DOCUMENT RESUME

ED 433 627 EC 307 372

AUTHOR Winnick, Joseph P.; Short, Francis X.
TITLE Project Target: Criterion-Referenced Physical Fitness Standards for Adolescents with Disabilities. Final Report.
INSTITUTION State Univ. of New York, Brockport. Coll. at Brockport.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE 1998-05-00
NOTE 226p.; For related documents, see EC 307 373-375.
CONTRACT HO23C30091-95
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC10 Plus Postage.
DESCRIPTORS Adolescents; Child Health; Children; *Criterion Referenced Tests; *Disabilities; Elementary Secondary Education; Performance Tests; Physical Fitness; *Physical Fitness Tests; Standards; *Test Construction; Test Content; Test Items
IDENTIFIERS *Testing Accommodations (Disabilities)

ABSTRACT

This final report discusses the outcomes of a project designed to extend the nation's current emphasis on health-related, criterion-referenced fitness testing and programming to children and adolescents with disabilities. It summarizes project activities leading up to the Brockport Physical Fitness Test and related measures. Activities included: (1) identifying and selecting populations requiring adjustments to physical fitness tests and standards; (2) identifying tentative test items and standards for physical fitness development; (3) validating test items and standards; (4) developing a criterion-referenced test of physical fitness and modifying items and standards associated with selected national physical fitness tests designed for typical individuals; (5) developing an educational component relevant to adolescents with disabilities; and (6) preparing materials for project dissemination. Appendices include 16 research reports that evaluated specific test items for the criterion-referenced health-related test for individuals with disabilities. Appendices also include the notes and agenda of the Advisory Committee meetings. (Contains 19 references.) (CR)

PROJECT TARGET:

CRITERION-REFERENCED PHYSICAL FITNESS STANDARDS FOR ADOLESCENTS WITH DISABILITIES

Final Report

- - -

Submitted to the Office of Special Education and Rehabilitative Services, U.S. Department of Education

by
Joseph P. Winnick
and
Francis X. Short

Department of Physical Education and Sport
State University of New York
College at Brockport

May 1998

This project was funded by the Office of Special Education and Rehabilitative Services, U.S. Department of Education, project number H023C30091-95. The contents in this final report are those of the authors and do not necessarily reflect the position or policy of the U.S. Department of Education and no official endorsement should be inferred.

PROJECT TARGET:

CRITERION-REFERENCED PHYSICAL FITNESS STANDARDS FOR ADOLESCENTS WITH DISABILITIES

Final Report

Submitted to the Office of Special Education and Rehabilitative Services, U.S. Department of Education

by
Joseph P. Winnick
and
Francis X. Short

Department of Physical Education and Sport
State University of New York
College at Brockport

May 1998

Project Target

Central Staff at the State University of New York
College at Brockport

Project Director:	Joseph P. Winnick
Project Coordinator:	Francis X. Short
	George Lawther (1993-94)
Graduate Assistants:	Kevin Biata
	Mary Powers
	Rob Korzeniewski
	Kevin Wexler
	Lori Erickson

Office of Special Education and Rehabilitative Services

Project Officer:	Melville Appell
	Doris Andres

Project Target Advisory Committee
and Panel of Experts

Kirk J. Cureton, Ph.D. - University of Georgia
Harold W. Kohl, Ph.D. - Baylor Sports Medicine Institute
Kenneth Richter, D.O. - Medical Director, United States Cerebral Palsy Athletic Association
James H. Rimmer, Ph.D. - Northern Illinois University
Margaret Jo Safrit, Ph.D. - American University
Roy J. Shephard, M.D., Ph.D., D.P.E. - The University of Toronto
Julian U. Stein, Ed.D. (retired) - George Mason University

Preface

This is the final report for U.S. Department of Education project number HO23C30091-95 entitled **Project Target: Criterion-Referenced Physical Fitness Standards for Adolescents with Disabilities** conducted at the State University of New York, College at Brockport from June 1, 1993-May 31, 1998. The project was designed to extend the nation's current emphasis on health-related criterion-referenced fitness testing and programming to children and adolescents with disabilities. Essentially, it was the purpose of this project to develop a health-related criterion-referenced physical fitness test appropriate for youngsters with disabilities. This report summarizes project activities leading to the development of the **Brockport Physical Fitness Test** and related materials. Accompanying this narrative under separate covers are copies of the **Brockport Physical Fitness Test Manual**, the **Brockport Physical Fitness Test Technical Manual**, the **Brockport Physical Fitness Test Training Manual**, and the **Brockport Physical Fitness Test Instructional Videotape**, all developed as part of this project.

Reflecting back over the past five years, Project Target was the most ambitious, challenging, exciting, rewarding, and difficult professional endeavor with which we have been associated. Perhaps on the surface the task appears easy: select a few physical fitness tests appropriate for kids with disabilities and "pick a score" they must attain in order to pass. How hard can that be? Well, as it turns out, "pretty hard". Selected test items had to come from the components of health-related physical fitness (and, as such, have some relationship to health status); measure some important ability in the life of a youngster with a particular disability; be easily administered in a field setting; have some relationship to the **Prudential FITNESSGRAM** (where possible); have a scoring system to accommodate a wide range of abilities (as much as possible); have evidence of validity and reliability; and measure an ability that could be influenced (i.e. improved) through physical activity.

"Picking a score" (i.e. setting a criterion-referenced standard) for each test item was more challenging yet. A common criticism of criterion-referenced tests is that the standards often are established arbitrarily, consequently student success ("pass" vs. "fail") may be defined arbitrarily as well. Of course, we were quite aware of this potential pitfall throughout the project, but sometimes awareness alone was insufficient in selecting a standard to accurately represent the status of some dimension of health. Where possible we relied on research literature to establish standards. In particular we studied with great care the **Prudential FITNESSGRAM Technical Reference Manual** to understand the standard-setting process used for that criterion-referenced test. Even though the **Prudential FITNESSGRAM** was designed for nondisabled participants and had the benefit of decades of physical fitness literature to draw upon, test developers still had to rely on "expert opinion" to establish standards for many of the items in that battery. Faced with a much smaller research base on fitness for youngsters with disabilities, we too had to rely on "expert opinion" in many cases. But we had some very good experts! Our Advisory Committee and consultants were some of the very best people available in physical fitness, adapted activity, and measurement. So, while some of the standards, in fact, may be arbitrary via "expert opinion", that opinion was considered and well-informed.

Health-related physical fitness, for both disabled and nondisabled individuals, will be a fertile area for research and development for many years to come. Hopefully the **Brockport Physical Fitness Test** (and its related materials) will undergo scrutiny and evolve over that period of time. It is quite possible that this test (or a similar one) will look very different in 25 years than it does today. Nevertheless, the funding of Project Target likely will remain the cornerstone for criterion-referenced fitness testing for youngsters with disabilities. It is difficult to imagine a research effort of this magnitude being completed independent of external funding. Appreciation is extended to the Office of Special Education and Rehabilitative Services, U.S. Department of Education, for supporting Project Target.

Joseph P Winnick
Francis X. Short
Brockport, NY
May 15, 1998

Table of Contents

Appendices

Introduction

The positive value of physical fitness is widely accepted throughout the United States today. Physical fitness is important for enhancing the quality and length of life. It assumes an important role in the play and development of children and youth, enhancing their participation in leisure time pursuits and athletic endeavors. It helps in carrying out the requirements of daily living. Although physical fitness is important for all persons, it is particularly important for adolescents and young adults with disabilities. Where higher levels of physical fitness are attained, individuals are more likely to participate in play and sport activities and receive the same physical and social benefits from these activities as nondisabled adolescents and young adults. Because an enhanced physical fitness status helps persons in carrying out the requirements of daily living, it is of particular importance in the transition from school to the work place.

The physical fitness of youngsters with disabilities has been the focus of several investigations (for instance, Francis and Rarick, 1959; Howe, 1959; Johnson and Londeree, 1976; Rarick and Dobbins, 1972; Rarick, Dobbins, Broadhead, 1976; Rarick and McQuillan, 1977; Sengstock, 1966; Berg, 1970; Berg and Bjure, 1970; Brown, 1975; Dobbins and Rarick, 1975; Short and Winnick, 1986a; Short and Winnick, 1986b; Winnick and Short, 1984; Winnick and Short, 1988). These and other studies in the literature provide considerable evidence that the physical fitness performance of persons with disabilities during the developmental period of their lives is significantly below that of their nondisabled peers, that their physical fitness can be improved and, in some instances, approximate the performance of able-bodied peers.

Although the physical fitness of American able-bodied youngsters exceeds that of their peers with disabilities, there is agreement that the physical fitness of able-bodied youngsters is below desirable levels. The American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD) is a national professional organization which has attempted to deal with this situation. One of its principle objectives is to support, encourage, and provide guidance for the development and conduct of programs in health, leisure, and movement-related activities to meet the needs of individuals in today's society. In 1988, AAHPERD presented **Physical Best**, a new comprehensive physical education and assessment program designed to motivate all children and youth to participate in physical activity to develop their "personal best".

In 1994, AAHPERD entered into an agreement with the Cooper Institute for Aerobics Research (CIAR) whereby the **Prudential FITNESSGRAM** (CIAR, 1992) physical fitness test would supplant the **Physical Best** health-related fitness assessment instrument. The **Prudential FITNESSGRAM** was developed by the staff at CIAR. The intent of the agreement was to consolidate fitness testing and eliminate user-confusion in cases where items and standards were disparate. (The **Physical Best** educational and reward components were retained as part of the agreement under the title **Physical Best**).

The **Prudential FITNESSGRAM** provides (as did the **Physical Best** test) criterion-referenced standards. A criterion-referenced test is defined as a test with a predetermined standard

of performance that is tied to a specified domain of behavior. Standards represent criterion behaviors. Those whose test scores equal or surpass the standard are classified as "masters" on the test (i.e., "fit") and those with scores falling below are designated as "nonmasters"(i.e., "unfit"). In a criterion-referenced orientation there are no constraints on the number of individuals who can be classified as masters.

According to Safrit (1990), criterion-referenced tests have certain advantages over norm-referenced tests. Norm-referenced tests tend to measure the status of a normative group and give data as to "what is" rather than "what should be". Norm-referenced tests rarely provide the extensive diagnostic feedback which is associated with criterion-referenced tests. Criterion-referenced testing is most closely associated with the mastery model of instruction. In such a model, teachers care much less about differences between persons and much more on helping as many students reach a particular mastery level as possible. The focus is on establishing mastery standards which reflect levels of physical fitness which are desirable and helping as many persons attain this level as is possible.

The **Prudential FITNESSGRAM** is designed to measure health-related physical fitness; that is, "the dimension of physical fitness that is associated with health status. Five components make up health-related physical fitness: aerobic capacity, muscular endurance, muscular strength, body composition, and flexibility "(Morrow, Falls, and Kohl, 1994, p. 104). The components of health-related fitness and their corresponding test items from the **FITNESSGRAM** are discussed below.

- Aerobic capacity refers to the ability of an individual to take up and utilize oxygen to process metabolic fuels during exercise. The recommended test item is the one-mile walk/run. The PACER (a 20-meter multistage shuttle run) is the optional item.

- Body composition relates to the relative amounts of muscle, fat, bone, and other vital parts of the body. The sum of the triceps and calf skinfolds is the recommended test of body composition. An optional measure is body mass index.

- Muscular endurance refers to the ability of muscle groups to exert a submaximal force for a series of repetitions or successive exertions. Muscular strength relates to the maximal amount of external force that a muscle or group of muscles can exert. Flexibility is the range of motion available at a joint or a series of joints. These three components are sometimes referred to collectively as musculoskeletal functioning. Recommended tests include curl-ups, push-ups, trunk lift, and back-saver sit-and-reach. Modified pull-ups, pull-ups, flexed arm hang, and shoulder stretch are optional items.

The **Prudential FITNESSGRAM** manual includes some general recommendations for youngsters with disabilities. However even a cursory review of the test items suggests that further work is needed if the needs of persons with disabilities are to be met. For example, persons confined to wheelchairs are not able to run a mile to demonstrate aerobic capacity. They may need to move

using a wheelchair or become involved in another kind of sustained physical activity. There also is considerable evidence that a distance of one mile is too long for persons who are moderately mentally retarded. The sit and reach test item is obviously inappropriate for persons with spastic cerebral palsy or missing lower extremities. Curl-up tests are usually inappropriate for individuals with paraplegia, and often even one curl-up cannot be executed by adolescents with moderate mental retardation. The pull-up or modified pull-up test is used as an indicator of upper body strength and endurance. However, it is very apparent that there are some individuals with amputations who do not have available two arms in order to perform such a test. For these individuals, test items must be adapted or different items must be used.

The selection of different test items may be dictated by a set of health-related needs that pertain to youngsters with a specific disability and are different than those of able-bodied youngsters. Youngsters in wheelchairs, for example, may need the requisite upper body strength and endurance to lift their bodies from their chairs and/or to propel their wheelchairs up a ramp. These needs likely will influence test item selection.

In addition to test item selection is the need for logically and empirically-based levels of mastery associated with each (i.e., standards). For example, mile run standards developed for nondisabled youngsters would need to be adjusted for blind youngsters who must run with some kind of guidance (partner, wire, etc.) since such assistance increases the energy demand of the activity. Similarly, range of motion standards developed for youngsters with cerebral palsy would need to reflect a "functional" orientation rather than an "optimal" orientation inasmuch as the condition frequently restricts range of motion potential.

The purpose of this project, then, was to identify those items which could serve to assess the components of health-related physical fitness and to identify mastery standards relevant for individuals with disabilities when such standards are different from those associated with able-bodied or nondisabled populations. No other comprehensive and systematic research effort has been conducted for this purpose. **Prudential FITNESSGRAM** is the only testing program recognized by physical education's largest national professional organization (AAHPERD) and most likely reaches the greatest number of youngsters throughout the United States. For this reason, **Prudential FITNESSGRAM** was the program which was emphasized as a reference for the purpose of this project.

The term selected for this project's title was Project Target. Project Target reflected the need to establish valid criterion-referenced standards which can serve as a target for the attainment of a desirable state of health-related physical fitness for persons with selected disabling conditions.

Objectives

In order to accomplish the purpose of the project, the following objectives were established at the outset:

1) Identify and select populations requiring adjustments to physical fitness tests and standards.

2) Identify tentative test items and standards for physical fitness development.

3) Validate test items and standards.

4) Develop a criterion-referenced test of physical fitness and modify items and standards associated with selected national physical fitness tests designed for nondisabled individuals. (The national tests of physical fitness initially included the **Prudential FITNESSGRAM**, the physical fitness test of the President's Council on Physical Fitness and Sport, and the Amateur Athletic Union test. Subsequent to the grant proposal, however, the American Alliance for Health, Physical Education, Recreation and Dance adopted, as its only test of health-related physical fitness, the **Prudential FITNESSGRAM**. The project, therefore, was coordinated most closely with the **Prudential FITNESSGRAM**.)

5) Develop an educational component relevant to adolescents with disabilities.

6) Prepare materials for project dissemination.

These objectives are used as the basis for organizing much of the narrative contained in the body of this final report.

Personnel

Project Staff

Dr. Joseph P. Winnick developed the original grant application and served as project director throughout the life of the project. Dr. Winnick was assigned 75% time on the project during quarter 1 of each year (summer) and 60% during quarters 2, 3, and 4 (roughly equivalent to the academic year). This was the third research grant related to physical fitness testing of individuals with disabilities directed by Dr. Winnick. The first of these led to the publication of **Project UNIQUE: Physical Fitness Testing of the Disabled**.

The initial project coordinator was Dr. George Lawther. Dr. Lawther was hired under a "one-year emergency appointment" and served as coordinator during the first year of the project (1993-94). Dr. Francis X. Short, originally a member of the Advisory Committee, took over as project coordinator during the second year of the grant and served in that capacity until the completion of

the project. Dr. Short worked on the project the equivalent of 1.5 months during quarter 1 each year. For quarters 2-4 he was assigned 40% time to the grant. Dr. Short worked with Dr. Winnick as project coordinator for each of the previous fitness-related research grants conducted at SUNY, College at Brockport.

Support to project staff was provided by graduate assistants and secretarial help. Graduate assistants included Mary Powers, Kevin Biata, Rob Korzeniewski, Kevin Wexler, and Lori Erickson. The majority of the secretarial support came from Melissa Zurlo and Paul Plavetzki.

Advisory Committee and Panel of Experts

The project had a very strong Advisory Committee which served as its panel of experts. Members of the committee/panel included:

Dr. Kirk J. Cureton (University of Georgia) who had been active with the development of the **Prudential FITNESSGRAM** and who wrote the technical reference material for the measures of aerobic capacity associated with that test;

Dr. Harold W. Kohl (formerly of the Cooper Institute for Aerobics Research and currently with the Baylor College of Medicine) who worked at CIAR (the developers of the **Prudential FITNESSGRAM**) and was a co-editor of the **Technical Reference Manual** for the **FITNESSGRAM**;

Dr. James H. Rimmer (University of Northern Illinois) an expert in adapted physical education and author of **Fitness and Rehabilitation Programs for Special Populations**;

Dr. Margaret Jo Safrit (American University), an expert in measurement and evaluation in physical education and exercise science and author of the chapter related to setting and interpreting criterion-referenced standards which appears in the **Prudential FITNESSGRAM Technical Reference Manual**;

Dr. Roy J. Shephard (University of Toronto and Brock University), a pre-eminent authority on physical fitness of persons with disabilities and author of **Fitness in Special Populations**;

Dr. Kenneth J. Richter (Mercy Hospital, Pontiac, Michigan), medical director of the United States Cerebral Palsy Athletic Association with a particular expertise in physical activity of persons with physical disabilities; and

Dr. Julian U. Stein (George Mason University - retired) who is a well-known authority in adapted physical education with extensive experience particularly in the area of mental retardation.

