Youth Fitness Testing: Validation, Planning, and Politics

B. Don Franks, James R. Morrow, Jr., and Sharon A. Plowman

Physical fitness testing has become a divisive issue among professional, governmental, and private agencies. The selection of test items has too often been the result of political compromise rather than measurement processes. This article outlines the physiological and measurement bases for the development of physical fitness tests. It raises questions that need to be addressed by physiological, psychological, and evaluation scholars. Roles for different individuals and groups are delineated. The process recommended would lead to the development of a better physical fitness test that could be supported and promoted by professional, governmental, and private organizations.

Testing the physical fitness of our youth has been a controversial topic during the past few years. It is almost beyond belief that this single issue has received so much attention. There have been numerous phone calls, meetings, memos, and letters about fitness testing among the top executives and officers of the American Alliance for Health, Physical Education, Recreation, and Dance (AAHPERD), the President’s Council on Physical Fitness and Sports (PCPFS), and the Institute for Aerobics Research (IAR) during the past couple of years. Along with many other professionals, we ourselves have spent hundreds of hours in various forms of communication on fitness testing as we tried to represent the AAHPERD Measurement and Evaluation Council, the AAHPERD Physical Fitness Council, and the AAHPERD Research Consortium in these discussions. We were also involved in the joint AAHPERD and PCPFS Task Force which cosponsored public hearings on the topic in Chicago in October 1986.

This paper will not attempt to rehash the events of the last few years on fitness testing except to reflect on those experiences to try to project a better means for developing, revising, enhancing, and implementing physical fitness development and testing. We want to share our views of the process that should be used.

About the Authors: B. Don Franks is with the School of HPERD, Long Field House, Louisiana State University, Baton Rouge, LA 70803. James R. Morrow, Jr., is with the Department of HPER, 123 Melcher Gym, University of Houston, Houston, TX 77004. Sharon A. Plowman is with the Department of PHED, Anderson Hall, Northern Illinois University, DeKalb, IL 60115.
for future test revision, development, and utilization. After a brief review of the recent history of fitness testing and some of the current options available, the paper will focus on what needs to be accomplished in the future. We believe that the substance of fitness testing must be moved from politics to planning. There must be a delineation of the roles of organizations whether they be governmental, private, or professional. Further, our best physiological, psychological, and measurement thinking is needed in order to develop the questions that need to be answered and to conduct relevant research so that future tests can be the best possible. Finally, the thinking of scientists and the resources of organizations must be used to develop long-range plans that will provide optimal content and promotion.

Today a fitness teacher/leader has several choices to make in terms of selecting a physical fitness test for children and youth. There are fitness tests from AAHPERD, IAR, and PCPFS as well as from individual states, nonprofit agencies, and private individuals/groups.

Table 1 includes the youth fitness tests developed by AAHPERD, IAR, and PCPFS. The original Youth Fitness Test (American Alliance, 1957) was developed by AAHPERD leaders and gained timely national visibility for fitness testing. This visibility was enhanced when, in the early 1960s, the PCPFS adopted the Youth Fitness Test as the basis for its Presidential Fitness Award. However, this test did not include a formal definition of physical fitness, had a limited physiological basis, and was not developed through sufficient measurement procedures. Based in part on criticism of the original Youth Fitness Test from measurement and exercise science specialists, AAHPERD started a process in 1975 to revise its physical fitness test. An appropriate measurement process was begun. This led to the 1980 Health Related Physical Fitness Test (American Alliance, 1980), which differentiated fitness related to functional health from performance on specific physical activities (e.g., athletics). For several years, both the Youth Fitness Test (slightly modified from the original version in 1976) and the Health Related Physical Fitness Test were promoted by AAHPERD. During that time the PCPFS recognized only the Youth Fitness Test and continued to recognize students for achieving the President’s Physical Fitness award on that test alone.

