because flexibility is truly not related to health/health risks, or is it because flexibility has not been included in many quality research studies on a theoretical basis? Or was the IOM decision based, at least in part, on the way in which the literature review was conducted? As stated in the IOM (2012) report: “Although the CDC (Centers for Disease Control and Prevention) performed a systematic search for flexibility, the articles on this component were not abstracted because of limited time and resources. When studies addressing cardiorespiratory endurance or musculoskeletal strength or endurance also included flexibility as a fitness component, however, the CDC abstracted such information” (p. 52). Thus, from a pool of “6016 studies addressing flexibility” (p. 195), the decision to exclude was made on “7 experimental, 5 quasi-experimental, 9 experimental (no control) and 4 longitudinal studies” (p. 52) published between 2000 and 2010. In all fairness, no matter how the literature could have been analyzed, definitive studies with large participant pools of varying ages relating flexibility specifically to any health issue would not have been found. This is a problem that needs to be corrected. Indeed, the IOM report calls for studies to investigate all aspects of flexibility and health in children and adolescents.

Some information is available. One study (Miyachi et al., 2007) that examined the relationships between age, flexibility, and metabolic syndrome components concluded that poor flexibility was independently associated with high plasma glucose in middle-aged and older women. Might further research solidify a relationship? In addition, two studies that were not cited have reported a connection between flexibility and arterial stiffening. Arterial stiffness (loss of the ability of arteries to expand and recoil with cardiac pulsation and relaxation) is an independent risk factor for hypertension, a variety of cardiovascular/coronary heart disorders, stroke, and mortality (Fernhall & Agiovlasitis, 2008; Yamamoto et al., 2009). Both a stretching training program (Cortez-Cooper et al., 2008) and high sit-and-reach values (Yamamoto et al., 2009) have been linked with less arterial stiffening. A major question with this research is what mechanism of action could possibly be responsible for these findings? How can flexibility be reflecting arterial stiffness? Is this an area worth pursuing, and if so, will it provide that elusive link? Finally, Farinatti and colleagues (Farinatti, Brandão, Soares, & Duarte, 2011) investigated the acute effects of stretching on heart rate variability (HRV). HRV is a well-recognized way to assess the autonomic modulation of the heart during recovery from exercise—particularly parasympathetic reactivation. Faster reactivation is linked with a lower risk for cardiovascular disease. The stretching sessions were found to enhance postexercise vagal modulation in individuals with low flexibility levels. Do these results warrant further investigation into the long-term effects of flexibility and flexibility training on autonomic modulation?

If flexibility is to be included as a musculoskeletal test item, what is the best test? Given that flexibility is joint-specific, studies are needed that investigate flexibility field test items other than just the sit-and-reach either singly or as part of a combination of items.

5. DO MUSCULOSKELETAL TEST ITEMS NEED TO BE ADJUSTED FOR BODY WEIGHT AND/OR BODY COMPOSITION?

When the original American Association for Health, Physical Education, and Recreation (AAHPER) Youth Fitness Test was published in 1958, it included the five items indicated in Table 1 plus a 50-yard dash for speed and a shuttle run for agility. Test results could be reported as percentiles based on age and sex alone or, for junior and senior high school boys and junior high school girls, percentiles based on the Neilson-Cozens Classification Index that was a weighted composite of age, height, and body weight. The classification index continued to be used in the 1965 edition, but by 1976, it was no longer available. A paper by Gross and Casciani (1962) had concluded that “In general, age, height, and weight have negligible value for classification purposes . . . ” for the AAHPER Youth Fitness Tests. The 1976 manual (AAHPER, 1976, p. 10) simply states that “the Classification Index . . . is omitted from this manual as research indicates that age is a more valid criterion.” Teachers also did not like the extra time that using the index required. Is it true today that age and sex alone are the best criteria for standards?

Winter (1992), Jaric (2003; Jaric, Mirkov, & Markovic, 2005), and Markovic and Jaric (2004, 2005, 2007) in a series of publications make the case that physical performance tests (which include many commonly used musculoskeletal physical fitness test items) need to be normalized for body size. Lighter individuals have an advantage in tests based on supporting their own weight, whereas heavier individuals have an advantage in overcoming an external load. Scaling (the partitioning out of differences in size) can be done by several techniques (Winter, 1992), but the most common in exercise physiology are the use of ratio standards and allometric power functions. Ratio standards simply divide a physiological variable by an anthropometric attribute—typically body weight as in expressing power in Watts as $\text{Watts}\cdot\text{kg}^{-1}$. Allometric scaling uses equations that are of the general form: $y = a\cdot x^{b}$, where "y" is the normalized value of the test score "a" adjusted by the body size variable "x" raised to the power of "b." The most commonly used body size variables are height and weight/mass.

An example of the results obtained from these two scaling techniques is a study by Vanderburgh, Mahar, and Chou (1995). The tested variable was handgrip strength. Non-normalized strength scores (such as handgrip) penalize lighter individuals. In this study, ratio scaling (dividing by body mass) was shown to "... yield a scaled score no better that the raw score in partitioning out the effect of body mass" (p. 83). Conversely, allometric scaling "... allowed for the computation of useful [handgrip strength] norms free of the confounding effect of body mass" (p. 83). These results are fairly typical, but ratio scaling continues to be used. The use of allometric scaling is more complex than ratio scaling. To begin, power functions need to be either theoretically assumed or experimentally calculated.

Different classifications of musculoskeletal tests undoubtedly have different allometric power functions. According to one system of classification (Jaric et al., 2005; Markovic & Jaric, 2004), tests of exertional force (e.g., grip strength), tests of rapid movements (e.g., standing long jump, countermovement vertical jump), and tests of supporting body weight (sit-ups, pushups, pull-ups, static trunk extension endurance) would have very different allometric power values, although the power function for the tests of rapid movement may be zero (Markovic & Jaric, 2005, 2007). In addition, theoretical values often differ from experimental values, and experimentally determined values can differ between studies.

What impact would allometric scaling have on the relationship between laboratory tests of muscular strength, endurance, and power and field tests selected to represent them in children and adolescents? What impact would allometric scaling for body size have on the relationship between health risk factors/markers and the various tests of musculoskeletal fitness, and how would this vary between the sexes and ages?