Consultants

Project staff also drew upon the expertise of a number of consultants who are authorities in their field. Among those who played significant roles as consultants to Project Target were:

Mr. Jeffrey Jones (President, United States Cerebral Palsy Athletic Association) and Dr. Michael Paciorek (Eastern Michigan University), who along with Advisory Committee member Richter, provided valuable information on fitness testing of youngsters with cerebral palsy;

Dr. Bo Fernhall (George Washington University) and Dr. Ken Pitetti (Wichita State University) who consulted on tests of aerobic functioning and were principal investigators of a validation study on Project Target measures of aerobic capacity using subjects with mental retardation;

Dr. Paul Surburg (Indiana University) who critiqued Project Target field tests and accompanying rationale, and contributed to the flexibility/range of motion chapter in the training manual;

Dr. Pat DiRocco (University of Wisconsin, La Crosse) who provided commentary on Project Target field tests and contributed toward a chapter on muscular strength and endurance in the training manual;

Dr. Timothy Lohman (University of Arizona) who provided standards for certain measures of body composition and critiqued the chapter of the technical reference manual which included information on skinfolds and body mass index;

Dr. Jeff McCubbin (Oregon State University) and Dr. Georgia Frey (Texas A&M University) who contributed to a chapter on aerobic functioning in the training manual; and

Dr. Stephen Klesius (University of South Florida) who filmed the administration of virtually all Project Target test items and edited the instructional videotape.

Project Activities

Project activities are summarized in the following section of this report according to the original objectives of the grant. In some cases sub-sections are used to better organize the content.

Objective 1: Identify and select populations requiring adjustments to physical fitness tests and standards.

This objective was accomplished early in the project. The target groups included in this project were youngsters with mental retardation and mild limitations in fitness (MR); those who have

spinal cord injuries (SCI); those with cerebral palsy (CP); those who have either a congenital anomaly or an amputation (CA/A); and those who have visual impairments resulting in blindness (VI). (Each of these classifications is defined in more detail in the test manual that is appended to this report.)

Objective 2: Identify tentative test items and standards for physical fitness development.
Objective 3: Validate test items and standards.

Objectives 2 and 3 are closely related and are combined here for discussion purposes. The pursuit of these objectives constituted the "heart and soul" of the project. This section of the report, consequently, is more lengthy than other sections and is divided into three sub-sections: **meetings and consultation, training activities, and research activities.**

Meetings and consultation

There was a significant conceptual aspect leading to the development of the **Brockport Physical Fitness Test**. Over the life of Project Target numerous judgments were made based on expert opinion, clinical practice, and/or apparent logic. This was true particularly with regard to the selection and development of alternative (or nontraditional) tests of physical fitness for youngsters with specific disabilities. For this reason, no fewer than nine formal meetings were conducted during Project Target to address specific issues related to fitness testing for youngsters with disabilities. These important meetings included project staff, members of the Advisory Committee, and/or consultants with expertise in specific areas of physical fitness or particular disabilities. The proceedings of each of these meetings is summarized below in chronological order to provide some sense of the issues that were addressed and the changes that took place during the development of the test. Some additional information (agendas, notes) from these meetings can be found in Appendix A.

October 3-4, 1993. (Brockport, NY) Attendees: Short, Stein, Rimmer, Kohl, Cureton, Winnick, Lawther. This meeting was essentially designed to introduce the project to the Advisory Committee, to discuss the projects orientation and thrust regarding physical fitness, and to begin setting general foundations which undergird the project. As a result of this meeting, the 11 recommendations appearing below were formulated:

1. The definition and/or conception of physical fitness used in this study should be comprehensive.

2. The Chrysler Fund - AAU Physical Fitness Test (health related aspects) should replace the AAHPERD Health Related Physical Fitness Test as the fourth national test to be made "accessible" through Project Target activities.

3. The results of the project should be made usable and user friendly to the practitioner.

4. The validity contingency coefficient of .80, specified in the original proposal as a standard for accepting cut off scores, should be reduced to .60, because a coefficient of .80 would be difficult to attain given the variability of the population associated with this project.

5. The conception of physical fitness used by the project should embrace a) the identification of zones of performance reflecting desirable levels of health, including the need to reduce health risks, b) a functional element of physical fitness, reflecting physical fitness for the performances of daily living tasks, including work, and c) the attainment of a base level of fitness needed for participation in exercise, sport and other physical activities.

6. The project should use the following health-related components of physical fitness as a basis for the selection of fitness tests and standards: a) aerobic or cardiorespiratory endurance, b)flexibility, c) body composition, and d) muscular strength/endurance.

7. Alternate tests and standards of physical fitness should be personalized, but include test items and standards that relate to the four components of physical fitness, mentioned above, to the extent possible and appropriate.

8. Guidelines serving as the basis for a personalized physical fitness test should be established. The following factors should be represented in these guidelines: a) disability, b) physical activity interests and abilities, c) daily living activities, and d) functional needs.

9. To the extent possible and appropriate, avoid the development of physical fitness test items and standards based upon a medical-categorical and/or disability-specific approach.

10. Select alternate physical fitness test items and standards that reflect desirable physical fitness levels. Test items that are important for the development of physical fitness but are not measures of desirable levels of physical fitness, may subsequently be used in programs to develop physical fitness i.e., training programs.

11. Use a definition, or orientation, or criterion-referenced testing which emphasizes the attainment of mastery standards. Mastery standards selected should be established on the basis of expert opinion, empirical data, i.e. performance of elite athletes, etc.

April 14, 1994. (Denver, CO) Attendees: Short, Safrit, Rimmer, Kohl, Stein, Winnick, Lawther. This meeting included a number of topics. First, the committee discussed issues pertaining to the health-related standards for aerobic capacity and body composition. Eventually it was decided that the **Prudential FITNESSGRAM** standards would be adopted for these components of fitness. It also was decided that the aerobic capacity standards could be reduced for certain disabilities if

justification existed. Tentative standards for youngsters who are blind were also considered. In essence, the tentative standards for youngsters who are blind were not adjusted in any way except in cases where the lack of sight negatively affected performance (e.g., distance running). Another topic discussed was tentative test items for youngsters with cerebral palsy. Kohl suggested the development of functional test items linked to activities of daily living. Finally, one aspect of the conceptual basis for the project was discussed. This element, the personalized approach, is based on the needs and desired fitness profiles for individuals. The committee was very supportive of this approach.

May 25-26, 1994. (Brockport, NY) Attendees: Paciorek, Jones, Richter, Short, Lawther, Winnick. Several important outcomes resulted from this meeting. First, it was agreed that the USCPAA Classification system for athletic participation could serve as an important resource for the Project Target physical fitness test. Secondly, the conference participants provided much input to the project staff in regard to the establishment of test items and standards for adolescents with cerebral palsy. This information was summarized and is presented in Appendix A. Thirdly, very strong support was given to the project from USCPAA. In essence, USCPAA agreed to cooperate with the project in various endeavors. This was especially important for the future testing of individuals with cerebral palsy as the project progressed. Additional materials pertaining to this meeting appear in Appendix A.

April 21-22, 1995. (Brockport, NY) Attendees: Cureton, Richter, Rimmer, Fernhall, Winnick, Short. This meeting focused on the measurement of aerobic functioning. Very early on there was consensus that attempting to predict maximum oxygen uptake from field-based data for youngsters with physical disabilities would be very difficult, if not impossible, at this time. Consequently, much of the meeting was devoted to an analysis of the Target Aerobic Movement Test (TAMT) as an alternate measure of cardiorespiratory endurance. (In essence, the TAMT requires the testee to work within a pre-determined target heart rate zone for a period of 15 minutes.) Additionally, the group discussed different approaches which might be used for adolescents with MR and a laboratory-based validation study was conceived. The agenda for this meeting as well as notes summarizing the meeting are included in Appendix A.

April 24, 1995. (St. Catherine's, Ontario) Attendees: Shephard, Winnick, Short. Advisory Committee member Shephard had been unable to attend the April 21-22 meeting in Brockport, but project staff were very interested in his opinions on aerobic functioning given his extensive background in this area. Consequently, Winnick and Short traveled to Brock University in Canada to meet with Shephard. Essentially Winnick and Short attempted to summarize the discussion of the Brockport meeting and asked Shephard to respond to some of the key points. Notes from this meeting also appear in Appendix A, but it seems fair to say that the central staff felt that Shephard's comments were positive regarding the directions established at the earlier meeting.

April 25, 1996. (Atlanta, GA) Attendees: Cureton, Stein, Rimmer, Fernhall, Winnick, Short. The agenda included a number of topics including the conceptual basis for the project, measuring the aerobic fitness of persons with mental retardation or blindness, and adjusting

musculoskeletal functioning standards for youngsters with mental retardation. A set of rough notes taken during the meeting appears in Appendix A. Perhaps the most significant discussion revolved around the issue of setting standards for MR youngsters on measures of strength and endurance. Project staff proposed that the standards that are recommended for the nondisabled be adjusted downward on a 2-4 year developmental basis for youngsters with MR in accord with the conclusion reached by Rarick in his work with MR subjects. Cureton suggested that such an approach would be flawed because people with MR never do "catch-up" with their nondisabled peers as a "developmental lag" approach would imply. He recommended instead that a "performance discrepancy" approach would more accurately reflect the differences between retarded and nonretarded youngsters on measures of strength and endurance. In such an approach a performance gap between retarded and nonretarded youngsters for a particular test item would be identified (e.g., a 25% difference) and standards would be adjusted accordingly . This approach was adopted for use in the project.

May 3-4, 1996. (Brockport, NY) Attendees: Richter, Shephard, Surburg, Winnick, Short. This meeting primarily dealt with the topic of assessing flexibility and range of motion in youngsters with physical disabilities. A set of rough notes taken during this meeting can be found in Appendix A. Considerable attention was given to the Active Range of Motion (AROM) test (which had been recently re-named the Target Stretch Test). In essence the group supported the use of this original test as a subjective measure of range of motion especially for use with youngsters with SCI or CP. A number of recommendations for improving the test and its presentation were discussed. It is also noteworthy that two new test items emerged from this meeting. It was suggested that a wheelchair push of some kind might be a good "functional" test for youngsters with class 2 CP. Also, Richter proposed a "reverse curl" for youngsters with SCI (low level quadriplegia). The test requires the youngster to lift a 1-pound weight one time using a "tenodesis" grasp. Both of these items are in the test battery. Finally, some discussion centered around the testing of subjects with CP at the upcoming Paralympic Games in Atlanta. (In reference to the recommendation to develop a wheelchair push item that emerged from this meeting, coordinator Short had some communication with Dr. Robert Waters of Rancho Los Amigos Hospital in Los Angeles during the summer of 1996. Dr. Waters has written extensively on the topic of energy expenditure and mobility. As a result, in part, of that interaction it was recommended that we would try a test that requires a youngster to negotiate a distance of 40 meters in 60 seconds or less and to do so at a heart rate that is below moderate intensity (i.e., a light intensity). Youngsters passing the test would demonstrate the ability to maintain a functional speed and the potential to sustain that speed over distances required in the community.)

March 20-22, 1997. (St. Louis, MO) Attendees: Safrit, Rimmer, Stein, Surburg, DiRocco, Winnick, Short. Actually this was a series of meetings where Advisory Committee members or consultants met with project staff either individually or in pairs to react to a draft of the test manual they had previously reviewed. The rough notes from these meetings can be found in Appendix A. It seems fair to say that the draft generally was well-received by the reviewers. Safrit, for instance, remarked "great job" and Rimmer thought the draft was "excellent." The greatest concern, however, seemed to be the length of the manual. More than once reviewers wondered whether a "short form"

was possible. In the absence of a short form it was recommended that the project staff work with the publisher to make the manual as "user-friendly" as possible. Suggestions included using colored paper or, possibly, tabs to help testers negotiate the manual. A number of editorial suggestions also were made.

April 18-19, 1997. (Brockport, NY) Attendees: Rimmer, Cureton, Stein, Shephard, Kohl, Richter, Winnick, Short. The majority of this meeting was devoted to flexibility/range of motion testing and standards. Specifically, a revision of the Target Stretch Test protocol whereby the scoring scheme would be revised from a 1-5+ system to a 0-2 system was adopted. In the revised system a "2" is meant to convey "normal" or "optimal" range of motion for a joint action; a "1" indicates a range of motion that is somewhat reduced, but still "functional;" and a "0" represents a more significantly reduced range of motion that is considered "nonfunctional." Considerable discussion centered around establishing "functional" standards for a variety of joint actions. A number of systems for establishing "functional" standards were reviewed including values from the American Medical Association and the New York State Department of Social Services. Ultimately, however, a "range of motion guide" that presents both "normal" and "functional" values for various joint actions was adopted by the group for use with the Target Stretch Test. This guide is used by physicians in the state of Michigan, but has little documentation to support it. It's use in Michigan, rather, appears to be part of an "oral tradition." The 40-meter push/walk test was considered and adopted as a measure of strength and endurance for certain youngsters with cerebral palsy. A "new" test item, the wheelchair ramp test, grew out of this discussion and was added to the battery. The meeting concluded with members of the group reacting to the draft of the test manual. Notes from this meeting can be found in Appendix A.

Training activities

Most of the field testing conducted during Project Target was accomplished by project staff with the assistance of graduate students or teachers working directly under their supervision. In some cases, however, it was deemed necessary to train professionals who would be collecting data on behalf of the project. During the grant period, five formal training sessions were conducted for the purpose of standardizing test procedures among individuals who would be involved in data collection for the project. On January 31, 1995, Short trained approximately 20 adapted physical education teachers from New York City in Brooklyn, N.Y. Nine of those teachers eventually tested subjects and those results were incorporated into the project. Next, Winnick and Short conducted a workshop for testers at the Michigan School for the Blind in Lansing on May 1, 1995. The eight testers who were trained were there in conjunction with the Michigan Sports Camp for the Blind taking place at the time. These testers, along with Winnick, Short, and some volunteers, tested 53 subjects at two sites.

Winnick traveled to Houston, Texas for a session on September 15, 1995 where he trained nine public school teachers. These efforts resulted in testing 38 youngsters with a variety of disabilities during the spring of 1996. The fourth training session took place on the Brockport campus in late September 1995 with graduate students in physical education serving as the trainees.

Both Winnick and Short conducted this session which was designed to prepare testers for data collection at the upcoming New York State Games for the Physically Challenged and for a future study at the Brockport Central School District. Finally, Winnick conducted a workshop in conjunction with the Illinois AHPERD conference on November 16, 1995. Since this session was held as part of a state conference, many attendees were there more for educational reasons than to specifically prepare to test subjects.

Research activities

Project staff (or their designees) collected data related to Project Target on 18 separate occasions over the life of the grant resulting in 16 research reports. Each of these reports can be found in Appendices B-Q. Although each of the studies had specific purposes of it's own, the general purpose of the research program was to acquire information regarding tests and standards. Each study was designed to look at one or more of the following aspects: the appropriateness of specific test items and procedures for specific target populations (i.e., the feasibility of the test); the reliability (including objectivity) of selected test items; the concurrent validity of selected test items; and the attainability of tentative standards by specific target populations.

A total of 1,542 subjects were tested as part of the project. Some demographic data on each of the research projects can be found in Table 1. As with the meetings and consultation section of this report, each of the 16 research reports is summarized below in chronological order within the categories of mental retardation, visual impairment, physical disabilities, and nondisabled in an effort to provide a sense of some of the issues that influenced the development of the BPFT.

Holy Childhood #1. Essentially, 11 tentative test items previously established for MR youngsters were administered to 55 subjects to determine the feasibility of the items and, where appropriate, the attainability of the test standards associated with the FITNESSGRAM. Results of the field testing suggested that body mass index (BMI), grip strength, back saver sit and reach, trunk lift, shoulder stretch and bench press can be appropriate criterion-referenced physical fitness tests for adolescents with MR. In fact, it appeared that FITNESSGRAM standards are attainable by those with MR on BMI, back saver, and trunk lift. Modified standards may be necessary for the others. The isometric push-up and bar hang both appeared to be appropriate items but are highly related, so the selection of just one item for the final battery should suffice. Subjects had difficulty with the modified pull-up and its elimination from the battery may be necessary. The feasibility of both the PACER and the six-minute run seemed good, but subject performance was especially poor on these two measures of cardiorespiratory endurance. This eventuality suggested that more attention be paid to measuring this component of fitness among adolescents with MR (See Appendix B).