### Table 1

<table>
<thead>
<tr>
<th>Test</th>
<th>Organization</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youth Fitness</td>
<td>AAHPERD</td>
<td>1957, 1976</td>
</tr>
<tr>
<td>Health Related Physical Fitness</td>
<td>AAHPERD</td>
<td>1980</td>
</tr>
<tr>
<td>Fitnessgram</td>
<td>IAR</td>
<td>1987</td>
</tr>
<tr>
<td>President’s Challenge</td>
<td>PCPFS</td>
<td>1987</td>
</tr>
<tr>
<td>Physical Best</td>
<td>AAHPERD</td>
<td>1988</td>
</tr>
</tbody>
</table>

*AAHPERD = American Alliance for Health, Physical Education, Recreation, and Dance; IAR = Institute for Aerobics Research; PCPFS = President’s Council on Physical Fitness and Sports.
A computer feedback program, Fitnessgram, was developed in 1979 and came to be used with both the Youth Fitness and Health Related Test in a cooperative effort by AAHPERD, PCPFS, and IAR. Starting about 1985, AAHPERD considered publishing only one test booklet and the PCPFS eventually announced its own test items. Discussions on physical fitness definition, test items, evaluation, and awards proved divisive to continued cooperation. As of Fall 1988, AAHPERD, IAR, and the PCPFS are each sponsoring a different physical fitness test.

This brief historical perspective illustrates that professional, governmental, and industrial cooperative programs have not been greatly successful. We will attempt to outline the characteristics, circumstances, and plans that might result in more effective test development, validation, dissemination, and impact.

**Testing Standards**

Our primary fitness objective is to have individuals of all ages develop and maintain a healthy level of physical fitness through a healthy lifestyle (e.g., see Corbin, 1987; Osness, 1987). The health goals for the nation (U.S. Department of Health and Human Services, 1980) emphasize both fitness activities and fitness testing. Studies such as the National Children and Youth Fitness Study I (1985) and II (1987) provide helpful information on factors related to fitness (Pate & Ross, 1987). The focus of this article is on physical fitness testing which can enhance fitness development, but we must remember that fitness development is our ultimate goal.

An excellent means of evaluating any test is to consider the characteristics found in the *Standards for Educational and Psychological Testing* (1985). These standards and associated psychometric properties are used to evaluate the technical adequacy and overall quality of the test. The standards are identified as primary, secondary, and conditional. All of the primary standards should be met and the secondary and conditional standards should be considered in light of their appropriateness. Consideration should be given to each of the properties evaluated by Keyser and Sweetland (1985), Mitchell (1983), and Sweetland and Keyser (1983), with a conviction to having a newly developed test appear in these volumes. Test developers should keep these standards in mind throughout the developmental processes and make decisions that will ultimately lead to a positive test evaluation by external reviewers. Broadly speaking, the relevant technical standards include (a) validity, (b) reliability, (c) development and revision, (d) scaling, norming, score comparability and equating, and (e) technical manuals and user’s guides (*Standards for Educational and Psychological Testing*, 1985). Sufficient evidence must be supplied with the test and associated materials so that external reviewers can determine if the standards are met.

**Definition**

The initial concern with development of a youth fitness test must be definitional. The definition of physical fitness has evolved over the last 35 years but remains largely operationally defined by the tests used to measure it. The tests used to measure physical fitness have, in turn, frequently been selected by the criteria of convenience (e.g., no equipment, little time involvement, ease of scor-
ing) and tradition (e.g., familiar) rather than by physiological soundness represent-
ing specific components of fitness.

Since development of the AAHPER Youth Fitness Test in 1956, understand-
ing of the construct called "physical fitness" has undergone considerable change. Any youth fitness test should validly reflect the current definition. Most recent evidence has related physical fitness to functional health. Position papers pre-
pared by the American College of Sports Medicine (1988) and the American Heart
Association (1986) illustrate this fitness definition. It is appropriate that the cur-
rent definition be one of health related fitness because it is the health related con-
struct that best reflects total test validity.

Components

Given that a definition is developed, the focus then turns to inclusion of test components that relate to the construct (i.e., operationally define physical fitness). Although the concept of health related physical fitness is sound and defen-
sible as far as the existing research goes, the terminology is probably unfortunate, giving the impression that somehow health related fitness has absolutely no relation-
ship to sport performance. The reality of course is quite different in that cardio-
vascular endurance, muscular strength and endurance, flexibility, and body composition form a basis for risk reduction from hypokinetic diseases as well as for participation in a variety of physical activities. These four components should be recognized as physical fitness without the constraints and confusions imposed by any descriptor adjectives. These components meet the criteria established by
AAHPERD (1980) in that they (a) "measure a range which extends from severe-
ly limited dysfunction to high levels of functional capacity" and (b) "measure capacities that can be improved with appropriate physical activity" (American
Alliance, 1980, pp. 3-4). These criteria reflect physiological concerns and should serve as a starting point for further refinement in this area.