There is no consistency in the literature. For example, Martinez-Gomez et al. (2012) used unadjusted values for handgrip strength when looking at relationships with health markers. Steene-Johannessen, Kolle, Andersen, and Anderssen (2013) adjusted handgrip strength for kilograms of body weight also when relating to health markers. Janz, Dawson, and Mahoney (2002) used allometric scaling for handgrip strength when investigating cardiovascular health and risk.

To make things even more difficult, allometrically adjusting performance values for body size does not adjust these values for factors such as differences in percent body fat or lean body mass, age, and maturation. Is percent body fat a more important confounder than body size?

Woods, Pate, and Burgess (1992) tested fourth- and fifth-grade students on five upper-body musculoskeletal endurance tests: pull-ups, flexed-arm hang, pushups, and two types of modified pull-ups. They provide evidence that percent body fat was a significant predictor of performance on all but one of the modified pull-up tests even after controlling for body weight. That is, after controlling for the effect of body weight, heavier participants were still at a disadvantage, but after controlling for the effect of body composition, heavier participants were not at a disadvantage.

Lloyd, Bishop, Walker, Sharp, and Richardson (2003) designed a study to determine the influence of body size and composition (measured as the sum of triceps and calf skinfold, body weight, and BMI) on the performance of the Fitnessgram test items. Of the neuromuscular items, trunk lift and back-saver sit-and-reach were not significantly correlated with any body size/composition measure, whereas both curl-ups and push-ups were. When adjusted scores were calculated by adding residual scores from a regression model to the mean of each performance variable, these adjusted scores were not correlated with any of the three measures of body size/composition. Adjusting the performance scores for the sum of skinfolds altered the classification (pass/fail according to the criterion referenced score) in both female participants (17.8% for the curl-up, 12.6% for the pushup) and male participants (17.5% for the curl-up, 20.5% for the pushup). In the adjusted scores, the advantage of leanness was removed in children with thinner skinfolds, whereas the disadvantage of thicker skinfolds was eliminated in the fatter children. This technique enabled each child to be evaluated relative to the expected performance of children with similar body composition or size.

Both Woods et al. (1992) and Lloyd et al. (2003) indicated that body composition should be considered when interpreting test performances. Given that there is typically a measure of body composition included in HRPF tests, should musculoskeletal test items all be adjusted for body composition so that the impact of body fatness is removed from the evaluation? Or should individuals of all levels of body composition be expected to perform the musculoskeletal tests as they live daily indicating functional fitness in handling their bodies? What impact would adjusting scores on the basis of percent body fat have on the relationship between laboratory tests of muscular strength, endurance, and power and field tests selected to represent them in children and adolescents? What impact would adjusting for percent body fat have on the relationship between health risk factors/markers and the various tests of musculoskeletal fitness? How would this vary between the sexes and ages and with and without prior allometric scaling for body size?

6. WHAT IS THE RELATIONSHIP BETWEEN HANDGRIP STRENGTH AND CENTRAL ADIPOSITY? WHAT, IF ANY, ARE THE IMPLICATIONS OF THIS RELATIONSHIP FOR HRPF TESTING?

The IOM report (2012) recommended handgrip strength for use in a national youth fitness survey and for consideration for inclusion in school physical fitness batteries. Handgrip strength has been used in Asian, Canadian, and European test batteries (see Table 1). Higher handgrip strength values are considered to be better than lower handgrip strength values.

Several studies have shown that handgrip strength values are significantly higher in overweight/obese male and female children and adolescents (Ara et al., 2010; Artero et al., 2010; Casajús, Leiva, Villarroya, Legaz, & Moreno, 2007; Deforche et al., 2003; Prista, Maia, Damasceno, & Beunen, 2003) than in normal-weight or underweight youth. In addition, studies have shown that handgrip strength values are significantly related to measures of body weight/body composition (Ara, Moreno, Leiva, Gutin, & Casajús, 2007; Milliken, Faigenbaum, Loud, & Westcott, 2008; Moliner-Urdiales et al., 2011; Vaara et al., 2012). In preadolescents and adolescents of both sexes, static strength and biological maturity are positively related (Beunen & Thomis, 2000; Taeymans et al., 2009). Overweight and/or obese youngsters may be early maturers. The most frequent explanation for the association between handgrip strength and overweight/obesity is that there is a greater amount of fat-free muscle mass in these individuals (Deforche et al., 2003; Moliner-Urdiales et al., 2011; Prista et al., 2003).

Additionally, and possibly more problematic, is the fact that several studies (Milliken et al., 2008; Moliner-Urdiales et al., 2011; Prista et al., 2003; Vaara et al., 2012) have shown positive associations between handgrip strength and central body fat either measured directly or indirectly by waist circumference. Abdominal adiposity (visceral fat) is an independent risk factor for diabetic/atherogenic abnormalities in youth (Kim & Lee, 2009). Although correlation certainly does not mean causation, what is the relationship between handgrip strength and abdominal adiposity, and if there is a consistent positive relationship, what is the physiological significance? What mechanisms might explain it? Does this mean that high handgrip strength values in overweight/obese individuals with large waist circumferences are not necessarily better than the lower values of normal-weight individuals in terms of HRPF? Should measures of waist circumference be included at least as an option in HRPF tests?

7. WHAT ARE THE BEST HEALTH-RELATED MUSCULOSKELETAL FIELD TEST ITEMS FOR USE IN THE SCHOOL SETTING?

As can be seen from Table 1, there has never been unanimity as to what musculoskeletal items should be included in physical fitness batteries in the school setting. Many of the test batteries have tried to achieve an anatomical distribution (upper arm and shoulder, girdle, core, and lower body) and construct distribution (strength, endurance, power, and flexibility). There was face validity in this approach for obtaining a comprehensive assessment of total-body musculoskeletal function, but as discussed previously, validity in relation to actual health risk factors/markers was, and is, minimal.

As mentioned previously, on the basis of the available health-related information, the recent IOM report (2012) recommended that schools should consider handgrip strength and standing long-jump tests, as well as alternative tests that have not yet been shown to be related to health but are valid, reliable, and feasible. These tests include the (Vermont) modified pull-up, pushups, and curl-ups. Curl-ups are suggested, not as an alternative to the standing long jump and handgrip, but as an addition as they measure a different construct—namely, core endurance. Reliability and validity information is available for most of these musculoskeletal test items (Artero et al., 2012; Castro-Piñero, Artero, et al., 2010; Plowman, 2014); however, there remains a need for greater understanding of the validity relationships between appropriate laboratory measures and field tests of muscular strength, endurance, and power (Milliken et al., 2008). In addition, should not tests other than the handgrip and standing long jump such as a vertical jump, plank exercises, and trunk extension endurance items be considered?