Holy Childhood #2. Holy Childhood Study #2, conducted during the fall of 1994, sought to determine if youngsters with MR could perform the curl-up test employing FITNESSGRAM procedures (which include the maintenance of a cadence) and whether the poor PACER performance noted in Study #1 could be improved as a result of an eight-week training

Table 1. Research Conducted as Part of Project Target

Location	Dates	Subject #	Group
Holy Childhood (#1) Rochester, NY	Spring 1994	55	MR
Holy Childhood (#2) Rochester, NY	Fall 1994	40	MR
Holy Childhood (#3) Rochester, NY	Spring 1995	33	MR
NYC Public Schools (A) New York, NY	Spring 1995	72	MR
G. Washington University Washington, D.C.	Summer 1995	34	MR
Houston Independent School District Houston, TX	Spring 1996	38	MR CP
Holy Childhood (#4) Rochester, NY	Spring 1996	42	MR
Holy Childhood (#5) Rochester, NY	Spring 1997	38	MR
Holy Childhood (#6) Rochester, NY	Spring 1998	40	MR
NYC Public Schools (B) New York, NY	Spring 1995	51	VI
Michigan Sports Camp Lansing & Kalamazoo, MI	Spring 1995	53	VI
NYSGPC (#1) Brockport, NY	Fall 1994	24	CP SCI
NYSGPC (#2) Brockport, NY	Fall 1995	25	CP SCI
Northern Illinois University Dekalb, IL	Summer 1995	32	SCI
Paralympics Atlanta, GA	Summer 1996	13	CP
ESGPC (#3) Brockport, NY	Fall 1996	24	CP SCI
ESGPC (#4) Long Island, NY	Spring 1997	12	CP
ESGPC (#5) Brockport, NY	Fall 1997	12	CP SCI
Brockport Central Schools Brockport, NY	Fall 1995 - Spring 1996	904	Nondisabled
		Total 1542	

program. Field testing determined that the curl-up is potentially a difficult item for subjects with MR. The cadence was a problem as was the necessity of sliding the hands along the mat as the curl-up was performed. This item may need to be eliminated or modified, or may simply require more instruction and practice prior to testing. Results of the PACER training experiment were discouraging. Only half of the subjects improved their PACER scores following the training program and about 32% actually got worse after training. The cardiorespiratory area, therefore, remained unresolved (See Appendix C).

Holy Childhood #3. The purposes of this study essentially were to see whether adolescents with MR could learn to do the curl-up with more instruction and practice and whether a modified PACER (running a 16m course instead of a 20m course) might have some promise as a cardiorespiratory measure for this population. Results of the curl-up training aspect of the study were mixed. About half of the subjects were able to follow the test procedures, however, subjects who were not able to follow the procedures initially had difficulty learning the procedures subsequently. In addition, FITNESSGRAM standards are very difficult for this group; approximately 10% of the subjects in this study attained the standard. The modification of the PACER appeared promising. By reducing the length of a lap by four meters MR subjects were able to run almost twice as many laps and cover 1.5 times the total distance over the course of the test. These findings were encouraging because by increasing the duration of the activity there is greater potential that the test is actually measuring cardiorespiratory endurance (See Appendix D).

New York City (A). At about the same time that Study #3 was taking place at Holy Childhood, testers were collecting data on subjects with MR (n=72) in the New York City Public Schools. Essentially the same 11 items that were administered in Study #1 were taken by the subjects in New York City (although the 600-yard run was substituted for the six-minute run and the curl-up for modified pull-up) in an effort to confirm the findings in the first study. There were, in fact, a number of similarities in the results of the two studies. BMI, sit and reach, trunk lift, shoulder stretch, bench press and grip strength all appeared to be "good" test items, although some modification in standards may be necessary. As with Study #1, a high degree of relationship was noted between the bar hang (later called extended arm hang) and isometric push-up suggesting that one could be eliminated from the final battery. Interestingly, the NYC teachers felt the curl-up test was appropriate and did not report the same problems as had been noted at Holy Childhood. NYC subjects also did better on the 20m PACER (12 laps completed versus 7), but the average length of the test was still under two minutes. Additional work with the 16m PACER seemed warranted. The 600-yard run actually appeared to be the superior measure of cardiorespiratory endurance in the NYC study (See Appendix E).

George Washington University. During the summer of 1995, Fernhall and colleagues conducted a validation study of the 16m PACER, the 20m PACER, and 600-yard run as measures of aerobic capacity (as estimated by laboratory-collected maximum oxygen uptake). Thirty-four subjects with MR were included in the study. Results were encouraging inasmuch as all three field tests were found to be very reliable and each of the field tests was found to be related to VO_{2max}. Accurate prediction of VO_{2max} appears to be enhanced if gender and BMI are included with a field test score in the regression equation. As was expected, however, the VO_{2max} values for the MR subjects were found to be below FITNESSGRAM standards thereby raising questions about the utility of measuring aerobic capacity for this group (See Appendix F).

Houston Independent School District (HISD). Twenty subjects with MR were tested on BMI, 16m PACER, 600-yard run/walk, Target Aerobic Movement Test, dominant grip strength, back saver sit and reach, trunk lift and shoulder stretch. Essentially the purpose of the testing was to confirm the feasibility of test protocols and the attainability of tentative criterion-referenced standards. Field testing supported the contention that, with practice, youngsters with MR can learn to perform appropriately the test items in question. Although subject numbers were low, results generally supported a reasonable level of attainability by youngsters with MR. The 16m PACER, however, was a notable exception. At least 50% of the subjects met standards used in the back saver (right and left), trunk lift, shoulder stretch, TAMT, and dominant grip strength. It is noteworthy that a total of 89% (17 of 19) successfully passed the TAMT. Conversely, only 25% (2 of 8) met minimum adjusted standards for the 16m PACER. (The HISD project also included nine subjects with CP; interested readers are referred to the synopsis in Appendix G for these results.)

Holy Childhood #4. The general purpose of this study was the continued field testing of selected test items of aerobic functioning (TAMT and PACER) and musculoskeletal functioning (curl-up, modified curl-up, 10-lb dumbbell press for ages 10-12, and 15-lb dumbbell press for ages 13-17). Forty-two youngsters with mental retardation and mild limitations in fitness participated in this study. One of the objectives of the study was to determine the test-retest reliability of two of the less common tests employed in the battery, the isometric push-up and the 15-lb dumbbell press. An alpha coefficient of .83 was calculated for the isometric push-up while coefficients of .98 were found for the dumbbell press, representing at least acceptable levels of reliability. Another important outcome of this study was the confirmation that the modified curl-up is an easier test to learn and perform than the curl-up test. Consequently, the modified version of the curl-up will be recommended for youngsters with MR (See Appendix H).

Perhaps the most interesting finding to come out of this study, however, was related to the PACER test. Youngsters with MR are notorious for poor performance on measures requiring aerobic capacity. Sometimes researchers suggest that at least part of the poor performance is due to motivational problems or an unwillingness to work through physical discomfort rather than (or in addition to) deficiencies in aerobic capacity. The subjects in this study, in fact, did poorly on the PACER, but heart rate monitors worn by the youngsters indicated that they were working quite hard at the conclusion of the test. Almost 70% of the subjects tested were working at approximately 85% of their maximum predicted heart rate, a value that usually serves as the dividing line between

"moderate" and vigorous" exercise. Low scores on the PACER, therefore, would seem to be due to low aerobic capacity rather than extraneous variables such as motivation and physical discomfiture.

Holy Childhood #5. In the spring of 1997 Project Staff returned to the School of the Holy Childhood in Rochester, N.Y. Although feasibility and attainability continued to be objectives of this study, the primary focus was to determine the reliability of selected test items and to determine the relationship between the 16m and 20m PACER tests. To this end, 38 subjects with MR were tested on two occasions on 16m PACER, dominant grip strength, 35-lb bench press, extended arm hang, flexed arm hang, and modified curl-up. Test-retest alpha coefficients were above .80 for extended arm hang and modified curl-ups. Coefficients above .90 were found for 16m PACER, grip strength, bench press, and flexed arm hang. Criterion-referenced reliability (i.e., consistency of classification, or P) was also determined for these items. P values above .70 were found for extended arm hang and modified curl-ups, above .80 for bench press and flexed arm hang, and above .90 for 16m PACER and grip strength. These test items seem to possess at least minimally acceptable levels of reliability for youngsters with MR. A Pearson r of .95 was calculated as an indication of the degree of relationship between the 16m PACER and 20m PACER (See Appendix I).

Holy Childhood #6. The final study conducted using youngsters with MR (n=40) as subjects once again took place at the School of the Holy Childhood in Rochester. Subjects were tested on two occasions (February 12 and February 26, 1998) on three test items: trunk lift, shoulder stretch, and back saver sit and reach. The purpose of the testing was to gauge the utility of the general standards for use with youngsters with MR on these items and to investigate the reliability of these items with this population. Graduate students trained in Project Target test procedures served as the testers. Testers administered the same tests on both occasions. Pass rates (sessions and genders combined) ranged from 92-98% for trunk lift, 46-69% for shoulder stretch, and 60-73% for back saver sit and reach. Alpha reliability coefficients were calculated for the shoulder stretch and back saver sit and reach. Depending upon the side of the body tested, coefficients ranged from .95-.96 for back saver and from .83-.94 for shoulder stretch. Proportions of agreement (P), a criterion-referenced estimate of reliability, was determined for all three items. P indicates the percentage of judgements (pass vs. fail) that remained the same over the two administrations of the test. P values were .89 for the trunk lift, and ranged from .83-.94 for the shoulder stretch, and from .89-.92 for the back saver. Kappa (k) is a statistic that adjusts P by removing the number of agreements that occur simply by chance to yield a more rigorous estimate of reliability. K values ranged from .67-.87 for shoulder stretch and .76-.83 for back saver sit and reach. (Due to the distribution of scores in the contingency table associated with the trunk lift, meaningful values for alpha and kappa could not be determined.) It was concluded that the general standards are appropriate for youngsters with mental retardation and mild limitations in fitness for all three items and that the items possess a sufficient level of reliability for this population (See Appendix J).

New York City (B). As part of the study described earlier, 51 subjects with visual impairments were tested in New York during the spring of 1995. This effort represented one of two

investigations conducted with youngsters with VI during the project. The general purpose of each of these investigations was to field test items and standards for children and adolescents with VI. The following items were included in the New York study: BMI, curl-ups, push-ups pull-ups, flexed arm hang, trunk lift, shoulder stretch, sit and reach, mile run, and 20m PACER. Among other things, the data collected on these items indicated that the 20m PACER should be preferred to the mile run as a measure of aerobic functioning. In fact, one of the more interesting findings in this study was the low correlation (r=.30) between the mile run and the PACER despite the fact that both generally are considered to be measures of aerobic capacity. Although perhaps the "preferred" measure, the pass rates for the 20m PACER were still low suggesting that an adjustment to the standards may be necessary. Measures and standards for both flexibility and body composition, on the other hand, seemed to be appropriate for these subjects. Pass rates for the measures of muscular strength and endurance ran lower than expected, but since none of these items was very sight-dependent there appeared to be little logic for adjusting the standards. Rather it appeared that the lower scores were due more to poor fitness than poor sight (See Appendix E).

Michigan Sports Camp. The second study conducted with VI subjects took place in conjunction with sport camps run by the Michigan Association for Blind Athletes. Fifty-three youngsters were tested either in Lansing or Kalamazoo. Once again the purpose was to field test tentative items and standards. Test items included BMI, curl-ups, push-ups, pull-ups, flexed arm hang, trunk lift, shoulder stretch, sit and reach, 20m PACER, skinfolds, Canadian Aerobic Fitness Test, and the 600-yard run/walk. Data analysis confirmed, to a large extent, what had been observed in New York City. Measures and standards associated with body composition and flexibility, for instance, can be used with youngsters with VI without modification. Also, performance on the 600-yard run suggested that a one-mile run simply would be too difficult (and time consuming) for youngsters with VI, especially the younger ones (10-12). Once again, all things considered, the PACER appeared to be the "best" measure of aerobic functioning for these subjects. Regardless of the measure used, however, some adjustment in standards seems to be warranted to account for the apparent additional energy expenditure due to running with a partner. As in New York, subjects in the Michigan study also had difficulty with the strength and endurance items, especially those involving the arms and shoulders. Still, there would seem to be little justification for adjusting the standards (See Appendix K).

New York State Games for the Physically Challenged #1. Data on subjects with physical disabilities (primarily cerebral palsy and spinal cord injuries) were collected on seven different occasions resulting in five research reports appearing in Appendix B. Five of the seven occasions were at the New York State Games for the Physically Challenged (NYSGPC), later called the Empire State Games for the Physically Challenged (ESGPC). Subjects (n=24) were first tested at the 1994 regional NYSGPC in Brockport. Inasmuch as the 1994 study was the first conducted with physically disabled subjects, its primary purpose was simply to field-test many of the musculoskeletal (strength, endurance, and flexibility) items being considered at the time. Depending on classification (CP or SCI), subjects took any of three goniometry tests, four functional tests of flexibility, two additional FITNESSGRAM items, and three "original" tests of strength and endurance. Among other things, the results suggested that the flexibility/range of motion area will need greater attention for

youngsters with CP. The Apley test, for instance, appeared to be a "better" test of shoulder flexibility for CP subjects than the shoulder stretch test. Goniometry appeared to have fair potential with the CP group; the test was feasible for youngsters with CP, but administration was a bit cumbersome and standards derived for the test will have to be modified from normal expectations. Two of the "original" tests demonstrated good success with both CP and SCI subjects. On the other hand, two of the FITNESSGRAM tests, sit and reach and trunk lift, will probably be eliminated (See Appendix L).

New York State Games for the Physically Challenged #2. The 1995 Games study was primarily concerned with field testing the AROM test as well as the AMT (later called the Target Aerobic Movement Test). The active range of motion test and the aerobic movement test are two "nontraditional" fitness tests that appeared to have potential for those with physical disabilities. In relation to the AROM, data on functional independence for activities of daily living (ADLs) also was collected. Results of the AROM testing were encouraging. Inter-rater reliability for most items was strong (86% overall agreement between two testers) and testers were able to subjectively classify (within 20%) a subject's ROM when compared to a measure of goniometry (93% accuracy). Furthermore, some of the AROM measures correlated with specific measures of functional independence. For instance, elbow extension had a phi coefficient of .69 with dressing upper body and wheelchair transfers, while shoulder extension had a phi of .75 with six different ADLs. Relationships such as these suggest that it might be possible to establish flexibility standards based upon a level of functional independence. Results of the AMT were also encouraging. Twenty of the 25 subjects who attempted the test met the criteria for "passing" suggesting that this might be the preferred measure of cardiorespiratory endurance for those with physical disabilities (See Appendix M).

Northern Illinois University. During the summer of 1995, project staff and Advisory Committee member Rimmer sought to learn more about the aerobic movement test. The AMT was administered on two occasions to 32 youngsters with spina bifida. Although some minor problems with heart rate monitors were experienced, the testing program demonstrated that the AMT is quite feasible with a population that uses wheelchairs for activities of daily living. Test-retest results suggested that the AMT was reliable and that the recommended standards are attainable for youngsters with this type of SCI (more than 90% of the subjects attained the criteria on each of the administrations). Results of this study were published in **Adapted Physical Activity Quarterly** (See Appendix N).

Paralympics and **Empire State Games for the Physically Challenged (ESGPC) #3.** In August 1996 project staff traveled to Atlanta, Georgia and tested 13 Paralympians with cerebral palsy. Two months later, 12 youngsters with CP were tested at the ESGPC in Brockport. (During ESGPC #3 five youngsters with SCI also were tested, but due to the low number, these test scores were not formally analyzed.) The data collected on these 25 subjects with CP were combined with the data collected on 13 other subjects with CP tested in Brockport during NYSGPC #2. Hence, 38 subjects with CP, tested over a 12-month period at three different sites, were combined for data analysis. One aspect of this research was to determine if the Target Stretch Test (formerly known

as the Active Range of Motion Test) possessed a reasonable level of inter-rater reliability (i.e., objectivity). Inasmuch as this test requires subjective judgements on the part of testers such a determination is important. The correlation coefficient for 175 paired observations was found to be .86 which was interpreted to represent an acceptable level of objectivity (See Appendix O).

Another element of this study was to determine the magnitude of the relationships between measures of musculoskeletal functioning and two indices of health. One index of health was self-reported functional independence on activities of daily living. The other was self-reported patterns of physical activity (where inactivity was considered to be a health-related risk factor). The correlation coefficients generated for these analyses were somewhat discouraging. Although there were a few exceptions, most coefficients between test items and indices of health were low to moderate. Some earlier pilot work had suggested that moderate to high relationships were possible. This finding suggests that project staff will probably have to rely on expert opinion when establishing criterion-referenced standards for measures of musculoskeletal functioning rather than on statistical relationships with measures of health.

Finally, project staff determined pass/fail rates for selected test items using the tentative standards which were in place at the time of the testing. Pass rates for the flexibility/range of motion tests and for the seated push-up generally were high (i.e., greater than 70%), but rates for the dumbbell press and grip strength were only moderately high (i.e., just over 50% for the minimal standards). Although some "tinkering" with the standards for youngsters with CP will probably still occur, these results suggest that the tentative standards are fairly appropriate for this group.

ESGPC #4 and **ESGPC #5.** A total of 24 subjects with physical disabilities were tested in 1997 in conjunction with the ESGPC. Twelve subjects were tested at the state games in Uniondale, N.Y. in May and 12 more subjects were tested at the regional games in Brockport in October. Of the 24 subjects tested, 19 had CP and five SCI. Originally the primary purpose of this study was to determine the test-retest reliability of selected items for youngsters with physical disabilities, but this objective was abandoned due to low subject numbers. Instead, some attainability data were collected and the administration of a number of items was videotaped as a "dry run" for an instructional videotape to be developed as part of the grant. Perhaps the most interesting aspect of this study was that four testers (three of whom had very little training) viewed the administration of a number of items from the Target Stretch Test on videotape and scored the items. Subsequent to their subjective scoring, coordinator Short used a goniometer on the still images from the videotape to determine a "criterion score." A total of 80 comparisons between subjective scores (20 scores for each of four testers) and criterion scores were made. The subjective scores matched the criterion scores 85% of the time indicating an acceptable level of criterion-related validity (See Appendix P).