It is possible that these identified components are not the only ones that should be included. For example, bone density bears consideration. Given the problem of osteoporosis in the elderly (postmenopausal women in particular), and in highly trained amenorrheic young women, and the fact that mechanical stress caused by a muscle contraction influences bone mineral content (Drink-
water, 1986), a strong case could be made for its consideration as a fitness com-
ponent. Whether this factor will be sufficiently related to some field test (e.g., of muscular strength) awaits further research.

The discussion to this point has centered on the physiological nature of a physical fitness battery. Careful attention should be given to cognitive understand-
ings related to fitness testing for inclusion in test materials. In addition to the physical education that students receive in classes, cognitive information should be provided and tested. Teacher and student oriented materials should be pre-
pared. Following lectures, laboratory, and physical conditioning sessions, stu-
dents should be held accountable for classwork. Therefore consideration should be given to inclusion of numerous cognitive test items from which the teacher can choose (e.g., see Sander & Burton, 1987). This will require review of the item characteristics (e.g., content, difficulty, and discrimination) for various grade levels. A call should be made for items to be included and a selection committee should be appointed to evaluate, select, and organize the test for both hand-scored
and microcomputer versions. Validity and reliability estimates must be provided for these cognitive items.

**Microcomputers**

Nearly all fitness tests developed since 1980 have microcomputer software (e.g., AAHPERD's Physical Best, 1988; Fitnessgram, 1988; Fit Youth Today, 1986). Consideration should be given to enhancing such software so that it can be educational for the student as well as the parent. Computer aided instructional strategies should be considered for inclusion with fitness testing software. Most computerized reports are currently computer managed instruction oriented in that they are time saving for the teacher or administrator when preparing test results for distribution to students and/or parents. Care should be taken to provide necessary, adequate, sufficient, and valid information to parents and students. Pilot testing of reports should be conducted to determine readability and which information to present. Too much information is just as useless as too little. Parents review many reports from schools; experts should assist developers in determining which reports have the greatest impact and are most likely to be read and utilized.

Computer programmers should be employed for software development because of the rapid changes occurring in computer technology. Although physical educators should provide advice for software development, it is unlikely that the expertise required in software development, test development, measurement, and physiology will be found in a single individual. Resource people should be available to assist users (a toll-free 800 number would be most helpful).

**Grading**

Two areas often discussed in the development of tests relate to grading procedures and "teaching for the test." Most experts now agree that students should not be assigned grades based upon fitness test performance. However, teaching for the test is not only appropriate but it also should be done repeatedly. Students ought to be aware of their likelihood of meeting the physical standards prior to the test date because of the continued practice they had when taking the test during their class sessions. Teaching for the test is true not only for the physical battery but also for any cognitive information that is presented and tested within the classes. If the content of delivered material is important and students are encouraged and expected to learn the materials, they should fully expect to demonstrate higher levels of cognitive behaviors on written examinations.

The educational aspect of physical fitness testing must not be ignored. Current tests result in students completing the same test items throughout their school years. Physical education is the only academic discipline in which the basic test does not change from kindergarten through high school. If the test is only given once a year the student will nevertheless have taken it more often than any other test. This can and should illustrate the importance of the fitness components included in the battery. This is a key concept that must of necessity be carried with students throughout life. Thus they can have a greater understanding of the health related fitness construct and its associated cognitive, physical, and behavioral aspects.
Test Selection

Once the construct of physical fitness is adequately defined and the components isolated, it then becomes necessary to identify specific test items that measure the components. It is not enough for officers, staff, and members of different organizations to simply discuss test items. All decisions must be based upon current scientific evidence. Careful consideration must be given to the validity, reliability, objectivity, and feasibility of each item. This consideration will involve exercise science researchers, physicians, physical educators, health educators, epidemiologists, and public school teachers.

Validity

Validity is the sine qua non of the test battery as well as individual test items and must be considered above all other characteristics. Test manuals should include specific logical and statistical evidence of validation. It is imperative that each item adequately measure the component it is intended to measure. Each item must reflect the four basic types of validity. Not only should a test development team be concerned with the item’s concurrent validity (relationship to criterion test) but careful consideration must be given to predictive validity (can predict criterion test or behavior) with regard to future functional health problems. The relationship between functional health and various fitness parameters has been demonstrated (e.g., Paffenbarger, Hyde, Wing, & Hsieh, 1986; Paffenbarger, Hyde, Wing, & Steinmetz, 1984). However, most of the research conducted on the health related physical fitness construct has been accomplished with adults.