It is also important to answer the question: What is the relationship among the various musculoskeletal field test measures? For example, Castro-Piñero and colleagues recently published (Castro-Piñero, Ortega, et al., 2010) a study assessing the usefulness of the standing long jump as a general index of muscular fitness. They found a coefficient of determination of $R^2 = .864$ between the standing long jump and vertical jump in 90 boys and girls ages 6 to 17 years old. When compared against upper-body power and strength tests (basketball throw, 90° pushup, and static shoulder strength; note the handgrip strength test was not included in this battery of tests), values for the standing long jump were $R^2 = .851$, .542, and.694, and for the vertical jump, values were $R^2 = .843$, .555, and .608, respectively. Thus, although they recommend that the standing long jump be considered the general index of muscular fitness in youth, could not the same be said for the vertical jump on the basis of these statistics? What other field test items might correlate strongly with the standing long jump?

As to feasibility, there is debate (Milliken et al., 2008) about whether the vertical jump or the long jump is easier to administer. There is a long history of using the standing long jump in European testing (Table 1, EUROFIT and ALPHA), and it is generally favored there. However, the vertical jump is an important skill in several popular sports in the United States (e.g., volleyball and basketball), so motivation based on familiarity and willingness to practice could be a consideration if the vertical jump is seen as more relevant by American students. España-Romero et al. (2010) did find a low reliability for the standing long jump in the school setting for elementary-aged children but felt that feasibility and safety were acceptable. In addition, the handgrip strength test requires equipment (an expense and possible problem for maintaining calibration in the schools), one-on-one testing, and a meaningful amount of time. The dynamometer must be adjusted for each child/adolescent, and both hands should be tested more than once. Similarly, either of the jumps also requires testing one individual at a time and may involve equipment (mats or a jumping standard) but should go quickly. Thus, the relationships between musculoskeletal test items that can be administered to large numbers of students at the same time and handgrip strength and jumping need to be further investigated. Is it the test (e.g., the handgrip and jump) or the construct (static upper-body strength and lower-body power) that is most related to health? If it is the construct, are there other field tests that might be more easily administered? Of course, once identified, the musculoskeletal tests should be tested for feasibility in the schools before adoption.

8. CAN A VALID, RELIABLE, FEASIBLE BACK EXTENSION OR PLANK TEST BE DEVELOPED TO ASSESS HEALTHY BACK FUNCTION?

The original health condition with which flexibility and core strength and endurance were linked in HRPF tests was LBP. Theoretically, lumbar, hamstring, and hip flexibility enable proper pelvic rotation in posterior and anterior rotation. Strong, fatigue-resistant abdominal muscles maintain proper pelvic position and reinforce the back extensor fascia providing support during forward flexion. Similarly, strong, fatigue-resistant back extensor muscles provide stability for the spine, maintain erect posture, and control forward flexion. Although the anatomical logic is strong, the prospective experimental evidence for prediction of LBP in both adults and youth is not (Plowman, 1992b, 2014). The assumption seems to be that the weak link in the nonestablishment of a direct relationship between muscular strength, endurance, power, and/or flexibility and LBP is the result of nonsensitivity of the physical fitness measures (sit-and-reach, trunk extension, curl-ups, etc.). Yet at least part of the problem could lie in the identification and description of LBP.

Calvo-Muñoz, Gómez-Conesa, and Sánchez-Meca (2013) have recently completed a meta-analytical study on the prevalence of LBP in children and adolescents. Their search resulted in 59 studies that met their criteria. The analysis resulted in a point prevalence of 12%, a 1-week period prevalence of 17.7%, a 1-year period prevalence of 33.6%, and, a lifetime prevalence of 39.9% for LBP in boys and girls younger than 18 years old. The mean age of the participants showed a positive statistically significant relationship ($R^2 = .457$) with prevalence rates. The most recently published studies (2001–2011) reported higher prevalence rates than do older studies (1980–2000), and those with a better methodology exhibited higher lifetime prevalence rates than did studies with poor methodology. The most important methodological problems were a lack of a clear definition and delimitation of LBP and failing to include specifications of pain such as frequency, intensity, and duration of the episodes. There is, of course, no way of knowing from these results whether the higher recent prevalence values are the result of changes in youth or in the research techniques. However, the authors concluded "… more attention should be devoted to develop and apply prevention programs for young children … [and there is a] need for efforts towards an early detection in LBP in children and adolescents" (Calvo-Muñoz et al., 2013, p. 23). The fact that a definitive link has not been found does not mean that one will not be found, and longitudinal prospective studies continue to be needed with improved quality in both the identification of LBP and the musculoskeletal test items. At the very least, it is important to raise the awareness of what constitutes healthy back function in schools.

To this end, can better musculoskeletal fitness test items be developed? The only moderately consistent reported predictor for both first-time and recurrent LBP and thus possibly the most important link in healthy back function is trunk extension, especially the Biering-Sorenson static trunk extensor endurance test with a maximal hold time of 240 s (Plowman, 1992b). The current Fitnessgram trunk-lift test has not shown validity, but a 90° trunk extension test has some promise as an acceptable test item (Hannibal, Plowman, Looney, & Brandenburg, 2006). However, that still requires individual testing and modest equipment. What is needed in the schools is an item for testing several students at once. Possible items for investigation include, but are not limited to, versions of prone or lateral plank exercises, exercise balls, bleachers, and rolled mat variations. Six different types of static back extension endurance-testing methods were reviewed by Moreau and colleagues (Moreau, Green, Johnson, & Moreau, 2001), but more needs to be done.

Trunk extension, of course, represents only one side of the anatomical core. The IOM (2012) report suggests the continued use of the curl-up test as a core strength measure in the schools. It is, of course, more of a core dynamic endurance test. Torso muscle endurance reference values have recently been presented (Dejanovic, Harvey, & McGill, 2012). It would be interesting to see how a combination of dynamic trunk extension and dynamic curl-up and/or static trunk extension or plank activity and static curl-up relate to back function.