Brockport Central School District. During late 1995 and early 1996, 904 nondisabled youngsters were tested on at least one test item related to muscle strength and endurance (most youngsters took a variety of the tests). The primary purpose of this project was to establish normative data on test items that might be considered "nontraditional" or "alternative" measures of fitness. Such items might have relevance for youngsters with disabilities, but no standards in the

literature by which to judge performance. These items included the 35-lb bench press, the 15-lb dumbbell press, the 10-lb dumbbell press, the extended arm hang, and the isometric push-up. Three more traditional measures (grip strength, flexed arm hang, and push-ups) were also taken to determine relationships with the nontraditional tests. Finally, the bench press was administered a second time to a sub-sample of boys in an effort to determine test-retest reliability (See Appendix Q).

This project was successful in that project staff was able to establish percentile values for boys and girls aged 10-17 for each of the test items. In the absence of research linking a specific level of strength/endurance to a specific index of health, test developers frequently rely on expert opinion to establish criterion-referenced, health-related standards. Expert opinion, in turn, is often influenced by norm-referenced data as was the case with the development of certain standards associated with the **Prudential FITNESSGRAM**. Descriptive data for each of the test items, intercorrelations for the traditional and nontraditional items, and results of the bench press reliability study can be found in the research report.

Objective 4: Develop a criterion-referenced test of physical fitness and modify items and standards associated with selected national physical fitness tests designed for nondisabled individuals.

This objective was attained with the development of The **Brockport Physical Fitness Test** which is appended to this report (as Appendix R) under separate cover. The test manual includes the following six chapters: Introduction, The Conceptual Framework, Using the Brockport Physical Fitness Test, Profiles, Test Selection Guides and Standards, General Recommendations for Test Administration and Test Items, and Testing Youngsters with Severe Disabilities.

A second document, **The Brockport Physical Fitness Test Technical Manual**, also is appended to this report (as Appendix S) under separate cover and supports the test manual as a reference by providing available information on the validity, reliability, and attainability of the test items and standards. Much of the information presented in the technical manual (as with the test manual) comes from the various meetings and research activity described in the previous section of this report. The technical manual includes the following five chapters: Introduction, Aerobic Functioning, Body Composition, Muscular Strength and Endurance, and Flexibility/Range of Motion.

Objective 5: Develop an educational component relevant to adolescents with disabilities.

The educational component of the project was conceptualized as consisting of three elements: a training manual, an instructional videotape, and a computer program to assist with test administration and record keeping. The training manual tentatively entitled the **Brockport Physical Fitness Test Training Manual** is appended to this report (See Appendix T) under separate cover. It consists of principles and activities for the development of health-related physical fitness and suggestions for adapting the principles and activities for the target populations covered by this

project. As mentioned earlier in this report, four adapted physical activity professionals from higher education contributed to chapters in the training manual under the direction of the project staff. The four included Dr. Jeff McCubbin with Dr. Georgia Frey (aerobic functioning and body composition), Dr. Pat DiRocco (muscular strength and endurance), and Dr. Paul Surburg (flexibility and range of motion). In addition to these chapters, the training manual includes an introductory chapter, a section with annotated written and audiovisual resources, a bibliography and appendices.

The instructional videotape also is included as an appendix to this report (See Appendix U). Dr. Stephen Klesius came to Brockport on October 31, 1997 to interview project staff and film the administration of Project Target test items for possible inclusion in a video presentation as part of a separate federally funded research project. Graduate assistants in adapted physical education (Lori Erickson, Michelle Shea, Tim Coyle, Wendy Kohler, Stephanie White, and Ellen Gill) and coordinator Short served as testers and a number of youngsters with and without disabilities served as participants. Filming took place over a two-day period. Dr. Klesius subsequently agreed to work as a consultant to the project for the purpose of editing the tape into an instructional video. Winnick and Short shot additional footage with Dr. Klesius on February 24, 1998 and he used the facilities at WUSF (University of South Florida) to construct the final product.

As a part of grant activities, a foundation for a test administration software component was developed. As will be discussed under Objective 6, Human Kinetics Publishers, Inc. has agreed to publish the test, technical, and training manuals. At this time plans call for an integration of some or all of the manuals as part of a software package that will probably employ CD-ROM technology. By using the software package, it is anticipated that testers will be able to do the following tasks: enter students into a database that will manage student records; classify students according to age, gender, and disability; select test items based on classifications; reference test validity and reliability documentation via links to the technical manual; review test item protocol; enter student test scores; compare test scores to previous administrations and to criterion-referenced standards; and summarize student test performance. Due to the relative complexity of the project (six target populations with specific definitions and classifications, 27 different test items, multiple sets of standards, etc.), the development of a user-friendly software package is viewed as a critical aspect of dissemination.

Objective 6: Prepare materials for project dissemination.

Information about Project Target has been (or will be) disseminated in both "formative" and "summative" ways. A number of presentations and publications related to this project were made over the life of the grant. These formative efforts were important in sharing project directions with professional audiences and receiving feedback on the development of the test. A list of presentations made during the grant follows.

Winnick, J.P. (1993). "Measurement and Evaluation in Physical and Motor Ability in the Educational Process of Persons with Disabilities." Presented at an International Course on Sport and Physical Education for Disabled People, Cordoba, Spain, October, 1993.

Winnick, J.P. (1994). "Measurement and Evaluation in Adapted Physical Education: Present and Future." Presented at the Annual Meeting of the National Association for Physical Education in Higher Education, San Antonio, Texas, January 6, 1994.

Winnick, J.P. and Short, F.X. (1994). "Research Related to Criterion-Referenced Physical Fitness." Presentation at the Annual Meeting of the National Consortium for Physical Education and Recreation for Individuals with Disabilities, Arlington, Virginia, July 14, 1994.

Winnick, J.P. (1994). "Project Target: A Criterion-Referenced Physical Fitness Program for Students with Disabilities." Presentation at the annual convention of the New York State Association on Health, Physical Education, Recreation and Dance, Sarasota Springs, New York, November 5, 1994.

Winnick, J.P. and Short, F.X. (1995). "Research Needs Related to Criterion-Referenced Health-Related Physical Fitness Tests for Adolescents with Disabilities." Presentation at the 1995 Conference of the National Association for Physical Education in Higher Education, Palm Springs, California, January 5, 1995.

Winnick, J.P. (1995). "Health-Related Criterion-Referenced Physical Fitness Test for Adolescents with Mental Retardation." Presented at the 4th national Conference on Adapted Physical Activity, Macomb, Illinois, March 23, 1995.

Winnick, J.P. and Short, F.X. (1995). "Criterion-Referenced Health-Related Physical fitness Testing for Individuals with Disabilities." Presentation at the annual convention of the American Alliance for Health, Physical Education, Recreation, and Dance, Portland, Oregon, March 31, 1995.

Winnick, J.P. (1995). "The Measurement and Evaluation of Physical Fitness in Adolescents with Disabilities." Presentation at the 10th International Symposium on Adapted Physical Activity, Oslo/Beitostolen, Norway, May 23, 1995.

Short, F.X. and Winnick, J.P. (1995). "A Health-Related Criterion-Referenced Physical Fitness Test for Adolescents with Spinal Cord Injuries." Presented at the Annual Meeting of the National Consortium for Physical Education and Recreation for Individuals with Disabilities, Alexandria, Virginia, July 15, 1995.

Winnick, J.P. "Health-Related Physical Fitness for Individuals with Disabilities." Keynote address at the 1996 International Symposium on Adapted Physical Activity. National Taiwan University, Taipei, Taiwan, May 15, 1996.

Winnick, J.P. "Personalized Health-Related Criterion-Referenced Physical Fitness Test Items and Standards for Individuals with Disabilities." Keynote address at the 1996 International Symposium on Adapted Physical Activity. National Taiwan University, Taipei, Taiwan, May 23, 1996.

Winnick, J.P. "Health-Related Physical Fitness of Adolescents Who Are Blind: Relationship to Sport." Presented at the Third Paralympic Congress, Atlanta, GA, August 15, 1996.

Winnick, J.P. and Short, F.X. "Standards for a Health-Related Criterion-Referenced Physical Fitness Test for Adolescents with Spinal Cord Injuries." Presented at the annual meeting of the National Consortium for Physical Education and Recreation for Individuals with Disabilities, Atlanta, GA, August 17, 1996.

Winnick, J.P. "Health-Related Criterion-Referenced Physical Fitness Measurement and Evaluation for Children and Adolescents with Disabilities." invited lecture at EWHA Women's University, Seoul, Korea, September 16, 1996.

Winnick, J.P. and Short, F.X. "Health-Related Criterion-Referenced Physical Fitness Measurement and Evaluation for Children and Adolescents with Disabilities." Presented at the annual meeting of the American Alliance for Health, Physical Education, Recreation, and Dance, St Louis, MO, March 24, 1997.

Winnick, J.P. and Short, F.X. "Taiwan Workshop on Physical Fitness for Youth with Disabilities," a three-day conference conducted at the National Taiwan Normal University, Taipei, Taiwan, May 20-22, 1998.

Publications related to the work of Project Target included the following:

Winnick, J.P. (1995). "Personalizing Measurement and Evaluation for Individuals with Disabilities." In Seaman, J. (ed.) Physical Best and Individuals with Disabilities: A Handbook for Inclusion in Fitness Programs. p. 21-30, Reston, VA: The American Alliance for Health, Physical Education, Recreation, and Dance.

Fernhall, B., Pitetti, K., Vukavich, M., Stubbs, N., Hensen, T., Winnick, J., and Short, F. (1996). Validation of cardiovascular fitness field tests in children with mental retardation. Medicine and Science in Sports and Exercise 28: 2 (suppl.) p. 550.

Winnick, J.P., Rimmer, J.H., Connor-Kuntz, F., and Short, F.X. (1996). The reliability of an aerobic movement test in children with spina bifida. Medicine and Science in Sports and Exercise 28: 5 (suppl.) p. 56.

Winnick, J.P., Rimmer, J.H., Connor-Kuntz, F., and Short, F.X. (1996). Feasibility of the Target Aerobic Movement Test in children with spina bifida. Research Quarterly for Exercise and Sport 67: 1 (suppl.) p. 124.

Rimmer, J.H., Connor-Kuntz, F., Winnick, J.P., and Short, F.X. (1997). Feasibility of the Target Aerobic Movement Test in children and adolescents with spina bifida. Adapted Physical Activity Quarterly 14: 147-155.

Fernhall, B., Pitetti, K., Vukovich, M., Stubbs, N., Hensen, T., Winnick, J., and Short, F. (1998). Validation of cardiovascular fitness tests in children with mental retardation. American Journal on Mental Retardation, 102(6), 602-612.

Summative dissemination will be accomplished primarily through Human Kinetics Publishers, Inc. (HKP), Champaign, IL. HKP arguably is the leading publisher of physical activity-related literature in the world and recently acquired the publication rights to **FITNESSGRAM**. (The Prudential Insurance Company, the previous sponsor, has withdrawn its support for future versions of the **FITNESSGRAM** program. **FITNESSGRAM** was developed by the staff of the Cooper Institute for Aerobics Research in Dallas, Texas.) HKP has agreed to publish the **Brockport Physical Fitness Test Manual** and related materials as part of their "national fitness testing program". Project staff are hopeful that the materials will have a 1999 publication date.

At this time it appears that the test and training manuals will be published in a traditional "print format," not unlike the way they appear in Appendix R and Appendix T. As discussed earlier, HKP also has agreed to develop software for the test to improve its utility and assist users with record keeping and interpretation tasks. Plans call for the technical manual (See Appendix S) to be reproduced in an "electronic format" on one or more compact (or floppy) discs as part of the software package. Human Kinetics also has expressed interest in working with the videotape developed as part of this project with the possibility of publishing it (or a variation) commercially.

Closing Statement

From June 1, 1993 until May 31, 1998 project staff worked on six objectives leading to a goal of developing a criterion-referenced health-related physical fitness test for youngsters with specific disabilities. The grant funded the equivalent of one full-time faculty member for this purpose during the college's academic year (quarters 2-4) and approximately 1.5 FTE faculty during the summer months (quarter 1). In addition, the grant funded one graduate assistant each year and covered necessary secretarial support. The grant also provided funds to support travel and stipends for work completed by an advisory committee and various consultants.

By the end of the funding period all six objectives had been achieved and the result was the development of the **Brockport Physical Fitness Test** and related materials. The related materials included the **Brockport Physical Fitness Test Technical Manual**, the **Brockport Physical Fitness Test Instructional Videotape**, and the **Brockport Physical Fitness Test Training Manual**. The test manual and related materials are provided as appendices under a separate cover to this final report. Materials developed as part of this grant will be available through Human Kinetics Publishers, Inc, Champaign, Illinois or the authors.

References

American Alliance for Health, Physical Education, Recreation and Dance (1988). Physical Best. Reston, VA: Author.

Berg, K. (1970). Effect of physical training on school children with cerebral palsy. Acta Paediatrica Scandinavica (Suppl. 204), 27-33.

Berg, K., & Bjure, J. (1970). Methods for evaluation of the physical working capacity of school children with cerebral palsy. Acta Paediatrica Scandinavica (Suppl. 204), 15-26.

Brown, A. (1975). Review: Physical fitness and cerebral palsy. Child: Care, health, and development, 1, 143-152.

Cooper Institute for Aerobics Research (1982). The Prudential FITNESSGRAM Test Administration Manual. Dallas, TX: Author.

Dobbins, D.A., & Rarick, G.L. (1975). Structural similarity of the motor domain of normal and educable retarded boys. Research Quarterly, 46, 447-456.

Francis, R.J., & Rarick, G.L. (1959). Motor characteristics of the mentally retarded. American Journal of Mental Deficiency. 63, 782-811.

Howe, C.E. (1959). A comparison of motor skills of mentally retarded children and normal children. Exceptional Children. 25, 352-358.

Johnnson, L., & Londeree, B. (1976). Motor fitness testing manual of the moderately mentally retarded. Washington DC: AAHPER.

Morrow, J.R., Falls, H.B., & Kohl III, H.W. (Eds.) (1994). The Prudential FITNESSGRAM Technical Reference Manual. Dallas, TX: Cooper Institute for Aerobics Research.

Rarick, G.L., & Dobbins, D.A. (1972). Basic components in the motor performance of educable mentally retarded children: Implications for curriculum development (Project No. 142714). Berkeley, CA: University of California, Department of Physical Education.

Rarick, G.L., Dobbins, D.A., & Broadhead, G.D. (1976). The motor domain and its correlates in educationally handicapped children. Englewood Cliffs, NJ: Prentice-Hall.

Rarick, G.L., & McQuillan, J.P. (1977). The factor structure of motor abilities of trainable mentally retarded children: Implications for curriculum development (Project No. 4233544). Berkeley, CA: University of California, Department of Physical Education.

Safrit, M.J. (1990). Introduction to Measurement in Physical Education and Exercise Science. St. Louis: Times Mirror/Mosby.

Sengstock, W.L. (1966). Physical fitness of mentally retarded boys. The Research Quarterly, 37 (1), 113-120.

Short, F.X., & Winnick, J.P. (1986a). The performance of adolescents with cerebral palsy on measures of physical fitness. In C. Sherrill (Ed.), Sport and Disabled Athletes. (pp.239-244). Champaign, IL: Human Kinetics.

Short, F.X., & Winnick, J.P. (1986b). The influence of visual impairment on physical fitness test performance. Journal of Visual Impairment and Blindness, May 1986, 729-731.

Winnick, J.P., & Short, F.X. (1984). The physical fitness of youngsters with spinal neuromuscular conditions. Adapted Physical Activity Quarterly, 1, 37-51.

Winnick, J.P., & Short, F.X. (1988). Project Unique II Final Report. Brockport, NY: State University of New York.

BROCKPORT

State University of New York
College at Brockport
Brockport, New York 14420

Physical Education and Sport
716) 395-2229

To: Frank Short, Roy Shephard, Jo Safrit, Julian Stein, Jim Rimmer, Harold Kohl, Kirk Cureton

From: Joseph P. Winnick, Ed.D.

Date: September 2, 1993

Re: Advisory Committee

I want to take this opportunity to thank all of you again for serving on the Project Target advisory committee. The first meeting of the committee is scheduled for October 3, 1993 (6:30 - 8:30 PM) and October 4, 1993 (8:30 AM - 3:30 PM). A tentative schedule for the advisory committee is presented on the attached sheet.

In order to prepare for the meeting, I am asking each of you to read the proposal which was written and approved, and the negotiation comments relative to the proposal presented in the last three pages attached. In your reading of the proposal, please jot down any reactions that you may have. On Sunday evening, I have reserved a portion of the program for each committee member to make a brief comment relative to the project. Your comments and reactions to the proposal or any other reactions and concerns that you may have relative to the objectives of the project or the methods and procedures employed by it are welcome. Another idea might be for you to identify your own efforts in relationship to criterion-referenced testing or your hopes or aspirations in regard to this particular topic. Essentially, I am thinking of a brief statement taking no longer than ten minutes for each committee member.