Though numerous fitness studies have been conducted with children, there is little research with school-age children that demonstrates the relationship between items of physical fitness and functional health that will have positive benefits throughout life. The matter of predictive validity is not well documented in children with future health characteristics as criteria. Test developers must consider this dilemma. This will have an impact on definition and fitness test development. Obviously the item should reflect construct validity in that it is related to the definition developed for the battery. Consideration of the item’s content validity (a logical choice for this characteristic), while not paramount, must be made.

Children and parents will be greatly influenced by the test battery simply based upon their perceptions of the test content. Time should be taken to educate teachers, students, and parents concerning physical fitness definition, components, and test items. Recall that distance runs were once thought dangerous for school-age children and that consideration was given to distances of less than ¼ mile as valid tests of aerobic performance. Distance runs of at least a mile are now commonly accepted by teachers, students, and parents alike. However, this has taken over 30 years of education.

Reliability

Reliability is demonstrated if the test evidences validity. However, separate consideration should be given to battery and individual item reliabilities (Safrit & Wood, 1987; Wood & Safrit, 1987). From a norm referenced perspective,
the items should reflect stability across time, assuming that no change in true score has occurred. Reliability estimates should be provided for subgroups based upon both age and gender. Additionally, standard errors of measurement should be included for various groups and subgroups of students. All reports of reliability should include thorough descriptions of the conditions under which the data were collected as well as reliability types (e.g., stability, internal consistency, alpha). Test users should be informed of the reliability problems associated with the use of change measures.

Objectivity

Objectivity, interrater reliability, must be considered for items not generally objective in nature. For example, skinfolds are now commonly used in assessing body composition in school-age children. Careful review of the research should be provided as evidence for adequate objectivity with these and similar measures. Once again, standard errors of measurement should be included. Generalizability theory (Brennan, 1983) studies might be most appropriate for evidence in these areas (i.e., intrarater and interrater reliability).

Feasibility

Representation from schoolteachers and administrators is essential in deciding which items to include in the battery. While an item might truly reflect validity, reliability, and objectivity, it may not work in the typical public school setting. This is demonstrated in early versions of fitness tests that included a swimming item. Perhaps no one would disagree that swimming might be an important skill for school-age children, but the item simply isn’t feasible in most school settings.

Population

Recent interest and discussion has centered on the appropriateness of fitness tests for special populations. Various groups have suggested fitness test batteries for subgroupings of special populations. It has even been suggested that “what is needed is a test battery for each possible combination of handicapping conditions” (Baumgartner & Horvat, 1988, p. 53). Although this suggestion is unrealistic, special populations must be considered in fitness test development. Programs such as the one being developed by the National Handicapped Sports and Recreation Association (1988) might provide appropriate alternatives in this area. However, the very nature of the processes involved in developing fitness tests and the numbers of subjects needed for determination of battery and item validity suggest that development of many such batteries is unlikely. Rather, teachers and researchers should be encouraged to recommend modifications in the test items so that special populations can be validly tested.

The health related concepts for nonspecial populations appear to be identical to those for special populations (i.e., heart disease, obesity, and musculoskeletal function). Suggestions for modification of typical items can be made for special populations. Classroom teachers should be provided guidelines for development of individualized standards for special students. Certainly, setting such standards will be difficult at best. Teachers should be encouraged to set standards based
upon improvement and behaviors that reflect enhancement in quality-of-life parameters. Comments regarding the reliability and validity research with such groups and test modifications should be included in testing materials. Test users should be cautioned in their use and interpretation until such modifications are more fully validated.

Measurement issues in nonschool settings should be considered. What is the purpose of the fitness test? Is it to be appropriate for use by untrained, non-school based personnel (e.g., scouting organizations)? If so, additional training and distribution strategies will be needed. Effects upon test validity and reliability should be reported when data are collected in these nontraditional educational settings and by those typically unfamiliar with testing procedures.