9. SHOULD MUSCULOSKELETAL FITNESS BE ASSESSED BY SOME TYPE OF INDEX INSTEAD OF A SERIES OF SEPARATE ITEMS?

Of the test batteries described in Table 1, only the Canadian Physical Activity, Fitness, & Lifestyle Approach (Canadian Society for Exercise Physiology, 2003) includes an index. For an evaluation of back health, the following five items are included: physical activity participation questionnaire, waist circumference, sit-and-reach, partial curl-ups, and back extension. In relation to this specific index, the question is: What is the validity? Generically, the question is: Is it possible that the lack of strong evidence between poor performance on muscular strength, endurance, and flexibility items and either first-time or recurring LBP is because test items have generally been looked at in isolation and not as a composite? Might this also be true for other health risks?

Recently, a great deal of interest has been expressed in a composite test called the Functional Movement Screen (FMS; Cook, Burton, & Hoogenboom, 2006a, 2006b). This series of seven tests assesses quality of fundamental movement patterns that require muscle strength, flexibility, endurance, range of motion, coordination, balance, and proprioception to theoretically identify an individual's limitation and/or asymmetries. The seven items are: deep squat, hurdle step, in-line lunge, shoulder mobility, active straight-leg raise, trunk-stability pushup, and rotary stability. The battery includes "clearing" tests for shoulder mobility, trunk-stability pushup, and rotary stability. Each test is graded 0 to 3 on the basis of the quality of the performance of the movement: 3 = performed without compensation; 2 = performed with compensation; 1 = could not perform; and 0 = pain. Interest is high in the FMS because some studies (Butler, Contreras, Butron, Plisky, & Kiesel, 2011; Chorba, Chorba, Bouillon, Overmyer, & Landis, 2010; Kiesel, Plisky, & Voight, 2007; Lisman, O'Connor, Deuster, & Knapik, 2013; O'Connor, Deuster, Davis, Pappas, & Knapik, 2011), but not all studies (Hoover et al., 2008; Schneiders, Davidsson, Hörman, & Sullivan, 2011; Sorenson, 2009), have indicated the ability of a cutoff score (generally taken as 14 out of the possible 21) to predict who is likely to get injured in subsequent physical occupations (athletics, firefighting, military) or to determine who had been injured within the past 6 months.

Only three studies have been found where children or adolescents have been tested with the FMS, and none of these have related to injuries or anything prospectively (Butler, Elkins, Kiesel, & Plisky, 2009; Duncan & Stanley, 2012; McFelea, Butler, Kiesel, Plisky, & Elkins, 2010). Does the FMS have potential for use in a physical fitness battery for children and adolescents? There are several practical problems with this battery as now configured: (a) Each individual needs to be tested separately on not only seven items, but five of these have left- and right-side evaluations—and generally three trials are allowed; (b) a minimum of 4 hr of training appears to be both necessary and sufficient for individuals doing the evaluation; (c) specific, although minimal, equipment is needed; and (d) criterion-referenced standards are not available for children and adolescents.

Interestingly, three of the seven items have counterparts that are being or have been used in physical fitness tests. The active straight-leg raise evaluates hamstring flexibility as does the sit-and-reach. The shoulder mobility test is basically the shoulder flexibility test used in the Fitnessgram. The "clearing" item for the trunk stability pushup is similar to the current Fitnessgram trunk lift, although it allows the arms to do the pushing up, and pushups, albeit differing in format, are part of both the FMS and many fitness tests. Given this background, might it be possible to develop a composite index from only a couple of these movements or even several more traditional musculoskeletal test items that are more feasible for the physical education setting that can be validated against health criteria?

10. WHAT CRITERION-REFERENCED CUTOFF VALUES ARE AVAILABLE FOR MUSCULOSKELETAL FITNESS TEST ITEMS FOR CHILDREN AND ADOLESCENTS? WHERE CRITERION-REFERENCED STANDARDS ARE NOT AVAILABLE, WHAT ARE THE OPTIONS?

Unfortunately, the answer to the first question is that there are currently no criterion-referenced cutoff values available for use as standards for any musculoskeletal fitness test item for American children and adolescents (IOM, 2012; Plowman, 1992a; Zhu, Mahar, Welk, Going, & Cureton, 2011). This is a definite need.

There are several techniques available for the establishment of criterion-referenced standards. These have been discussed at length in at least five references (IOM, 2012; Larson, Welk, & Eisenmann, 2014; Plowman, 1992a; Zhu, 2013; Zhu et al., 2011). Unfortunately, for the musculoskeletal portion of health-related fitness, it is not as easy as just following the steps in any given technique. To begin, it is best to have a "gold standard" criterion test for the construct that has been validated against a health outcome. Static, dynamic, and isokinetic maximal voluntary contraction using a variety of equipment, for example, have been used as criterion measures for strength. However, there is no agreement on which test is best and no absolute validation against any health outcome for any laboratory tests in this area even remotely comparable to maximal oxygen consumption for cardiovascular health. Without a criterion measure, the classic technique of validation of field tests against the criterion test cannot be done. Is there one best criterion measure each for muscular strength, muscular endurance, power, and flexibility, or is the complication of different body parts too much to ever come to an agreement on this?

Without one or even four criterion measures (one for each construct), can the contrasting group procedure (Berk, 1976) be used to determine cutoffs? What little evidence exists at the moment has yielded less-than-encouraging results (Looney & Plowman, 1990; Rutherford & Corbin, 1994).

If field tests such as the handgrip, standing long jump, curl-ups, pushups, or sit-and-reach can be validated against health risks/markers, can they be used directly for the establishment of percentile norms? The IOM report (2012) recommends that cut points for tests of the musculoskeletal fitness components be set temporarily by being derived from the "20th percentile" (p. 212). Which available normative percentile values are broad enough and current enough to use in this manner? There are some recent Australian and European norms available for a variety of musculoskeletal test items (Castro-Piñero et al., 2009, 2013; Catley & Tomkinson, 2013; Dejanovic et al., 2012; Ortega et al., 2011; Sauka et al., 2011), but can these simply be transferred for use with American youth?

If it is agreed that the handgrip and standing long-jump test items are the only musculoskeletal items that are going to be administered, then perhaps a new national survey of school children in the United States could provide these needed data for these items. However, although these tests have been part of EUROFIT since 1988 (Council of Europe, 1988), they have not been routinely administered in the United States since 1980 (the long jump was part of the AAHPER[D] Youth Fitness Test, but the handgrip has never been part of a U.S. test; see Table 1). Currently, at least in the Fitnessgram, individual schools/teachers have some degree of choice in which musculoskeletal tests are administered. Choice can be a good thing. However, if these options continue or are expanded, the variety of upper-body, core, and lower-body strength, endurance, and power test items available increases the probability of inconsistent results in pass/fail classifications even if cutoffs can be established for each separate test (Zhu et al., 2011). The primary test-centered equating method for setting cutoff scores (Zhu, Plowman, & Park, 2010) has been shown to be a way of equating test results from different items. However, it requires a criterion measure and an identified primary test. "For the construct whose criterion . . . measure has not been well identified and whose field tests have not been well validated (such as upper body muscular strengths tests . . .), researchers should be cautious before applying the new [test equating] methods" (Zhu et al., 2010, p. 408). What can be done to solve this problem?