In order to be present for the meeting, it will be necessary for you to arrive at the Rochester Airport before 4:00 PM on Sunday. Make arrangements to depart on Monday, no earlier than 4:30 in the afternoon. It takes about 20 minutes from Brockport to the airport. I would like to have you make your own arrangements in terms of flights and let us know what these are in advance so that we might be able to make arrangements to pick you up and take you back to the airport. Arrangements for meals are up to you.

For your efforts, we will pay you a stipend of $100.00 for pre-conference work and $100.00 for Monday. (Unless receiving this kind of stipend violates some regulation for you). We will reimburse you fully for lodging (or make arrangements for you) and for your meals. It would be advisable for you to keep receipts for your airline ticket, tolls, parking, and meals.

If you have difficulty making a flight out of Rochester on Monday, please feel free to stay over until Tuesday morning. We will again make arrangements for you to get to the airport and will provide for your lodging on Monday night.

The reason each of you is on the advisory committee is because I feel that you have something to offer to make this project successful and to make it serve the needs of individuals with disabilities. As you think about the contributions that you can make, please feel free to take initiatives on behalf of the project. For example, if you have publications which you think would be important, please bring a copy with you. If you have a list of references which would be important in seeking data in relationship to the project, please bring this with you. Of course these are only two examples. The point is, please feel free to share with the project anything that you think would be important. I am looking forward to this very important meeting and having this group function together.

I have tried to cover all of the important points in relationship to the meeting, but if you have any other questions, please call me at (716) 395-2383 or the Project Coordinator George Lawther at (716) 395-2629. The fax number here is (716) 395-2246.

P.S. Before this memo was mailed it has come to my attention that staying over on Saturday night will save a ton of money. If you are willing and able, come on Saturday. We'll take a trip to Niagara Falls on Sunday. If you can't make it on Saturday, it's fine. Officially we start Sunday night.

Project Target
Tentative schedule for the Advisory Committee Meeting

Sunday, October 3, 1993

Arrive in Rochester, N.Y. before 4:00 PM

6:30 - 8:30 PM - Evening Meeting

- Introduction
- Orientation to the Project - J.P. Winnick
- Committee Member Briefs - each committee member
- What Others are Doing Related to C-R Testing - group discussion

Monday, October 4, 1993

8:30 AM - 12:30 PM - Morning Meeting

- Methods and Procedures
  - Definitions and Components of Physical Fitness
  - Reference Points for Fitness Standards
  - Test Items and Standards
  - Classifications
  - Discussion of Standards X Fitness Components

12:30 - 1:45 PM - Lunch

1:45 - 3:30 PM - Afternoon Meeting

- Strategies for Validation
  - Research Design
  - Data Collection
  - Data Analysis

3:30 PM - Departure

BROCKPORT
tate University of New York
ollege at Brockport
rockport, New York 14420
dapted Physical Education Projects
716) 395-2383

TO: Members of the Project Target Advisory Committee:
Dr. Kirk Cureton, Dr. Harold Kohl, Dr. James Rimmer, Dr. Roy Shephard, Dr. Frank Short, and Dr. Julian Stein.

FROM: Dr. Joseph P. Winnick, Project Director

DATE: October 13, 1993

RE: Project Target advisory committee meeting

The first meeting of the Project Target advisory committee was held on October 3rd and 4th, 1993 at Brockport, New York. Persons attending this meeting included Dr. Kirk Cureton, Dr. Harold Kohl, Dr. James Rimmer, Dr. Frank Short and Dr. Julian Stein. As a result of the discussions and deliberations, the following recommendations were made for the project:

(1) The definition and/or conception of physical fitness used in this study should be comprehensive, so that it encompasses transition goals related to physical fitness.

(2) The Chrysler Fund - AAU Physical Fitness Test (health related aspects) should replace the AAHPERD Health Related Physical Fitness Test as the fourth national test to be made "accessible" through Project Target activities.

(3) The results of the project should be made usable and user friendly to the practitioner.

(4) The validity contingency coefficient of .80, specified in the original proposal as a standard for accepting cut off scores, should be reduced to .60, because a coefficient of .80 would be impossible to attain. The variability of the population associated with this project is a factor in the selection of a validity coefficient standard.

(5) The conception of physical fitness used by the project should embrace a) the identification of zones of performance reflecting desirable levels of health, including the need to reduce health risks, b) a functional element of physical fitness, reflecting physical fitness for the performance of daily living tasks, including work, c) the attainment of a base level of fitness needed for participation in exercise, sport and other physical activities.

(6) The project will use the following health-related components of physical fitness as a basis for the selection of fitness tests and standards: a) aerobic or cardiorespiratory endurance, b) flexibility, c) body composition, and d) muscular strength/endurance.

(7) Alternate tests and standards of physical fitness should be personalized, but include test items and standards that relate to the four components of physical fitness, mentioned above, to the extent possible and appropriate.

(8) Guidelines serving as the basis for a personalized physical fitness test should be established. The following factors should be represented in these guidelines: a) disability, b) physical activity interests and abilities, c) daily living activities, and d) functional needs.

(9) To the extent possible and appropriate, avoid the development of physical fitness test items and standards based upon a medical-categorical and/or disability-specific approach.

(10) Select alternate physical fitness test items and standards that reflect desirable physical fitness levels. Test items that are important for the development of physical fitness but are not measures of desirable levels of physical fitness, may subsequently be used in programs to develop physical fitness i.e., training programs.

(11) Use a definition, or orientation, of criterion-referenced testing which emphasizes the attainment of mastery standards. Mastery standards selected should be established on the basis of expert opinion, empirical data (including contrasting groups) that validates test items and standards representative of desirable levels of physical fitness, norm-referenced data which provides justification for the selection of test items and standards, and other performance-based data, i.e. performance of elite athletes, etc.

I would like to express my appreciation to all of the members of the advisory committee for their willingness to be a part of this project. Special thanks are extended to those who were able to attend our first meeting earlier in this month. I feel that much was accomplished, and I will look forward to more valuable input from you all in the future, as the work on Project Target progresses.

BROCKPORT
tate University of New York
ollege at Brockport
rockport, New York 14420
dapted Physical Education Projects
716) 395-2383

TO: K.J. Cureton, Bill Kohl, Jim Rimmer, Margarit Safrit, R.J. Shephard, J. Stein

DATE: March 8, 1994

FROM: J.P. Winnick

RE: Project Target meeting in Denver

As you recall, you are all members of the Project Target advisory committee which also functions as the project's panel of experts. In connection to this later responsibility, I am mailing you 2 manuscripts in preparation for our meeting in Denver.

The first enclosure is a questionnaire which I would like you to read, fill out, and be ready to discuss. I am hoping this questionnaire will serve as a good starting point to see where we have agreements or disagreements about criterion-referenced health-related assessment. Eventually, we need to take a stand on the issues represented by entries in the questionnaire.

The second document represents my thinking about test items and standards related to blind adolescents. Although I have some reluctance in making statements about "a class of persons," it is necessary for the purposes of the grant. Basically in that document, I am saying that blind adolescents use the same test items and standards as sighted persons relative to criterion-referenced health-related physical fitness. I am taking this position even though their performance as a group may be less than average when compared to sighted persons. I am also saying that I would be happy to recommend different test items as long as they validly measure fitness components. I have no problem with giving tactual assistance during performance as long as such assistance does not significantly influence performance. I have included in that manuscript some information on the performance of blind youngsters. That information gives us some security that test items can be performed by this population and that the standards are attainable, i.e. they are realistic.

I realize that all of you will not be available in Denver. If you come to the meeting in Denver, you will be provided a $150 honorarium for your preparation and meeting time. For those of you unable to meet in Denver, I can provide a honorarium of $50 for those of you who 1) fill in the questionnaire and return it by April 20; 2) read and provide feedback you may have on the manuscript dealing with blind adolescents.

The meeting time for the Denver meeting is 8:00-10:30 A.M., April 14. The meeting will take place in my room at the Radisson Inn in Denver. I will also be available there until April 16.

I am looking forward to meeting with you.

xc: Dr. Short
Dr. Lawther

Distributed 4/22/94

April 21, 1994

Dear Project Target Advisory Committee Members: Dr. Cureton, Dr. Kohl, Dr. Rimmer, Dr. Safrit, Dr. Shephard, Dr. Short, Dr. Stein

I want to thank each of you for returning the questionnaire and/or meeting in Denver. A summary of our Denver meeting is attached. I have one minor request of you. Please mail me your copy of the "Tentative Health-Related Fitness Standards for Blind Adolescents" manuscript dated 3-8-94. Write any comments on the manuscript. (Dr. Stein has already given me his). I thank you in advance. I'll keep in touch.

Cordially,

Joseph P. Winnick, Ed.D.
Professor of Physical Education & Sport

Project Target
Advisory Committee Meeting
April 14, 1994

Attendance: J. Winnick, G. Lawther, F. Short, J. Safrit, J. Rimmer, W. Kohl, J. Stein

Agenda Item #1 **Discussion of Fitness Domains and Standards for Criterion-Referenced Physical Fitness**

The first step in the process was to record results on the Project Target Questionnaire. The second was to obtain committee agreement on issues surrounding health-related standards. In regard to the latter, the following was agreed:

a. The acceptable VO2 reference standard for aerobic capacity should be 42 ml./kg./min for nondisabled boys and 35 to 39 ml./kg./min. for nondisabled girls, ages 10-17. Project Target will adjust these standards for individuals with disabilities when there is physiological justification to do so.

b. The criterion-referenced standards relative to body composition will reflect optional to moderately high percent body fat (10-25% for boys and 12-32% for girls). Skinfold and body mass index measures will reflect these standards when healthy fitness zones are created.

Agenda Item #2 **Criterion-Referenced Health-Related Standards for Blind Adolescents**

As its second agenda item, the committee considered the tentative health-related physical fitness standards for blind adolescents presented by the project director in a paper dated 3-8-94. In regard to this topic, the following were agreed upon:

The committee generally agreed with the tentative standards presented in the document. However, before final adoption the committee suggests:

a. providing information from Project Unique indicating the percent of youngsters passing/failing standards

b. that the Canadian Aerobic Fitness Test (a step-up test estimating VO2 be used) in instances where there is concern about the use of the mile run or pacer because of guidance which may be necessary in these tests.

c. that the pacer be field-tested to determine the efficacy of guideropes or partner-assisted procedures. The 3-8-94 document will have to be revised to reflect 10-20% and 17-32% body fat standards agreed upon in item 1.

Agenda Item #3 **Reaction to Tentative Approach Regarding Adolescents with Cerebral Palsy**

For this part of the meeting the project director passed out a one-page initial attempt in identifying test items for individuals with cerebral palsy. The purpose of this part of the agenda was to get some reaction to current thinking regarding this area. The following points were noted:

a. Some tentative test items reflected a "performance assessment" which was encouraged by Dr. Safrit.

b. Dr. Kohl encouraged that activities of daily living be analyzed in order to select test items appropriate for this population. As an example, he encouraged the development of functional test items which relate to "getting out of a wheelchair" or the "transferring from a wheelchair to bathroom facilities."

Other

Throughout the meeting the committee indicated its support for the conceptual basis for project approaches and activities. Specifically, the committee supported a personalized approach based upon desired physical fitness profiles. Also, the idea of developing alternate test items to measure components of physical fitness was supported.

Project Target
Aerobic Functioning Meeting

State University of New York
College at Brockport

April 21-23, 1995

Attended by:

Kirk Cureton
Bo Fernhall
Ken Richter
Jim Rimmer
Frank Short
Joe Winnick

Agenda: Aerobic Functioning Meeting

State University of New York College at Brockport

April 21-23, 1995

Friday, April 21, 1995

3:00 • Introduction

3:15 • The Project Target approach applied to adolescents with cerebral palsy

3:45 • Project Target Aerobic Movement test
• Basic approaches for measuring aerobic functioning

4:15 • Aerobic tests for adolescents with mental retardation

4:45 • Organization

Saturday, April 22, 1995

9:00-1:00 • Agreement on tests and procedures - Aerobic Movement Test
• Agreement on validation study on MR
• Flexibility/ROM for Project Target

2:15-4:15 • Other tentative test items and standards
  • Spinal Cord
  • Amputations/Anomalies
  • Mental Retardation

PROJECT TARGET
AEROBIC FUNCTIONING MEETING
APRIL 21-22, 1995

SUMMARY

Most of Friday, April 21, was devoted to presentations by Short and Winnick designed to acquaint the conferees with important issues and approaches currently being considered for the project. Much of Saturday, April 22, consisted of discussion on the principle concepts. Highlights of those discussions are summarized below for a variety of topics.

Aerobic Movement Test Procedures

- There was general agreement that the AMT is probably the best approach available for assessing aerobic function for youngsters with CP and it has potential for other groups as well.
- In response to a question about what is measured by the AMT, Cureton suggested that the test measures "aerobic behavior" (rather than aerobic capacity, aerobic power, etc.)
- The interpretation is that a youngster who can pass the AMT is able to do the type of activity at an intensity and duration that is believed to promote good cardiovascular health.
  - This would require that a youngster work in their target heart rate zone for 15 minutes. (ACSM aerobic training guidelines)
- In response to discussion that seriously "detrained" disabled youngsters might find the 15 minute standard unattainable, the group also adopted a 6 minute functional standard.
  - The interpretation is that one who could pass the 6-minute standard can do the type of activity at an intensity and duration that promotes the ability to carry-out sustained activities of daily living.
  - In the absence of any literature support for such a standard it was recommended that the validity of the standard was established by a "panel of experts."
- Establishing the THRZ for SCI quads is more difficult than for other groups: A suppressed HR requires a lowered THRZ, but some might have a resting HR already in the THRZ.
  - Recommendations included eliminating the upper end of the THRZ and/or adding 20 b/m to the youngster's resting HR to establish the THRZ.
  - Another suggestion was to use 85-100 as the THRZ for those with resting HRs below 65 or to add 20 b/m for those with resting HRs above 65.
- Reduce predicted maximum HR by 15 b/m (rather than by 10) for those who use arms-only forms of exercise.
- Richter suggested that a reference to autonomic dysreflexia be included in the precautions section for those with spinal lesions above T6.
- Fernhall speculated that predicted maximum HR for those with Down Syndrome should be lower (<175), but there is no hard data to confirm at this time.
- When a HR monitor is used, timing should begin when the youngster enters the THRZ rather than use a standard warm-up time for all youngsters.

- When a HR monitor is not used the following procedures were discussed:
  - have a 3-minute warm-up
  - do pulse checks at 2, 4, 6, 9, 12, & 15 minutes after warm-up
  - take a radial pulse or tape a stethoscope to the chest
  - take pulse rates for 6 seconds

Validation Study with MR Subjects

- There was a sense that pursuing aerobic capacity (max VO_2) is still worthwhile for youngsters with MR.
- Fernhall has data on a half-mile run, but Winnick argued that the 600-yard run might make more sense because it is apparently of sufficient duration for a cardiovascular measure and motivational problems (resulting in more walking) would be lessened.
- Winnick also suggested that the PACER and modified PACER (16M) be investigated as possible tests of aerobic capacity.
- Fernhall agreed to look at the 600, the PACER, and modified PACER at GWU over the summer and early fall.
- Short (with some support from Cureton) questioned pursuing max VO_2 with this population on the grounds that even if it can be predicted we would not know the significance of the value.
  - others felt that if we could at least get a valid value a basis for a test standard could be determined later.

Flexibility/ROM

- The functional ROM tests seem worthy of pursuit based on group discussion.
- Richter made a few suggestions including:
  - use the Apley test as a precursor or screening test for the shoulder-related "clock tests" for classes 3-8.
  - use "shoulder elevation" as a functional test of ROM
  - if youngsters fail the Apley follow-up with shoulder elevation, external rotation, and horizontal abduction
  - make trunk rotation on optional item due to the prevalence of spinal fusion among SCI youngsters
  - consider a dumbbell press from the prone position
  - measure hip extension via the Thomas test
  - consider reducing the number of test items among the functional ROM tests because setting standards across CP classes will be difficult
  - suggested that Medicare guidelines may be one basis for establishing C-R standards

Spinal Cord Injuries

- Richter argued (with support from Rimmer) that the low level quadriplegia category be modified to C6-C8 inclusive
- Richter suggested that "shoulder elevation" was a better indicator of shoulder function than tests of flexion, extension, and horizontal adduction and abduction.
  - he suggested that the Apley be given first with shoulder elevation and external rotation as follow-ups
- Richter observed that those in the LLQ category could do shoulder ROM tests

Congenital Anomalies and Amputations

- Richter noted that hip adduction was sometimes a problem with lower limb amputees and suggested it be added to the ROM battery.

PROJECT TARGET
MEETING WITH DR. ROY SHEPHARD
APRIL 24, 1995

SUMMARY

Project staff met with Dr. Shephard in St. Catharine's, Ontario to discuss the assessment of aerobic functioning and the Project Target approach. Discussion followed a brief presentation and some of the points are listed below.