**Evaluation of Test Scores**

Nearly all of the recently developed national fitness test batteries have chosen to utilize criterion referenced standards for student achievement rather than norm referenced comparisons that were commonly used prior to 1986. Though this change is admirable, it is not without potential problems. Previously, students were simply compared to normative data based upon same-sex students of similar age. As the name implies, a student's performance is now referenced to a criterion for each test item. Thus, criterion referenced decisions serve as a type of formative evaluation information and diagnostic information for the teacher, student, and parent regarding the student's performance on specific health related test items. However, criterion referenced standards have been chosen by expert opinion based on minimal research results, clinical information, and established normative scores rather than absolute standards. This process is time consuming and inexact at best (Glass, 1978; King & Aufsesser, 1988; Safrit, Baumgartner, Jackson, & Stamm, 1980).

Fitness test performances based on normative standards have typically not increased throughout the nation, and in some instances have declined (e.g., Morrow, Crowhurst, Monaghen, Morava, & Pyfer, 1986; National Children and Youth Fitness Study, 1985, 1987; President's Council, 1986). However, there are little data to substantiate that the performance demonstrated by these youths is not sufficient for healthy living (at their current age as well as once they have reached adulthood). Simons-Morton, O'Hara, Simons-Morton, and Parcel (1987) suggest that children are the most habitually active members of society. It must be decided whether the current level of physical activity is directly related to future behaviors and degenerative diseases. Only then can truly valid criteria be determined for test items. Until such data become available, careful criteria review should be completed periodically to determine if chosen standards are valid.

Criteria should not be changed simply because students do or do not meet the criteria. Groups responsible for criteria should extensively describe and explain the processes used in setting standards so that future researchers will have full understanding of the methods. This process has already begun with criteria being set by FYT (Fit Youth Today, 1986), AAHPERD (1988), and IAR (1988). Dotson's (1988) recent article on criterion standards for aerobic endurance is a good example of the type of rationale that will lead to discussion, further research, and revised future standards.
Inclusion of criterion referenced standards results in additional issues of validity and reliability related to the determination of whether or not the student has met the standard. With criterion referenced standards, the issue of evaluation is comparison to what should be rather than to what is (as is done with norm referenced evaluation). The issue of validity is one of accuracy of classification (i.e., fit or unfit), and that of reliability is consistency of classification with criterion referenced standards (Safrit, 1986). Considerably more evidence is needed before confident decisions can be made in this area, particularly concerning young school-age children. If these “cutoff” scores (i.e., fit or unfit) are to be used, specific reference should be made to reliability and standard errors of measurement estimates near these points, as they might differ somewhat from the reliability estimates otherwise provided.

Research Plan

There is a basic need in physical fitness testing for a preplanned logical approach to research in the area, perhaps coordinated by some body such as AAHPERD’s Research Consortium and/or a committee of the American College of Sports Medicine. Funding is needed not just to survey the status quo of fitness levels, as has been done in the past, but to systemically attempt to answer some of the following questions: What is the consensus of fitness experts on the definition of physical fitness? What criteria should be used for “expert”? How can this consensus be determined? What is the consensus of fitness experts on the components of fitness based on the definition? How can this consensus be determined?

Based on the current understanding of the definition of physical fitness and the fitness components, the following research questions need to be addressed: (a) Does each fitness component relate to tests of positive health? (b) Does it provide the basis for lower risks of major health problems? (c) Does it provide the basis for participation in a variety of activities? (d) Does it improve with regular physical activity? Are there other potential fitness components that meet these criteria (a–d)? What are the best field tests for each component? What additional validity/reliability/objectivity studies are needed? What are the criterion standards for each test item for females and males of different ages? What research is needed to determine more valid standards? What are the effects of growth and maturation on each fitness test item? How can these effects be considered in evaluating test scores? If criterion referenced evaluation is to be utilized, what cutoff scores are best? What recommendations for physical activity can be made to improve each fitness component for females and males of different ages and different levels of fitness?