CONCLUSION

The recent IOM report (2012) concluded that despite " . . . a lack of high-quality studies supporting a strong link between any specific musculoskeletal fitness test item and health outcomes in youth" (p. 153), musculoskeletal fitness should be tested in both a future national survey and the schools. The specific test items selected were based on the best available evidence. The challenge now is for researchers to perform high-quality studies using a variety of appropriate designs, statistics, and test items to determine whether musculoskeletal fitness is truly health-related in children and adolescents and/or predictive of adult health. The additional challenge is to find the best field test or tests to measure this construct and cutoff values that are accurate.

REFERENCES

American Alliance for Health, Physical Education, and Recreation. (1976). *AAHPER Youth Fitness Test manual* (rev. 1976 ed.). Reston, VA: Author.

American Alliance for Health, Physical Education, Recreation and Dance. (1980). *Health Related Physical Fitness Test manual*. Reston, VA: Author.

American Association for Health, Physical Education, and Recreation. (1958). *AAHPER Youth Fitness Test manual*. Washington, DC: Author.

American Association for Health, Physical Education, and Recreation. (1965). *AAHPER Youth Fitness Test manual* (rev. ed.). Washington, DC: Author.

Ara, I., Moreno, L. A., Leiva, M. T., Gutin, B., & Casajús, J. A. (2007). Adiposity, physical activity, and physical fitness among children from Aragón, Spain. *Obesity*, *15*, 1918–1924.

Ara, I., Sanchez-Villegas, A., Vicente-Rodriguez, G., Moreno, L. A., Leiva, M. T., Martinez-Gonzalez, M. A., & Casajús, J. A. (2010). Physical fitness and obesity are associated in a dose-dependent manner in children. *Annals of Nutrition & Metabolism*, *57*, 251–259.

Artero, V., España-Romero, V., Castro-Piñero, J., Ruiz, J. R., Jimenez-Pavón, D., Aparacio, V. A., . . . Ortega, F. B. (2012). Criterion-related validity of field-based muscular fitness tests in youth. *Journal of Sports Medicine and Physical Fitness*, *52*, 263–272.

Artero, E. G., España-Romero, V., Ortega, F. B., Jiménez-Pavón, D., Ruiz, J. R., Vicente-Rodríguez, G., . . . Castillo, M. J. (2010). Health-related fitness in adolescents: Underweight, and not only overweight, as an influencing factor. *The AVENA Study. Scandinavian Journal of Medicine & Science in Sports*, *20*, 418–427.

Australian Council for Health, Physical Education and Recreation. (1996). *Australian Fitness Education Award. User's manual and curriculum ideas*. Adelaide, Australia: Author.

Benson, A. C., Torode, M. E., & Singh, M. A. F. (2008a). The effect of high intensity progressive resistance training on adiposity in children: A randomized controlled trial. *International Journal of Obesity*, *32*, 1016–1027.

Benson, A. C., Torode, M. E., & Singh, M. A. F. (2008b). Effects of resistance training on metabolic fitness in children and adolescents: A systematic review. *Obesity Reviews*, *9*, 43–66.

Berk, R. A. (1976). Determination of optimal cutting scores in criterion-referenced measurement. *Journal of Experimental Education*, *45*, 4–9.

Beunen, G., Ostyn, M., Simon, J., Renson, R., Claessens, A. L., Eynde, B. V., … van't Hof, M. A. (1997). Development and tracking in fitness components: Leuven Longitudinal Study on Lifestyle, Fitness and Health. *International Journal of Sports Medicine*, *18*(Suppl. 3), S171–S178.

Beunen, G., & Thomis, M. (2000). Muscular strength development in children and adolescents. *Pediatric Exercise Science*, *12*, 174–197.

British Association of Exercise and Sport Sciences. (2004). BASES position statement on guidelines for resistance exercise in young people. *Journal of Sports Sciences*, *22*, 383–390.

Butler, R. J., Contreras, M., Butron, L., Plisky, P. J., & Kiesel, K. B. (2011). Functional Movement Screen and physical performance scores during a firefighter academy: Implications for injury screening [Abstract]. *Medicine & Science in Sports & Exercise*, *43*, 795.

Butler, R. J., Elkins, B., Kiesel, K. B., & Plisky, P. J. (2009). Gender differences in Functional Movement Screen and Y-balance test scores in middle school aged children [Abstract]. *Medicine & Science in Sports & Exercise*, *41*, 183.

Calvo-Muñoz, I., Gómez-Conesa, A., & Sánchez-Meca, J. (2013). Prevalence of low back pain in children and adolescents: A meta-analysis. *BMC Pediatrics*, *13*, 14–25.

Canadian Association for Health, Physical Education, and Recreation. (1966). *The CAHPER Fitness-Performance Test manual: For boys and girls 7 to 17 years of age*. Ottawa, ON, Canada: Author.

Canadian Society for Exercise Physiology. (2003). *Canadian Physical Activity, Fitness & Lifestyle Approach: CSEP-Health & Fitness Program's health-related appraisal and counselling strategy* (3rd ed.). Ottawa, ON, Canada: Author.

Casajús, J. A., Leiva, M. T., Villarroya, A., Legaz, A., & Moreno, L. A. (2007). Physical performance and school physical education in overweight Spanish children. *Annals of Nutrition & Metabolism*, *51*, 288–296.

Castro-Piñero, J., Artero, E. G., España-Romero, V., Ortega, F. G., Sjöström, M., Suni, J., & Ruiz, J. R. (2010). Criterion-related validity of field-based fitness tests in youth: A systematic review. *British Journal of Sports Medicine*, *44*, 934–943.

Castro-Piñero, J., Girela-Rejón, M. J., González-Montesinos, J. L., Mora, J., Conde-Caveda, J., Sjöström, M., & Ruiz, J. R. (2013). Percentile values for flexibility tests in youths aged 6 to 17 years: Influence of weight status. *European Journal of Sport Science*, *13*, 139–148.