- in regard to the 6-minute standard for the AMT Shephard suggested that the word "functional" might not be a good name for the standard
- he suggested we find a better word than "fail" in the AMT test description
- he wondered if Borg's RPE scale could be helpful in monitoring the intensity of the AMT in the absence of the heart rate monitor (suggested 13-14 on the scale)
- Shephard seemed to like the idea of a 16m PACER for MR kids; he suggested that we might even use longer pauses because such an approach would better approximate the way children actually engage in aerobic activity (short bursts of energy interspersed with short periods of rest)
- he thought the CAFT might have applicability for MR youngsters and referred the staff to the Reid article on MR adults
- supported the use of a 600-yard run (rather than 800) for youngsters with MR
- seemed comfortable with the approach of adding 20 b/m to the resting HR of quads as a basis for establishing a THRZ (at least he did not object)
- agreed that a 15 b/m reduction in predicted maximum heart rate was a reasonable adjustment for those who engage in arm exercise
- he did not feel that youngsters with Down Syndrome should be expected to have lower predicted maximum heart rates than the general population
- he thought that a 10-second pulse rate check was better than 6-second for the modified AMT

Notes of
Advisory Committee Meeting
Atlanta, Georgia
April 25, 1996

Members Present: Cureton, Stein, Rimmer, Fernhall

I. Project Progress

A. Summary report by F.X. Short

B. Concern by Stein about definition/classifications of MR

II. Theoretical Basis for the Project

A. Consider taking "favorably and unfavorably" out of definition

B. Consider adding aquatics to PE/LTAs

C. Dependency and Inactivity might be changed to Independence and Active in overall flow chart

D. Consider deleting medically-related under Physiological H-R Fitness

E. "Epidemiological data" added to source of standards

III. Aerobic Fitness for the Blind

A. Terminology

1. Maximal Aerobic Power is technically more correct than capacity: Cureton; agreed with Shephard's recommendation, but said, "power" was poorly understood

2. Use "capacity" but explain that we know that power is more correct

3. Cureton suggested using the term Aerobic Functioning. Aerobic Capacity and Aerobic Behavior could be subcategories

B. 600 yard run/walk?

1. Cureton opposed: also thinks there are too many test items

2. Fernhall and Rimmer agreed to eliminate the 600

3. Analyze aerobic functioning items to make sure we are not redundant

IV. MSF and Mental Retardation

A. Stein: put BMI in English system for computer software

B. 2-4 year adjustment (Rarick)

1. Stein: use more recent data

2. Cureton: identify a deficit (20%?) make adjustments and leave it alone

3. Fernhall: don't adjust - there's no basis for it

4. Stein: be careful with Down Syndrome

5. Fernhall: yes, this is a concern

6. Cureton: these standards distort reality (see point #2)

7. Fernhall: MRs don't catch-up so a developmental lag approach is a problem; Cureton's idea is better

8. Winnick: look at difference of each age to compute the specific standard

V. Aerobic Functioning and Mental Retardation

harder to use weight as a rationale for adjusting AC standards for MR

- because more fat = less health

Cureton: why not just do the AMT

- more defensible
- don't force the AC standards
- Fernhall agrees

10-15% below predicted max HR: Fernhall

- use fick equation to calculate max VO^2 with a reduced HR

maybe get to the functional side

- job performance/physical activity

got to account for Down's: Fernhall and Stein

- physiological diff.: lower max HR

table for METs for a particular sport or work task

- adjust for age

Fernhall: could use a 16M PACER to predict future success with a job because job energy expense is known and you can "back calculate"

VI. Post Meeting Discussion

Consider using METs as the criterion for the PACER or the Modified PACER - relative to aerobic behavior for individuals with disabilities. A 6-MET reference appears most appropriate

Summary Comments
Project Target
Range of Motion/Flexibility
Meeting
May 3-5, 1996

1. Individualized standard for C1 = AROM recommended

2. Use a qualifying statement (footnote) to leave final decision on appropriateness to test administrator.

3. In Atlanta, be sure to record events of participating athletes.

4. In regard to the Apley test:
   a. Score of three expected for C3 to C8.
   b. Eliminate test item for C1 (with assistive device, WADS)
   c. C1 non-WADS can attain a 1.
   d. C2U - a score of two is expected.
   e. C2L - the Apley test is out for this group, i.e. delete.

5. Thomas Test:
   a. Only using Thomas test for hip extension. Use for C5 to C8. Use P/F scoring. C5 to C8 expected to pass.
   b. C1 to C4 - refer to AROM for hip extension.
   c. May have a modified scoring system on Thomas for C1 to C4 based upon degree.

6. Delete Timothy and Tucker tests.

7. Elbow and other joints - recommended AROM

8. Seated Push-ups - should emphasize weight off seat - rather than be overly concerned with straight arm position.
   a. Class 3 and Class 4 - can do it except for hemiplegia. Use a footnote stating not to use seated push-up for hemiplegic. Use 5-second criterion.
   b. Class 2 - lower cannot do it (C2U), delete it.
   c. Class 1 - cannot do it.
   d. 2U = indicate 1 to 2 seconds as criterion with statement saying that there should be a clearance.

9. Grip Strength
   a. 2U, 3-8
   b. Can use norm-referenced standards.
   c. 20th percentile for classes 4, 7, 8 is okay.

11. Wheelchair Push - may be particularly acceptable for Class 2U and 2L. Classes 3 and 4 may use dominant grip or dumbbell press (except for Athetoid). Another idea is to do dumbbell press by lowering rather than raising the dumbbell. If wheelchair push is used for classes 3 and 4, a time factor might be considered. Otherwise use 50-100 m/yd. and P/F.

12. Trunk Rotation - Delete it for SCI because it is more disability than fitness oriented.

13. Shoulder Stretch/Apley
    a. HLP, LLP, and C6 to C8 can pass Apley.
    b. Prefer Apley to Shoulder stretch for all groups.

14. Seated Push-up - SCI
    a. Keep 5-second criterion for HLP, LLP, and ambulatory category.
    b. Put O/TA for LLQ

15. Grip Strength - HLP, LLP, Ambulatory (20th percentile of normal is okay)

16. One-pound Reverse Curl (C6 - C8)
(active grasp, wrist extension, elbow extension)

17. Dumbbell Press/Bench Press
    a. OK using 10 lb. weight - but may consider 5 lb. weight - check data for dumbbell.
    b. SCI - no need for HLP and LLP distinctions.
    c. 20th percentile of normal OK for paras. and ambulatory on dumbbell and bench press.

18. AROM
    a. Add forearm pronation to AROM as an optional item for SCI.
    b. Delete horizontal abduction.
    c. Delete hip extension because we have the Thomas test.
    d. If you score a five on AROM, you might be "most functional".
    e. Require a warm-up.
    f. Delete the degrees on instrument.
    g. Use present 5-3-1 scoring system on present form
    h. positioning for CP - do the best you can
        hand position
        arm position
        rest of body position
        sit, stand, lie
    i. drawings
        1. get rid of leg, stumps - external rotation
        2. too much forward head on elbow extension (left side) picture
        3. shoulder extension - right side - take out lean

19. Atlanta Testing
    a. 85% will be classified
    b. Ken says no problem relating to classification
    c. needs:
        -desk
        -2 tables
        -consent forms
            -English
            -Spanish
            -Translator spare
        -seated push-up bars
        -need credentials - for visit areas
        -promotion - shirts
        -brochures - on Target

20. Recommendations for AROM
    a. SCI - for ambulatory paras. recommend hip extension and an ambulatory category.
    b. SCI - for quads. recommended shoulder abduction, shoulder external rotation, forearm pronation and a criterion score of 5.
    c. SCI - for wheelchair users, recommended shoulder abduction and external rotation and a criterion score of 5.

21. AROM recommendations for CP
    a. For ambulatory CP, check out knee extension.
    b. For Classes C2 - C4, refer to AROM or angle test on Thomas Test.
    c. For CPs, AROM uses individualized criterion-referenced standard (ICRS).
    d. Elbow extension, shoulder abduction, external rotation, supination are particularly important AROM test possibilities for adolescents with CP.

Advisory Committee Meeting
St. Louis, MO

March 20-22, 1997
Notes

Margaret Safrit (March 20)

-"great job"
-should try to reinforce validity and reliability information especially for the "recommended" test items
-not surprised that we had difficulty tying musculoskeletal functioning items into indices of health
    -Plowman had similar problems with back extensions and sit and reach
-suggested that we might try to factor analyze stretch test and ADL data
-certain aspects of the manual "stuck" in her mind
    -personalization
    -flexibility for teachers
    -task analysis
    -Target Aerobic Movement Test
-thought that the idea of providing both minimal and preferred standards was "very important" and "good"
-asked if she thought the TAMT was a test of physical fitness or physical activity, she said that the fact that intensity was so closely monitored suggested to her that it was a physical fitness test, but acknowledged that it might have elements of both

Jim Rimmer and Julian Stein (March 21)

-"excellent"
-"great accomplishment"
-both (especially Stein) made editorial suggestions
    -use of bullets
    -don't say "it is important"
    -eliminate quotation marks around key words, etc.
-define extent flexibility (pg. 24)
-include the BMI table from Rimmer's book to reduce the need for calculations
-if we include triceps-only skinfold data, we should probably indicate that it is the least desirable skinfold approach
-can the narrative be reduced?
    -is a "short form" possible?
-include the 10-17 age range in the introduction
-consider using "youth" instead of "youngsters" throughout
-include latex allergies as a safety concern
    -might be a problem with heart rate monitors
    -might use a yellow "warning" symbol throughout the manual
-probably should reference Bouchard & Shephard at least for parts of the conceptual model
-consider using 10 seconds (rather than 5) as the minimal score for the seated push-up

Paul Surburg and Pat DiRocco (March 22)

-need to define the terms minimal, preferred, recommended, and optional
-can the manual be reduced or made more user-friendly?
    -could use tabs or colored paper to help
    -maybe a better intro on page 34
        -don't need to read every page of the disability-specific profiles (some redundancy)
-field test the manual with a novice teacher
-Thomas test needs to be revised to consider the position of the lower leg as well as the upper leg
-both seemed to prefer the Michigan Medicare standards to the 20% criterion for the stretch test
-both preferred the 2-1-0 scoring system to the 5+ system
-some discrepancy/confusion existed between stretch test narrative and sketches
    -DiRocco provided some specifics
-might consider trying to get hooked-up with state associations as a vehicle for dissemination (i.e., in-service presentations)

Advisory Committee Meeting
April 18 - April 19, 1997

Members Present: Stein, Cureton, Rimmer, Shephard, Kohl, Richter

Agenda Item #1: Report of Project Activities from May 1996 to May 1997

Reactions:

*Dr Shephard asked project central staff to analyze mean performance in groups in addition to correlation coefficients as reliability studies are conducted.

*Dr Cureton suggested that procedures dealing with standards associated with the 16m PACER and the 9-minute run walk be carefully reviewed. His concern was that the results be accurately depicted in the manual. (Dr Winnick indicated that he would discuss and consult with Dr Cureton on this item in the near future following the meeting).

Agenda Item #2: Target Stretch Test

Reactions:

*Dr Short began this agenda item by bringing the group up-to-date on progress related to the development of the Target Stretch Test (TST). As a part of this presentation, he reviewed the results of field tests involving subjects with cerebral palsy (1995-96) in Brockport and Atlanta. He also presented extensive background data which might be used to identify a target standard for flexibility on the TST. Finally, he recommended a possible standard that could serve as the basis for "minimal" standards on the TST.

*Following considerable discussion, the Advisory Committee recommended the following in regard to a minimal standard:

-adopt the "functional" values associated with the "Michigan Medicare" document provided by Dr Richter

-each of these values would represent "a clinically accepted level of range of motion that is obtainable and meets minimum requirements for functional activity"

-Dr Kohl noted that these values should be viewed as conservative since they generally are used with adults; youngsters aged 10-17 are believed to be more flexible.

-Dr Richter suggested using "individualized standards" for classes C1-C2.

-in adopting these standards the committee felt that these values are currently in use and this "field testing" adds to their validity and also noted that these values roughly approximate other expressions of impaired ROM.

-The standards expressed as a percentage of optimal range of motion are as follows: shoulder abduction (72%), external rotation (86%), shoulder extension (88%), elbow extension (88%), pronation (72%), supination (75%), knee extension (92%), and wrist extension (75%).

Agenda Item #3: Discussion of the 40m push/walk, the Tucker Test, the Timothy Test, and a Wheelchair Incline Push Test for Individuals with Cerebral Palsy. Results of this part of the meeting resulted in the following recommendations:

Reactions:

*The Tucker and Timothy Test items need not be included in the Brockport Physical Fitness Test.

*In addition to heart rate, consider self-perceived exertion, the ability to talk during and following the test, tester observation of exertion, as ways of reflecting moderate effort during the "flying" 40m push/walk.

*Consider adding a "wheelchair push up an incline" as a recommended test item for C3 youngsters with Cerebral Palsy. The 40m push would be an optional item for this class if this is done.

*Use the 40m push as a recommended item for C2U and C2L.

*Use the dominant dumbbell press instead of the 40m wheelchair push for C3 individuals with severe hemiplegia.

*Add shoulder extensions as a TST item recommended for individuals with cerebral palsy (superscript 2 in Table 12 in the manual).

Agenda Item #4: Discussion of the Apley and Thomas Tests as modified

Reactions:

*Dr Richter expressed concern about "flabby thighs" in relationship to the scoring system of the Thomas.

*Dr Richter suggested adopting wording from Kendall, Kendall, & Wadsworth with regard to knee extension during the Thomas.

*Dr Richter recommended that the standard for classes C1 and C2L be combined and that a score of "2" serve as the standard.

Agenda Item #5: Reactions to the Proposed Test Manual

Reactions:

*The verbiage should be cut down (Stein)

*Bullet and other creative formats should be used for clarity and user friendliness (Stein)

*The manual should have colored sections to group relevant information (Stein)

*Include a convenient chart in the manual to compute BMI (Rimmer)

*Consider the possibility of using a loose-leaf notebook (Stein)

*Consider the possibility of a shortened version of the test manual (Richter)

*Include a step by step approach in selecting test items for subjects (Richter and Cureton)

*Liked the reference in the manual to Individual Education Program (IEP) (Richter)

*Put " " in front of key safety recommendations as test items are described (Kohl)

*Include a list of resources at the end of the manual (Kohl)

*Consider the possibility of putting tables related to standards at the end of the manual (Kohl)

*Have a section on general recommendations for test administration at the beginning of the test item section. (i.e., Encourage and reinforce performance) (Kohl)
*Review the manual for "unit of measurement" concerns. (Shephard)
*Include a section in the manual on warm-up, clothing, rest between trials and test items, etc. (Shephard)
*Be more specific about the meaning of "qualified" personnel (p.2) (Shephard)
*Consider a glossary of terms at end of manual. (Shephard)
*Use the term wheelchair users rather than confined to wheelchair in the manual (Rimmer)
*Use "motorized" wheelchair rather than "electric" wheelchair (Rimmer)
*Acceptable to eliminate the CAFT for blind youngsters (entire group)
*Several edits were identified relative to the TAMT (Shephard and Cureton)
*After careful analysis, consider reviewing nine-minute run (Winnick and Short)
*Consider eliminating section in the manual on Goniometer Test of Selected Items (Winnick)
*Delete Appendix A in the manual because it is already in earlier tables (Shephard)
*Strengthen the section on safety considerations (Shephard)
*Sport-related activities should not be conceptualized as "daily activities" since this is a primary distinction between health-related and skill-related fitness (Cureton and Stein)
*Include the USCPAA phone number for testers to call if they need help with classification, etc. (Richter)
*Consider the World Health Organization definitions (Richter)
*Maybe some of the conceptual framework can be moved to a technical report/manual (Kohl)

The reaction to the manual in regard to overall presentation, conceptual information, theoretical approach for personalization was very positive (all committee members).

file:admeet.497
disk:project target #4

Study #1
School of the Holy Childhood
Rochester, New York
Spring, 1994

Purpose and Overview

During the spring semester of 1994, a study was conducted at the School of the Holy Childhood, Rochester, N.Y. to field-test selected test items on a tentative criterion-referenced health-related physical fitness test designed for individuals with mental retardation and mild limitations in physical fitness. Field testing was conducted to determine if the tentatively selected test items would be appropriate in a test of criterion-referenced health-related physical fitness. Specifically, test items were administered to: 1) determine whether subjects were able to learn how to perform the test item in the time permitted; 2) determine whether subjects were able to score on test items to the extent needed to fall within a measurable achievement continuum; 3) determine the percentage of subjects reaching healthy fitness zones associated with the Prudential FITNESSGRAM; 4) compare results of males and females; 5) compare results of individuals of different ages; 6) compare results of individuals with and without Down Syndrome. The study was conducted from February to May, 1994.

The study included 55 subjects between the ages of 10 and 17. Data were analyzed on 9 boys and 15 girls with Down Syndrome (DS) and 9 boys and 11 girls without Down Syndrome (Non-Down Syndrome or NDS) for a total of 44 subjects. Eleven subjects were not included for data analysis because it was clear that they did not function within the definition of mental retardation employed in the study.

Methods and Procedures

Test items field tested included the 1) body mass index; 2) the Progressive Aerobic Cardiovascular Endurance Run (PACER); 3) the 6-minute run/walk; 4) right and left grip strength; 5) modified pull-ups; 6) the isometric push-up (ages 10-12); 7) the bench press (ages 13-17); 8) the back saver; 9) the trunk lift; 10) the shoulder stretch; and 11) a bar hang test.

The field test was conducted by two physical education teachers of the subjects, one graduate assistant on Project Target, and the project's coordinator and director. All activities were conducted under the overall direction of the project director. Test items were administered in accordance with the Project Target Test Manual dated December, 1993.