Little evidence exists for a relationship between functional health throughout life and the fitness levels of school-age children, particularly at elementary school levels. However, this should not keep one from developing field batteries that can be utilized while such relationships are further investigated. Obtaining adequate evidence for such a relationship will literally take a lifetime of tracking students. Most of the relevant research with children has been cross-sectional in nature rather than longitudinal. This evidence has taken over 30 years to accumulate for adults and will necessarily take nearly twice as long for young school-age children.
**Development Strategy and Public Relations**

Implementation strategies must be developed and public relations activities must be extensively included if the new battery is to receive widespread attention and acceptance. This will require work by experts in these areas. After nearly 6 years of use, AAHPERD’s Health Related Physical Fitness Test still had not been widely adopted (Safrit & Wood, 1986). State of the art materials will require large capital expenditures to increase adoption. Steps should be taken to include the test in test reviews and critiques (e.g., Keyser & Sweetland, 1985; Mitchell, 1983; Sweetland & Keyser, 1983). Of interest is that the AAHPERD Health Related Fitness Test (1980) does not appear in such texts while the AAHPERD Youth Fitness Test (1976) continues to appear.

**Differentiated Roles**

Professional organizations such as AAHPERD or the American College of Sports Medicine must provide the substance for physical fitness testing, as well as for what should and should not be included in fitness programs. These organizations have the professionals who have developed and are continuing to study the types of physiological and measurement issues raised earlier in this paper. The exercise science members of these groups are best suited to answer the question, “what does the current evidence say about . . . ?” In addition, members of these groups are most likely to do the careful scientific studies needed for advances in fitness testing. A good example concerns the evaluation of body composition. When the health related test was being developed, it was determined that according to the definition of physical fitness the amount of fat a person had was one component of fitness. It was further decided that based on current evidence, the skinfold measurement of fat was the best field test. Further, it was concluded that there was not enough information to get accurate estimates of percent fat from skinfold measurements in children.

Several studies were undertaken (with Lohman and Wilmore the major investigators) that have provided the methods to estimate percent fat from skinfolds in children. A practical problem was raised with the use of the subscapular skinfold (especially with male teachers testing female students). The scientists (e.g., Lohman and others) found that the use of the triceps and calf skinfolds were as accurate as the triceps and subscapular skinfolds, thus solving another problem. This example would have been a better one if the professional organizations had funded the research. But in this case the individuals were responsible for their own funding, some from grants and some from their host universities. Another role of professional organizations is to have the practitioners assist with questions of feasibility and implementation (e.g., see Kopperud, 1986).

Agencies of the government have two major roles to play (Executive Order 12345, President’s Council, 1987). The first one, which has been done so well by the President’s Council on Physical Fitness and Sport, is to raise the public’s awareness through its promotion of fitness. The second role is to work with professional organizations in determining priorities, as was done in the 1990 health objectives for the nation, and to fund needed research in fitness.

Nonprofit agencies such as Young Men’s/Women’s Christian/Hebrew Associations, Boys/Girls Clubs, American Amateur Union, youth sports, and other
volunteer agencies have enormous influence on the children and youth of our country. They have a responsibility to work with professional organizations to select the best fitness tests and activities to be used in their programs, and to give feedback concerning problems associated with the recommended tests and activities.

Business and industry have cooperated with many programs in the past, providing the resources for research and promotional activities for fitness. They should work closely with the professional organizations to ensure that they are promoting the best tests/activities available.

The major point is that it is up to the members of professional organizations who have the academic background and research skills to provide the substance for fitness development and testing. The role of the practitioner members of professional organizations is to enhance the practical implementation of fitness development and testing. The role of government and private industry is to help coordinate promotional efforts and to provide support for the research that will be the basis for future revisions.

Implementation

The people and processes needed to successfully address the issues related to youth fitness test development are many and varied. These issues will require a number of experts to assist in developing the most valid fitness battery. A steering committee of 8 to 10 experts from the disciplines of medicine, exercise physiology, measurement, psychology, research, epidemiology, and public school personnel should be developed. This committee should consist of widely respected experts who are known for their interest and work in youth fitness. It is not enough to simply choose those who are willing to work with the committee.

Each expert must be committed to participation throughout a 3- to 4-year developmental period. Each steering committee member should work with a small group of 5 or 6 experts from their respective disciplines who develop strategies, positions, and recommendations for the steering committee. Each issue reviewed by the expert teams should be collectively presented before the steering committee, who can then query individuals and teams about questions that cross the various disciplines. The steering committee should then make final decisions regarding the format, inclusion, and nature of the test. Once these decisions are made, the proposed battery should be reviewed by the expert groups and then presented before national bodies as the proposed plan. After input from these meetings, the steering committee should make final decisions regarding the battery.

We believe the implementation of these recommendations would result in a cooperative plan that would improve the validity of fitness testing. More important, better testing would lead to an enhanced functional health for our society.

References