Castro-Piñero, J., González-Montesinos, J. L., Mora, J., Keating, X. D., Girela-Rejón, M. J., Sjöström, M., & Ruiz, J. R. (2009). Percentile values for muscular strength field tests in children aged 6 to 17 years: Influence of weight status. *Journal of Strength and Conditioning Research*, *23*, 2295–2310.

Castro-Piñero, J., Ortega, F. B., Artero, E. G., Girela-Rejón, M. J., Mora, J., Sjöström, M., & Ruiz, J. R. (2010). Assessing muscular strength in youth: Usefulness of standing long jump as a general index of muscular fitness. *Journal of Strength and Conditioning Research*, *24*, 1810–1817.

Catley, M. J., & Tomkinson, G. R. (2013). Normative health-related fitness values for children: Analysis of 85347 test results on 9-17-year-old Australians since 1985. *British Journal of Sports Medicine*, *47*, 98–108.

Chorba, R. S., Chorba, D. J., Bouillon, L. E., Overmyer, C. A., & Landis, J. A. (2010). Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *North American Journal of Sports Physical Therapy*, *5*, 47–53.

Cook, G., Burton, L., & Hoogenboom, B. (2006a). Pre-participation screening: The use of fundamental movements as an assessment of function—Part 1. *North American Journal of Sports Physical Therapy*, *1*, 62–72.

Cook, G., Burton, L., & Hoogenboom, B. (2006b). Pre-participation screening: The use of fundamental movements as an assessment of function—Part 2. *North American Journal of Sports Physical Therapy*, *1*, 132–139.

Cortez-Cooper, M. Y., Anton, M. M., DeVan, A. E., Neidre, D. B., Cook, J. N., & Tanaka, H. (2008). The effects of strength training on central arterial compliance in middle-aged and older adults. *European Journal of Cardiovascular Prevention and Rehabilitation*, *15*, 149–155.

Council of Europe. (1988). *EUROFIT: Handbook for the EUROFIT tests of physical fitness*. Rome, Italy: Council of Europe, Committee for the Development of Sport.

Cureton, K. J., Plowman, S. A., & Mahar, M. T. (2014). Aerobic capacity assessments. In S. A. Plowman & M. D. Meredith (Eds.), *FITNESSGRAM®/ACTIVITYGRAM® reference guide* (4th ed., pp. 6-1–6-226). Dallas, TX: Cooper Institute.

Deforche, B., Lefevre, J., DeBourdeaudhuji, I., Hills, A. P., Duquet, W., & Bouckaert, J. (2003). Physical fitness and physical activity in obese and nonobese Flemish youth. *Obesity Research*, *11*, 434–441.

Dejanovic, A., Harvey, E. P., & McGill, S. M. (2012). Changes in torso muscle endurance profiles in children aged 7 to 14 years: Reference values. *Archives of Physical and Medical Rehabilitation*, *93*, 2295–2301.

Dumith, S. C., Van Dusen, D., & Kohl, H. W. (2012). Physical fitness measures among children and adolescents: Are they all necessary? *The Journal of Sports Medicine and Physical Fitness*, *52*, 181–189.

Duncan, M. J., & Stanley, M. (2012). Functional movement is negatively associated with weight status and positively associated with physical activity in British primary school children. *Journal of Obesity*, *2012*, 697563. doi:10.1155/2012/697563

España-Romero, V., Artero, E. G., Jimenez-Pavón, D., Cuenca-Garcia, M., Ortega, F. B., Castro-Piñero, J., … Ruiz, F. B. (2010). Assessing health-related fitness tests in the school setting: Reliability, feasibility and safety: The ALPHA study. *International Journal of Sports Medicine*, *31*, 490–497.

Faigenbaum, A. D., Kraemer, W. J., Blimkie, C. J. R., Jeffreys, I., Micheli, L. J., Nitka, M., & Rowland, T. W. (2009). Youth resistance training: Updated position statement paper from the National Strength and Conditioning Association. *Journal of Strength and Conditioning Research*, *23*(Suppl. 5), S60–S79.

Faigenbaum, A. D., & Myer, G. D. (2010). Resistance training among young athletes: Safety, efficacy and injury prevention effects. *British Journal of Sports Medicine*, *44*, 56–63.

Farinatti, P. T. V., Brandão, C., Soares, P. P. S., & Duarte, A. F. A. (2011). Acute effects of stretching exercise on the heart rate variability in subjects with low flexibility levels. *Journal of Strength and Conditioning Research*, *25*, 1579–1585.

Fernhall, B., & Agiovlasitis, S. (2008). Arterial function in youth: Window into cardiovascular risk. *Journal of Applied Physiology*, *105*, 325–333.

Franks, B. D. (1989). *YMCA Youth Fitness Test manual*. Champaign, IL: Human Kinetics.

Going, S. B., Lohman, T. G., & Eisenmann, J. C. (2014). FITNESSGRAM®/ACTIVITYGRAM® reference guide. In S. A. Plowman & M. D. Meredith (Eds.), *FITNESSGRAM®/ACTIVITYGRAM® reference guide* (updated 4th ed., pp. 7-1–7-12). Dallas, TX: Cooper Institute.

Gross, E. A., & Casciani, J. A. (1962). The value of age, height, and weight as a classification device for secondary school students in the seven AAHPER youth fitness tests. *Research Quarterly*, *33*, 51–58.

Hannibal, N. S. III, Plowman, S. A., Looney, M. A., & Brandenburg, J. (2006). Reliability and validity of low back strength/muscular endurance field testsin adolescents. *Journal of Physical Activity & Health*, *3*(Suppl. 2), S78–S89.

Hoover, D., Killian, C. B., Bourcier, B., Lewis, S., Thomas, J., & Willis, R. (2008). Predictive validity of the Functional Movement Screen™ in a population of recreational runners training for a half marathon [Abstract]. *Medicine & Science in Sports & Exercise*, *40*(Suppl. 5), S219.

Institute of Medicine. (2012). *Fitness measures and health outcomes in youth*. Washington, DC: National Academies Press.

Janz, K. F., Dawson, J. D., & Mahoney, L. T. (2002). Increases in physical fitness during childhood improve cardiovascular health during adolescence: The Muscatine Study. *International Journal of Sports Medicine*, *23*(Suppl. 1), S15–S21.