The subjects in the study were assigned to various classes. The frequency of physical education each week varied amongst the subjects as did the amount of time used for learning and performing test items by different subjects. Field-testers were instructed to introduce test items to subjects and teach subjects how to perform test items until the subjects were able to perform test items appropriately. Once field-testers were convinced that the subjects understood how to

perform the test items, data were collected and recorded.

Results

Since the essence of this project is to study specific test items in terms of a criterion-referenced health-related test for individuals with mild limitations in physical fitness, the first part of this section will be presented according to the test items field-tested. The basis for the comments presented there are based upon the raw data collected, mean performance of subjects on selected test items (see Table 1), the percentage of subjects attaining the 1992 Prudential FITNESSGRAM standards, the percentage of subjects scoring zero on test items, and the observations and reactions of field-testers and project staff.

Test Item Analysis

Body Mass Index (BMI). The body measurement of height and weight of subjects to determine BMI appeared to valid, reliable, and economical. Approximately 2/3 of subjects were within healthy fitness zones established by the Prudential FITNESSGRAM. The one exception was female subjects with Down Syndrome. A total of 40% of these subjects were within healthy fitness zones. A total of 89% (8 of 9) with scores outside of the healthy fitness zone had BMI scores over rather than below the acceptable zone. The percentage of nondisabled adolescents who "passed" the FITNESSGRAM criteria from the National Children and Youth Fitness Study (NCYFS) was at least 73% for females and 82% for males throughout the ages of 10-17 (Looney and Plowman, 1990).

PACER. Based on observations of the field testers, it was clear that the PACER was a test item understood by the subjects and which they performed appropriately with physical assistance. Unfortunately, the results of the subjects did not compare well with Prudential FITNESSGRAM standards. Every subject was able to complete at least one lap but only 3 of 44 subjects (7%) attained minimal healthy fitness zone values. Two of the three who reached the standard were NDS females with mental retardation. A total of 64% (28 of 44) of the subjects were able to complete 6 or more laps (less than one minute) which is 120 m.

Six-minute Run/Walk. Traditionally, scores of adolescents with mental retardation on long distance runs have not compared favorably with nondisabled peers. This has been attributed to low aerobic ability, motivation, and pacing. The results of this study are consistent with previous studies. Taken as a group, the subjects ran at a pace slightly below one yard/second. In order to reach the 20th percentile for boys and girls, subjects would have to run 3.9 yds/sec (males) and 1.8 yd/sec. (females) using the Fleishmann Test (Fleishmann, 1964) as a basis or 4.14 yds/sec (males) and 3.2 yds/sec. (females) using the 1976 AAHPER Youth Fitness test (Hunsicker & Reiff, 1976) as a basis.

Table 1
Mean Scores of Subjects on Selected Physical Fitness Test Items

Variable	Group			
	Non-Down Syndrome Females	Non-Down Syndrome Males	Down Syndrome Females	Down Syndrome Males
Number	11	9	15	9
Age	13.3	13.3	13.5	12.9
BMI	21.4	22.3	26.5	20.4
PACER (laps)	9.5	7.3	5.9	8.7
Six-minute Run/Walk (yds. completed)	705	720	565	673
Right Grip (kg.)	16.0	21.0	15.0	18.4
Left Grip (kg.)	15.6	18.2	13.7	16.2
Modified Pull-up (No.)	.82	1.6	.40	2.9
Isometric Push-up (Sec.)	8.0	22.6	10.0	15.0
Bench Press (No.)	4.7	18.8	8.7	15.8
Back Saver (cm.)				
Right	29.9	19.9	35.8	32.8
Left	30.2	19.6	36.5	33.3
Trunk Lift (in.)	10.5	9.0	10.7	10.4
Bar Hang (sec.)	12.8	12.4	13.2	28.2

BMI = Body Mass Index

Table 2
Number of Subjects Reaching Prudential FITNESSGRAM Standards

Variable	Group			
	Non-Down Syndrome Females	Non-Down Syndrome Males	Down Syndrome Females	Down Syndrome Males
BMI	7/11 (64%)	6/9 (67%)	6/15 (40%)	6/9 (67%)
PACER	2/11 (18%)	0/9 (0%)	1/15 (6%)	0/9 (0%)
Modified Pull-ups	1/11 (9%)	0/9 (0%)	0/15 (0%)	0/9 (0%)
Back Saver				
Right	8/10 (80%)	5/9 (56%)	13/15 (83%)	8/9 (89%)
Left	6/10 (60%)	5/9 (56%)	14/15 (93%)	9/9 (100%)
Trunk Lift	10/11 (91%)	6/9 (67%)	14/15 (93%)	9/9 (100%)
Shoulder Stretch				
Right	6/11 (55%)	4/9 (44%)	3/15 (20%)	7/9 (78%)
Left	4/11 (36%)	3/9 (33%)	4/15 (27%)	5/9 (56%)

Table 3
Number of Subjects Scoring Zero on Test Items

Variable	No. Taking Test	No. Scoring Zero	Percentage Scoring Zero
PACER	44	0	0%
Grip Strength			
Right	44	0	0%
Left	44	0	0%
Modified Pull-ups	44	24	55%
Isometric Push-ups	18	3	17%
Bench Press	25	4	16%
Bar Hang	44	2	5%

Test Item Analysis

Body Mass Index (BMI). The body measurement of height and weight of subjects to determine BMI appeared to valid, reliable, and economical. Approximately 2/3 of subjects were within healthy fitness zones established by the Prudential FITNESSGRAM. The one exception was female subjects with Down Syndrome. A total of 40% of these subjects were within healthy fitness zones. A total of 89% (8 of 9) with scores outside of the healthy fitness zone had BMI scores over rather than below the acceptable zone. The percentage of nondisabled adolescents who "passed" the FITNESSGRAM criteria from the National Children and Youth Fitness Study (NCYFS) was at least 73% for females and 82% for males throughout the ages of 10-17 (Looney and Plowman, 1990).

PACER. Based on observations of the field testers, it was clear that the PACER was a test item understood by the subjects and which they performed appropriately with physical assistance. Unfortunately, the results of the subjects did not compare well with Prudential FITNESSGRAM standards. Every subject was able to complete at least one lap but only 3 of 44 subjects (7%) attained minimal healthy fitness zone values. Two of the three who reached the standard were NDS females with mental retardation. A total of 64% (28 of 44) of the subjects were able to complete 6 or more laps (less than one minute) which is 120 m.

Six-minute Run/Walk. Traditionally, scores of adolescents with mental retardation on long distance runs have not compared favorably with nondisabled peers. This has been attributed to low aerobic ability, motivation, and pacing. The results of this study are consistent with previous studies. Taken as a group, the subjects ran at a pace slightly below one yard/second. In order to reach the 20th percentile for boys and girls, subjects would have to run 3.9 yds/sec (males) and 1.8 yd/sec. (females) using the Fleishmann Test (Fleishmann, 1964) as a basis or 4.14 yds/sec (males) and 3.2 yds/sec. (females) using the 1976 AAHPER Youth Fitness test (Hunsicker & Reiff, 1976) as a basis.

Grip Strength. Although the concept of squeezing or gripping to exert maximum force required teaching, subjects were able to perform appropriately and clearly obtain scores within an achievement continuum. The mean performances of male and female subjects with or without Down Syndrome were usually within 2 and no greater than 3 kg. from each other. In comparing mean values in this study with 20th percentile values attained by nondisabled subjects in the Project UNIQUE study (Winnick & Short, 1985), it is evident that the subjects with mental retardation are below the 20th percentile of their nondisabled peers. In view of the small number of subjects involved in the present study, it appears advisable to collect additional data before any conclusions are drawn.

Modified Pull-up. Subjects in the present study had considerable difficulty in learning to appropriately perform modified pull-ups. The mean performance for females was less than one completed and the mean performance of males was less than three completed. Only one student out of 44 was able to reach Prudential FITNESSGRAM minimum healthy fitness zone standards and over 50% of all subjects failed to appropriately complete one modified pull-up.

Back Saver. A test item in this battery on which this sample particularly individuals with Down

Syndrome, had success was the back saver. A total of 79% (34 of 43 subjects) reached or surpassed the Prudential FITNESSGRAM standards. A review of mean scores in Table 1 suggest that subjects with Down Syndrome were extremely flexible. In analyzing scores of subjects with Down Syndrome, there is concern about subjects being "too flexible." There may be a need to discourage performance beyond that set on the Prudential FITNESSGRAM or to set an upper healthy fitness zone limit.

On the back saver with right leg extended, 8 of 43 subjects had a reach of 40 cm or greater and 1 subject has a reach of 50. On the back saver with left leg extended, 7 of 43 subjects had a reach of 40 or over and 2 subjects had a reach of 50. The percentage of nondisabled adolescents who "passed" the FITNESSGRAM criteria from the NCYFS was at least 85% for males and 94% for females throughout the ages of 10-17 (Looney & Plowman, 1990).

Trunk Lift. The results in the study indicate that this sample performed the trunk lift with much success. A total of 89% (39 of 44) subjects reached the Prudential FITNESSGRAM criterion-referenced standard. Mean scores of subjects with DS exceeded slightly the scores of NDS subjects. For the present study it was acceptable for testers to hold the feet down during the trunk lift.

Shoulder Stretch. Although designed as an indicator of upper body flexibility, subjects in this sample had more difficulty with this item than other items measuring flexibility. Intercorrelations between these items and other items measuring components of flexibility were relatively low. In the total sample, less than half were able to pass the right shoulder stretch or the left shoulder stretch, respectively. Performance did not appear to be inhibited by difficulty in learning the task. A total of 45% (20 of 45) passed the right shoulder stretch and 36% (16 of 44) passed the left shoulder stretch. Interestingly, the mean performance of males slightly exceeded the performance of females.

Isometric Push-up. Based upon ability to learn the test item and to reach a score on an achievement continuum, the isometric push-up appears the be an appropriate test item for subjects in this sample. Only 3 of 18 subjects (17%) scored a zero on the test item. Typical in the comparison of males and females, the mean performance of males exceeded that of females. Testers agreed that careful attention needs to be given to teaching adolescents how to perform the test item.

Bench Press. The results on the bench press were similar to those on the isometric push-up test event though subjects who took the bench press test did not take the isometric push-up. The mean performance of males exceeded that of females. Only 4 of 25 students (16%) scored 0 on the test. In this sample, subjects needed to be taught the bench press. Many subjects required repeated instruction on different days before they were able to appropriately perform the test item.

Bar Hang. At the beginning of the study, the flexed arm hang was one test item administered to the sample to measure upper body strength/endurance. It became clear that this sample had considerable difficulty with the flexed arm hang and a bar hang was substituted. In the bar hang test the subject simply hangs from a bar as long as possible. Subjects are not required to hold a flexed arm hang over a bar. Interestingly, the mean performance of male subjects with DS exceeded that of the other subjects in the sample. Subjects clearly attained scores on an achievement continuum. Only 5% (2 of

44 subjects) were unable to hold for at least one second. Subjects had little or no difficulty learning to perform appropriately. A correlation coefficient of .70 was found between the bar hang and modified pull-up performance. Since there is a relationship between what these tests measure and since the bar hang greatly reduces the number of subjects scoring zero, the bar hang is a preferred test item.

Intercorrelations

The intercorrelations between test items are presented in Appendix B. For the most part, correlation coefficients support the notion that different items are measuring different aspects of fitness. An exception is the relationship between the isometric push-up and the bar hang. The $r = .92$ between these two test items suggest that the two tests are measuring the same component of physical fitness and thus, a test need not include both test items. As a matter of interest, the intercorrelations between test items measuring flexibility were relatively low. This supports the contention that flexibility is specific to areas of the body.

Summary and Conclusions

The results in this study support the use of the following test items on a criterion-referenced health-related test of physical fitness for adolescents with mental retardation and mild limitations in physical fitness: BMI, grip strength, back saver, trunk lift, shoulder stretch, and the bench press. The isometric push-up and the bar hang appear to be appropriate tests items but since they may measure the same aspect of physical fitness, both may not be needed. The modified pull-up appeared to be difficult for subjects to perform appropriately. The PACER is a test items which is quickly learned and performed appropriately. However, this sample was very low in performance when compared with the Prudential FITNESSGRAM healthy fitness zones. The six-minute run/walk appeared to be satisfactory but motivation and pacing are concerns about the test. Also, the subjects in this sample failed to reach standards associated with even the 20th percentile of nondisabled peers. The area of aerobic functioning remains problematic.

References

Fleishmann, E.A. (1964). Examiner's Manual for the Basic Fitness Tests. Englewood Cliffs, NJ: Prentice Hall.

Hunsicker, P. & Reiff, G.G. (1976). AAHPER Youth Fitness Test Manual - Revised Edition. Washington, DC: American Alliance for Health, Physical Education, and Recreation.

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), 3: 215-223.

Winnick, J.P., & Short, F.S. (1985). Physical Fitness Testing of the Disabled: Project UNIQUE. Champaign, IL: Human Kinetics.

Disk: Project Target#3
File: study1 9-18-97

Comments on Teaching Test Items
Holy Childhood Study #1

In administering the tentative criterion-referenced physical fitness test is was necessary for field-testers to help subjects understand and learn to perform test items. The information presented below discusses procedures used related to the process.

1. **The PACER**

   The students were directed to run the width of the gymnasium at the sound of a whistle. Ten seconds were allowed between whistles, and the students were instructed to wait at the sideline and not to run until they heard the whistle each time. Teachers and aides ran with the students, giving them directions and encouragement while they were running. Partners would run with the students and hold the hand of the students from time to time, when necessary. As learning progressed, students were asked to continue the exercise for longer and longer periods of time. Starting with a one-minute time period, the length of the exercise gradually increased to four minutes. Although not all of the students were able to continue for the full amount of time, they were encouraged to walk, and to continue participating with the rest of the class. After the exercise was completed, the students were instructed to continue walking as a cool-down procedure.

2. **Six Minute Run/Walk**

   A 100-yard rectangular running area was marked off using traffic cones at each corner. The students were instructed to run in a counterclockwise direction around the cones for a period of six minutes. If the students became too tired to run, they were told to walk until they were rested enough to run again, but to keep moving for the entire six minutes. One helper ran alongside each participant, giving directions and encouragement. The students were conditioned for the test by running for progressively longer periods of time, until they could continue the exercise for the entire six minutes.

3. **Modified Pull-up**

   The students practiced with the modified pull-up apparatus. The teacher emphasized proper technique and gave physical assistance from time to time to help students perform the test correctly.

4. **Isometric Push-up**

   The instructor first demonstrated and explained the correct procedure for executing an isometric push-up. Then, the student was asked to lie on a mat in a prone position. The hands were placed directly under the shoulders. The student was told to push up to a complete (but not locked) extension of the arms keeping the body in a straight line. The instructor checked to see that the elbows were not locked. Manual assistance was used to help the student achieve the correct front-leaning-rest position. A mirror was used to help students see their position.

5. **Grip Strength**

   Students first practiced squeezing nerf balls, one in each hand. They were then instructed to squeeze as hard as possible. The instructor then used a handshake to check the strength of each student's grip for the left hand and the right hand, and to determine whether or not the student was squeezing with a maximal effort. After determining that the student had learned the concept of squeezing as hard as he/she possibly could, the instructor demonstrated and explained the procedure for using the grip dynamometer. The correct grip setting was determined for each student. Students were then allowed

to practice with the grip dynamometer until they understood the concept of the test, being encouraged to put forth a maximum effort on each trial.

6. **35-Pound Bench Press**

With the student lying in a supine position on a mat, the teacher used a broomstick to teach the proper technique for the bench press. Manually, the teacher resisted the student's effort to push the bar upward. Correct form in executing the press was emphasized by guiding the broomstick while the subject pushed. The instructor then demonstrated and explained the correct technique for the bench press, using a weight lifting bench with a rack and a 35-pound barbell. Students were given an opportunity to practice while the instructor spotted and gave verbal direction, as well as manual assistance, when necessary. Spotters stood on each side of the participant to help the student learn the correct technique for the bench press.

7. **Bar Hang**

Students were given physical assistance to get the correct starting position on the bar. Many required correction for the overhand (pronated) grip.

8. **Back Saver Sit and Reach**

A warm-up routine, with stretching exercises, was used prior to performing the test (e.g. toe touches). In many cases, manual assistance was needed to help students attain the proper position for the test. The teacher demonstrated and explained what the students were expected to do. The students were then allowed to practice, while the teacher gave support through verbal direction, encouragement and physical assistance. The instructor placed a hand on the knee of the leg of the student being tested, to be sure the knee did not bend during the test.

9. **Shoulder Stretch**

After a stretching warm-up, the teacher demonstrated and explained the procedure for the shoulder stretch to the students. Then students attempted to perform the test. Physical assistance was given to help students understand the concept of the test.

10. **Trunk Lift**

A warm-up of stretching exercises for the trunk, preceded the testing process. The teacher demonstrated the trunk lift. Then the students were allowed to practice on a mat. Physical assistance was given, where needed, to insure that the proper procedure was followed. It was permissible for the teacher or a partner to hold the participant's feet during the test.

STUDY #2
SCHOOL OF THE HOLY CHILDHOOD
ROCHESTER, NEW YORK
FALL, 1994

Purpose and Overview

During the Fall of 1994, a study was conducted at the School of the Holy Childhood, Rochester, NY to 1) field-test the curl-up test item in the Prudential FITNESSGRAM test, 2) field-test the Progressive Aerobic Cardiovascular Endurance Run (PACER) test item in the Prudential FITNESSGRAM test, 3) determine if PACER test results change after eight weeks of training, and 3) compare heart rates of subjects prior to and following paced-run training sessions. The study began on October 13 and ended on December 16, 1995.