Jaric, S. (2003). Role of body size in the relation between muscle strength and movement performance. *Exercise and Sports Sciences Reviews*, *31*, 8–12.

Jaric, S., Mirkov, D., & Markovic, G. (2005). Normalizing physical performance tests for body size: A proposal for standardization. *Journal of Strength and Conditioning Research*, *19*, 467–474.

Kell, R. T., Bell, G., & Quinney, A. (2001). Musculoskeletal fitness, health outcomes and quality of life. *Sports Medicine*, *31*, 863–873.

Kiesel, K., Plisky, P. J., & Voight, M. L. (2007). Can serious injury in professional football be predicted by a preseason functional movement screen? *North American Journal of Sports Physical Therapy*, *2*, 147–152.

Kim, Y.-M., & Lee, S.-J. (2009). Physical activity and abdominal obesity in youth. *Applied Physiology, Nutrition, and Metabolism*, *34*, 571–581.

Kraus, H., & Hirschland, R. P. (1954). Minimum muscular fitness tests in school children. *Research Quarterly*, *25*, 178–188.

Larson, K. R., Welk, G. J., & Eisenmann, J. C. (2014). Diagnostic performance of BMI percentiles to identify adolescents with metabolic syndrome. *Pediatrics*, *133*, e330–e338. doi:10.1542/peds.2013-1308

Lisman, P., O'Connor, F. G., Deuster, P. A., & Knapik, J. J. (2013). Functional Movement Screen and aerobic fitness predict injuries in military training. *Medicine & Science in Sports & Exercise*, *45*, 636–643.

Lloyd, L. K., Bishop, P. A., Walker, J. L., Sharp, K. R., & Richardson, M. T. (2003). The influence of body size and composition on FITNESSGRAM® test performance and the adjustment of FITNESSGRAM® test scores for skinfold thickness in youth. *Measurement in Physical Education and Exercise Science*, *7*, 205–226.

Looney, M. A., & Plowman, S. A. (1990). Passing rates of American children and youth on the FITNESSGRAM® criterion-referenced physical fitness standards. *Research Quarterly for Exercise and Sport*, *61*, 215–223.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports Medicine*, *40*, 1019–1035.

Maia, J. A. R., Lefevre, J., Claessens, A., Renson, R., Vanreusel, B., & Beunen, G. (2001). Tracking of physical fitness during adolescence: A panel study in boys. *Medicine & Science in Sports & Exercise*, *33*, 765–771.

Malina, R. M. (1997). Tracking of physical activity and physical fitness across the lifespan. *Research Quarterly for Exercise and Sport*, *67*(Suppl. 3), 48–57.

Markovic, G., & Jaric, S. (2004). Movement performance and body size: The relationship for different groups of tests. *European Journal of Applied Physiology*, *92*, 139–149.

Markovic, G., & Jaric, S. (2005). Scaling of muscle power to body size: The effect of stretch-shortening cycle. *European Journal of Applied Physiology*, *95*, 11–19.

Markovic, G., & Jaric, S. (2007). Is vertical jump height a body size-independent measure of muscle power? *Journal of Sports Sciences*, *25*, 1355–1363.

Martinez-Gomez, D., Eisenmann, J. C., Gomez-Martinez, S., Veses, A., Romeo, J., Veiga, O. L., & Marcos, A. (2012). Associations of physical activity and fitness with adipocytokines in adolescents: The AFINOS Study. *Nutrition, Metabolism & Cardiovascular Diseases*, *22*, 252–259.

Matton, L., Thomis, M., Wijndaele, K., Duvigneaud, N., Beunen, G., Claessens, A. L., … Lefevre, J. (2006). Tracking of physical fitness and physical activity from youth to adulthood in females. *Medicine & Science in Sports & Exercise*, *38*, 1114–1120.

McFelea, J. T., Butler, R. J., Kiesel, K. B., Plisky, P. J., & Elkins, B. (2010). Can a six-week training program improve dynamic balance and functional movement in middle school-aged children? [Abstract]. *Medicine & Science in Sports & Exercise*, *42*, 715.

Meredith, M. D., & Welk, G. J. (2010). *FITNESSGRAM® & ACTIVITYGRAM® test administration manual* (updated 4th ed.). Champaign, IL: Human Kinetics.

Mikkelsson, L., Kaprio, J., Kautiainen, H., Kujala, U., Millelsson, M., & Nupponen, H. (2006). School fitness tests as predictors of adult health-related fitness. *American Journal of Human Biology*, *18*, 342–349.

Milliken, L. A., Faigenbaum, A. D., Loud, R. L., & Westcott, W. L. (2008). Correlates of upper and lower body muscular strength in children. *Journal of Strength and Conditioning Research*, *22*, 1339–1346.

Miyachi, M., Sanada, K., Yamamoto, K., Kawano, H., Gando, Y., Tanimoto, M., … Tabata, I. (2007). Age, flexibility, and metabolic syndrome [Abstract]. *Medicine & Science in Sports & Exercise*, *39*(Suppl. 5), S238.

Moliner-Urdiales, D., Ruiz, J. R., Vicente-Rodriguez, G., Ortega, F. B., Rey-Lopez, J. P., España-Romero, V., … Moreno, L. A. (2011). Associations of muscular and cardiorespiratory fitness with total and central body fat in adolescents: The HELENA Study. *British Journal of Sports Medicine*, *45*, 101–108.

Moreau, C. E., Green, B. N., Johnson, C. D., & Moreau, S. R. (2001). Isometric back extension endurance tests: A review of the literature. *Journal of Manipulative and Physiological Therapeutics*, *24*, 110–122.

Morrow, J. R. Jr., Going, S. B., & Welk, G. J. (Eds.) (2011). FITNESSGRAM®: Development of criterion-referenced standards for aerobic capacity and body composition. *American Journal of Preventive Medicine*, *41*(Suppl. 2), S63–S143.

O'Connor, F. G., Deuster, P. A., Davis, J., Pappas, C. G., & Knapik, J. J. (2011). Functional movement screening: Predicting injuries in officer candidates. *Medicine & Science in Sports & Exercise*, *43*, 2224–2230.

Ortega, F. B., Artero, E. G., Ruiz, J. R., España-Romero, V., Jiménez-Pavón, Vicente-Rodriguez, G., … Castillo, M. J. (2011). Physical fitness levels among European adolescents: The HELENA Study. *British Journal of Sports Medicine*, *45*, 20–29.

Plowman, S. A. (1992a). Criterion referenced standards for neuromuscular physical fitness tests: An analysis. *Pediatric Exercise Science*, *4*, 10–19.

Plowman, S. A. (1992b). Physical activity, physical fitness, and low back pain. *Exercise and Sport Science Reviews*, *20*, 221–242.

Plowman, S. A. (2014). Muscular strength, endurance, and flexibility assessments. In S. A. Plowman & M. D. Meredith (Eds.), *FITNESSGRAM®/ACTIVITYGRAM® reference guide* (updated 4th ed., pp. 8-1–8-55). Dallas, TX: Cooper Institute.

Plowman, S. A., Sterling, C. L., Corbin, C. B., Meredith, M. D., Welk, G. J., & Morrow, J. R. Jr. (2006). The history of FITNESSGRAM®. *Journal of Physical Activity & Health*, *3*(Suppl. 2), S5–S20.

President's Council on Fitness, Sports and Nutrition. (2010). *The President's Challenge: Choose a challenge: Physical fitness test*. Retrieved from http://www.presidentschallenge.org/challenge/physical

President's Council on Physical Fitness and Sports. (1987). *The Presidential Physical Fitness Award Program instructor's manual featuring the President's Challenge*. Washington, DC: Department of Health and Human Services.

Prista, A., Maia, J. A. R., Damasceno, A., & Beunen, G. (2003). Anthropometric indicators of nutritional status: Implications for fitness, activity, and health in school-age children and adolescents from Maputo, Mozambique. *American Journal of Clinical Nutrition*, *77*, 952–959.

Ruiz, J. R., Castro-Piñero, J., Artero, E. G., Ortega, F. B., Sjostrom, M., Suni, J., & Castillo, M. J. (2009). Predictive validity of health-related fitness in youth: A systematic review. *British Journal of Sports Medicine*, *43*, 909–923.

Ruiz, J. R., Castro-Piñero, J., España-Romero, V., Artero, E. G., Ortega, F. B., Cuenca, M. M., … Castillo, M. J. (2011). Field-based fitness assessment in young people: The ALPHA Health-Related Fitness Test Battery for children and adolescents. *British Journal of Sports Medicine*, *45*, 518–524.

Rutherford, W. J., & Corbin, C. B. (1994). Validation of criterion-referenced standards for test of arm and shoulder girdle strength and endurance. *Research Quarterly for Exercise and Sport*, *65*, 110–119.

Sauka, M., Priedite, I. S., Artjuhova, L., Larins, V., Selga, G., Dahlström, O., & Timpka, T. (2011). Physical fitness in Northern European youth: Reference values from the Latvian Physical Health in Youth Study. *Scandinavian Journal of Public Health*, *39*, 35–43.

Schneiders, A. G., Davidsson, A., Hörman, E., & Sullivan, S. J. (2011). Functional Movement Screen™ normative values in a young, active population. *The International Journal of Sports Physical Therapy*, *6*, 75–82.

Shingo, N., & Takeo, M. (2002). The educational experiments of school health promotion for the youth in Japan: Analysis of the "sport test" over the past 34 years. *Health Promotion International*, *17*, 147–160.

Sorenson, E. A. (2009). *Functional Movement Screen as a predictor of injury in high school basketball athletes*. (Unpublished doctoral dissertation). University of Oregon, Eugene.

Steene-Johannessen, J., Kolle, E., Andersen, L. B., & Anderssen, S. A. (2013). Adiposity, aerobic fitness, muscle fitness, and markers of inflammation in children. *Medicine & Science in Sports & Exercise*, *45*, 714–721.

Taeymans, J., Clarys, P., Abidi, H., Hebbelinck, M., & Duquet, W. (2009). Developmental changes and predictability of static strength in individuals of different maturity: A 30-year longitudinal study. *Journal of Sports Sciences*, *27*, 833–841.

Tomkinson, G. R., Léger, L. A., Olds, T. S., & Cazoria, G. (2003). Secular trends in the performance of children and adolescents (1980–2000): An analysis of 55 studies of the 20 m shuttle run in 11 countries. *Sports Medicine*, *33*, 285–300.

Tremblay, M., & Lloyd, M. (2010). Physical literacy measurement: The missing piece. *Physical and Health Education Journal*, *76*(1), 26–30.

Vaara, J. P., Kyröläinen, H., Niemi, J., Ohrankämmen, O., Häkkinen, A., Kocay, S., & Häkkinen, K. (2012). Association of maximal strength and muscular endurance test scores with cardiorespiratory fitness and body composition. *Journal of Strength and Conditioning Research*, *26*, 2078–2086.

Vanderburgh, P. M., Mahar, M. T., & Chou, C. H. (1995). Allometric scaling of grip strength by body mass in college-age men and women. *Research Quarterly for Exercise and Sport*, *66*, 80–84.

Warburton, E. R., Gledhill, N., & Quinney, A. (2001a). The effects of changes in musculoskeletal fitness on health. *Canadian Journal of Applied Physiology*, *26*, 161–216.

Warburton, E. R., Gledhill, N., & Quinney, A. (2001b). Musculoskeletal fitness and health. *Canadian Journal of Applied Physiology*, *26*, 217–237.

Winter, E. M. (1992). Scaling: Partitioning out differences in size. *Pediatric Exercise Science*, *4*, 296–301.

Woods, J. A., Pate, R. R., & Burgess, M. L. (1992). Correlates of performance on field tests of muscular strength. *Pediatric Exercise Science*, *4*, 302–311.

Yamamoto, K., Kawano, H., Gando, Y., Iemitsu, M., Murakami, H., Sanada, K., … Miyachi, M. (2009). Poor trunk flexibility is associated with arterial stiffening. *American Journal of Physiology-Heart and Circulatory Physiology*, *297*, H1314–H1318.

Zhu, W. (2013). Science and art of setting performance standards and cutoff scores in kinesiology. *Research Quarterly for Exercise and Sport*, *84*, 456–468.

Zhu, W., Mahar, M. T., Welk, G. J., Going, S. B., & Cureton, K. J. (2011). Approaches to development of criterion-referenced standards in health-related youth fitness tests. *American Journal of Preventive Medicine*, *42*(Suppl. 2), S68–S76.

Zhu, W., Plowman, S. A., & Park, Y. (2010). A primer-test centered equating method for setting cut-off scores. *Research Quarterly for Exercise and Sport*, *81*, 400–409.