Subjects for the study included 40 adolescents with mental retardation and mild limitations in physical fitness. Subjects included 16 males and 24 females between the ages of 10 and 17. The study was conducted during the physical education classes of the subjects. A total of 5 classes were involved in the study. One class (class 1) was conducted on 18 days and the remaining classes were conducted for 43 days (see Table 1).

The field test was conducted by the two physical education teachers of the subjects in the study under the overall direction of Project Target staff. The Prudential FITNESSGRAM methods and procedures were employed in administering the curl-up and PACER test items. During testing and training, subjects were permitted to run with a partner in the PACER test. The physical education teacher and project staff were the only partners used.

Methods, Procedures, and Results

Curl-up Test

In regard to the curl-up test item, field testers were asked to first teach students to perform the test item correctly. Once students were able to perform the test item appropriately, field-testers were requested to test subjects to determine the number that they were able to perform.

In the process of implementing the methods and procedures for the curl-up test item, it was soon discovered that subjects had difficulty performing the test properly. Field-testers estimated that only 10 of the 40 subjects had learned and were able to correctly perform the test item. In view of this, it was decided to simply administer two trials of the test toward the end of the study. Trial 1, or a pre-test trial, was given between November 28 and December 2 and a post-test, consisting of two trials, was given between December 5 and December 9. Two trials were administered in the post-test to determine if results would be affected by holding the legs down during the execution of the curl-up test.

Table 1
Subjects X Number of Days of Physical Education

Classes	No. of Days	Males	Females
1st period	18	4	10
2nd period	43	2	4
3rd period	43	3	2
7th period	43	4	7
8th period	43	3	1
	Total	16	24

The results of testing the 10 subjects who learned to perform the test correctly appears in Table 2. As a group, these subjects improved from pretest (16) to post-test (21) and performed more curl-ups with feet held down (26) than not held down (21) in the post-test. Of these subjects 2 males and 1 female (with feet not held down) reached minimal health fitness zone standards associated with the Prudential FITNESSGRAM.

Table 2
Test Results on the Curl-up Test

Subject No.	Age	Gender	Pre-test (feet not held)	Post-test (feet held)	Post-test (feet not held)
1	16	F	5	20	7
2	17	F	2	13	0
3	14	M	10	12	12
4	17	M	22	25	20
5	13	M	13	75	*39
6	11	M	12	10	*26
7	14	F	6	8	10
8	13	F	4	24	3
9	13	M	13	30	16
10	16	F	75	45	*75
Average			16	26	21

***Subjects reaching Prudential FITNESSGRAM standards in the post-test (feet not held down)**

Based upon the observations of the Project Target staff and field-testers, the following additional comments appear warranted:

1) The recommended cadence to perform approximately one curl-up every three seconds appeared to be a problem for this sample. Some subjects preferred performing faster and others would have enhanced performance if they would have been permitted to perform more slowly. It was suggested that it may be advisable to permit greater flexibility in regard to cadence.
2) The learning of the curl-up was inhibited by the fact that this sample has learned to perform the sit-up using related but clearly different procedures. This appeared to result in negative transfer of learning.
3) The Prudential FITNESSGRAM test procedure suggests that a helper place the hands under the head of subjects as they return to the start position and that another tester stand facing the subject during test administration. Field-testers concluded that holding hands behind the head had no value and that the procedure may have distracted the participating subjects.
4) Subjects in this sample had considerable difficulty in inhibiting the use of the arms over their head to give them momentum in performing a curl-up.
5) It was difficult to teach students not to come to a full sit-up position in the upward movement in the curl-up because of these previous ways of performing the sit-up.

PACER TEST

In regard to the PACER test, field-testers were requested to teach students to perform the test accurately. Whenever necessary, subjects performed the test item with the assistance of a partner - usually the field tester.

The ability to perform the PACER test (usually with assistance) was demonstrated early. The pre-test for this test was administered between October 24 and October 28 and the post-test was administered between December 12 and December 16. In between these tests, a training program was administered to each of the classes for eight weeks. Class 1 experienced training twice each week and the remaining classes experienced training on a daily basis.

The eight-week training program was established to determine if change in the subject's performance would occur between the pre-test and post-test. The program, summarized in Table 3, involved increasing the number and intensity of PACER laps. In the first week, the program involved performance of 6 laps at nine seconds each and two laps at eight seconds each for a total training time of 70 seconds. During training, a whistle was blown to correspond to nine or eight-second lap intensities. Students were encouraged to complete the 20-meter distance in the PACER test within the time periods specified. If they were unable to do so, they were required to keep moving (jog or walk). Students were instructed to change directions whenever the whistle was blown. Clearly, subjects did not always complete laps before the whistle signal was given to change directions.

Table 3
The PACER Training Program

Week	No. and Time of Laps	Duration (sec.)
October 13-14	6 laps @ 9 seconds	54
October 17-20	6 laps @ 9 seconds	54
	2 laps @ 8 seconds	16
October 24-28	6 laps @ 9 seconds	54
	3 laps @ 8 seconds	24
November 1-4	6 laps @ 9 seconds	54
	4 laps @ 8 seconds	32
November 7-10	6 laps @ 9 seconds	54
	5 laps @ 8 seconds	40
November 14-18	6 laps @ 9 seconds	54
	6 laps @ 8 seconds	48
November 21-23	6 laps @ 9 seconds	54
	7 laps @ 8 seconds	56
November 28-December 2	6 laps @ 9 seconds	54
	8 laps @ 8 seconds	64
December 5-9	6 laps @ 9 seconds	54
	8 laps @ 8 seconds	64
December 12-16	6 laps @ 9 seconds	54
	8 laps @ 8 seconds	64

Table 4
Test Results on the PACER Test

	No. of Subjects	Average Score (laps)	No. of Subjects Reaching Prudential FITNESSGRAM Standards
Pre-test			
Males	16	8	0
Females	22	6	1
Total	38	7	1
Post-test			
Males	16	9	0
Females	23	7	0
Total	39	8	0

However, the results suggest that the training program was inadequate to result in positive changes in laps performed on the PACER test. In fact, 19 of 38 (50%) performed more laps in the post-test; 12 of 38 (31.6%) performed fewer laps in the post-test; and 7 of 38 (18.4%) performed the same number of laps in the post-test and pre-test.

At the completion of the training program field-testers analyzed the PACER test item and formulated comments. They emphasized that physical assistance was important in helping subjects and that students liked holding the hands of teachers during the run. They also recommended running a small number of students during testing to reduce distractions during the run.

Heart Rate Changes

The final part of the study was designed to compare heart rates of subjects prior to and following the daily PACER training program. This was done to gain insight on the intensity of the participants' activity as reflected by heart rate.

Heart rate, using a heart monitor, was taken directly prior to each training session while the individual was seated. Immediately following training (within 5 seconds), heart rate was again taken. Results of heart rate changes are summarized in Table 5.

Table 5

Summary of Heart Rate Changes Prior to and Following the PACER Training Program

Subject No.	Gender	Age	No. of Observations	Pre-run average Heart Rate	Post-Run average Heart Rate	Post-test PACER Laps Completed
1	Female	15	5	82	130	7
2	Female	15	19	89	129	13
3	Female	17	6	80	139	7
4	Female	11	17	74	114	7
5	Male	14	6	105	136	9
6	Male	16	11	107	145	10
		Total	**64**	**89**	**132**	**9**

The average post-exercise heart rate of 132 beats per minute suggests that subjects were working below maximal heart rate. Since the number of laps performed during training sessions was not counted, the amount of external work as measured by laps performed is not available. However, it appears that the cardiorespiratory demand of this activity was well below maximal effort.

Summary of Results and Discussion

In regard to the curl-up test item, subjects 1) had considerable difficulty in learning to perform the task correctly, 2) were negatively affected by the recommended cadence, 3) exhibited difficulty in adjusting to the lesser momentum which is generated by the curl-up rather than a traditional sit-up, and 4) had much less difficulty performing the task with the feet held down. It may be necessary to take more time for students to learn the task, provide greater flexibility in the cadence, or even allow students to select the sit-up test procedure they have learned and adjust standards for evaluation based on these methods and procedures.

Subjects in this study were able to learn the PACER test. However, the performance of subjects was well below Prudential FITNESSGRAM standards. As a group the sample increased one lap from pre-test to post-test. A total of 50% of the subjects performed more laps in the post-test. It is hypothesized that subjects would perform for longer time periods if the distance used in the test was reduced. However, motivation is evidently also important since post-run recovery heart rates reflected effort considerably below that considered maximal. On the other hand, many subjects appeared to be working more anaerobically than aerobically. In many cases, they appeared to be moving at full effort with the longest strides they could muster. If they were performing anaerobically, it is obvious that they would not be able to repeat many laps.

The results of this study support the need for much additional work in regard to the frequency, duration, and intensity of a PACER or a paced run with this population.

October 16, 1997
a:study2

STUDY # 3
SCHOOL OF THE HOLY CHILDHOOD
ROCHESTER, NEW YORK
SPRING, 1995

Purpose and Overview

During the spring of 1995, a follow-up of the study conducted in the fall of 1994 was conducted at The School of the Holy Childhood, Rochester, NY. This study conducted to continue field-testing the curl-up test item used in the Prudential FITNESSGRAM test, and the Progressive Aerobic Cardiovascular Endurance Run (PACER) test item in the Prudential FITNESSGRAM test. The curl-up component of the study began on February 12 and ended on May 31, 1995.

Subjects for the curl-up test included 33 adolescents with mental retardation and mild limitations in physical fitness. Subjects for the PACER test included 29 subjects who were also in the curl-up test. Subjects for the curl-up included 14 males and 19 females between the ages of 10 and 17, whereas, the subjects taking the PACER included 11 males and 18 females. The study was conducted during the physical education classes of the subjects. A total of 5 different classes were involved in the study. The number of sessions that students attempted to learn and/or performed the curl-up varied from class to class (see Table 1).

Table 1
Number of Days of Physical Education Classes

CLASSES	NO. OF DAYS	MALES	FEMALES
2nd period	23	2	1
3rd period	24	3	6
4th period	19	2	1
7th period	20	5	4
8th period	19	2	7
	Total	14	19

The field tests were conducted by two physical education teachers of the subjects in the study under the overall direction of Project Target staff. The Prudential FITNESSGRAM methods and procedures where employed in administering the curl-up and PACER test items. On the PACER, subjects ran 16 meters rather than the 20 meters used in the Prudential FITNESSGRAM. During the PACER , subjects were permitted to run with a partner. The physical education teacher and project staff were the only partners used.

Methods, Procedures, and Results

Curl-up Test

In regard to the curl-up test item, field testers were asked to teach students to perform the test item correctly at the beginning of the study. Once students were able to perform the test item appropriately, field-testers were requested to test subjects to determine the number they were able to perform. Field-testers determined that 17 of the 33 subjects were able to correctly perform the curl-up test the first time it was administered to them.

Subjects unable to initially perform the curl-up correctly were provided opportunity to learn how to perform it in their physical education classes from February 12 to May 31. The number of classes available to learn the curl-up varied with each subject. Depending on the class, 19 to 23 days of physical education were available. From February 12, the starting date for the curl-up test, to May 31, the ending date for instruction and testing, only three more subjects learned how to perform the test item.

Nineteen subjects were available for both pre- and post-testing. One female and one male met the minimal healthy fitness zone standard identified for the Prudential FITNESSGRAM. Test results for the curl-up test appear in Table 2. At the conclusion of the study 20 of 33 subjects could properly perform the curl-up.

Table 2
Test Results on the Curl-up Test for Subjects Learning the Test Item (N=19)

SUBJECTS	N	POST-TEST (FEET HELD)	POST-TEST (FEET NOT HELD)	PERCENT REACHING FITNESSGRAM STANDARDS
Males	7	16.9	9.7	14
Females	12	11.6	6.2	08
Combined	19	13.5	7.5	11

Based upon the observations of the Project Target staff and field-testers, the following additional comments appear warranted:

1) The recommended cadence to perform one curl-up every three seconds appeared to be a problem performing faster and others would have enhanced performance if they would have been permitted to perform more slowly. It may be advisable to permit greater flexibility in regard to cadence.

2) The learning of the curl-up was inhibited by the fact that this sample had learned to perform the sit-up using clearly different procedures. This appeared to result in negative transfer of learning.

3) The Prudential FITNESSGRAM test procedure suggests that a helper place the hands under the head of subjects as they return to the start position and that another tester stand facing the subjects during test administration of the curl-up. Field-testers concluded that holding the hand behind the head may have distracted the participating subjects.

4) Subjects in this sample had considerable difficulty in inhibiting the use of the arms over their head to give them momentum in performing a curl-up.

5) It was difficult to teach students not to come to a full sit-up position in the upward movement in the curl-up because of their previous experiences in performing the sit-up.

6) It should be noted that too much motivation from the field-testers distracts subjects rather than motivates them. An appropriate level of motivation during the curl-up test is crucial.

PACER Test

In the fall of 1994, 39 adolescents with mental retardation and mild limitations in physical fitness at the School of the Holy Childhood were tested to determine the number of laps they could perform in the PACER test. The average of laps completed by the sample was 8 and the average distance run was 160m. In view of the fact that the PACER was established as an indicator of aerobic capacity and the results reflected relatively brief involvement on the part of subjects, it was decided to further investigate the PACER test in the spring of 1995. Specifically, it was decided to study the effects of decreasing the distance of the test from 20m to 16m. The hope was that subjects would perform for a longer total distance thereby increasing their time in aerobic activity. The time for lap completion originally used by the Prudential FITNESSGRAM was maintained.

In the spring of 1995, 25 of the subjects originally involved in the fall 1994 study took part in the study. Subjects were scheduled to be tested in 4 trials at two-week intervals beginning approximately March 28, 1995. Results appear in Table 3. The number of subjects completing the trials varied due to illness, behavioral problems, medication, etc. As presented in Table 4, a total of 21 subjects completed the 20m test in the fall and a 16m trial in May of 1995. Results indicate that the total laps completed increased from 6.8 to 13.0 and total distance run increased from 135.2 to 207.2m. Sixteen of 21 subjects increased the distance completed using the 16m distance in May.

Table 3
PACER Results in Fall 1994 and Spring 1995

			Fall 1994		Spring 1995							
					3/28/95		4/11/95		4/25/95		5/16-5/18/95	
Subj No.	Age	Gender	No. 20m Laps	Distance	No. 16m Laps TRIAL #1	Distance	No. 16m Laps TRIAL #2	Distance	No. 16m Laps TRIAL #3	Distance	No. 16m Laps TRIAL #4	Distance
1	17	M	7	140	9	144	Absent		Behavior		10	160
2	12	F	3	60	2	32	Absent		6	96	8	128
3	11	M	3	60	3	48	Sick		3	48	5	80
4	10	M	Too young at start		4	64	6	96	8	128	6	96
5	10	M	Too young at start		7	112	7	112	Sick		8	128
6	10	F	Too young at start		7	112	18	288	10	160	12	192
7	11	M	9	180	15	240	Sick		19	304	23	368
8	13	M	2	40	5	80	3	48			5	80
9	12	M	Behavior		2	32	6	96	3	48	5	80
10	14	F	3	60	Absent		5	80	3	48	4	64
11		F	Not in test		17	272	17	272	17	272	18	288
12		F	Not in test		1	16	Conflict		2	32	1	16

			Fall 1994		Spring 1995							
					3/28/95		4/11/95		4/25/95		5/16-5/18/95	
Subj No.	Age	Gender	No. 20m Laps	Distance	No. 16m Laps TRIAL #1	Distance	No. 16m Laps TRIAL #2	Distance	No. 16m Laps TRIAL #3	Distance	No. 16m Laps TRIAL #4	Distance
13	13	F	9	180			9	144	9	144	12	192
14	15	F	9	180	10	160	16	256	17	272	On allergy medicine	
15	13	F	3	60	3	48	3	48	Too distracted		1	16
16	11	M	12	240	24	384	33	528	42	672	43	688
17	10	F	7	140	5	80	8	128	7	112	8	128
18	13	F	4	80	1	16	5	80	Too distracted		7	112
19	13	M	9	180	17	272	17	272	19	304	25	400
20	13	F	0	0	3	48	4	64	Would not run		5	80
21	12	F	10	200	18	288	27	432	Sick		13	208
22	14	F	13	260	13	208	23	368	17	272	23	368
23	16	M	10	200	9	144	12	192	12	192	10	160
24	16	F	7	140	8	128	9	144	8	128	10	160
25	15	F	3	60	7	112	7	112	6	96	7	112
26	11	M	7	140	13	208	20	320	26	416	24	384

			Fall 1994		Spring 1995							
					3/28/95		4/11/95		4/25/95		5/16-5/18/95	
Subj No.	Age	Gender	No. 20m Laps	Distance	No. 16m Laps TRIAL #1	Distance	No. 16m Laps TRIAL #2	Distance	No. 16m Laps TRIAL #3	Distance	No. 16m Laps TRIAL #4	Distance
27	17	F	14	280	16	256	20	320	15	240	22	352
28	12	F	7	140	8	128	9	144	8	128	7	112

Table 4
Comparison of 20m and 16m PACER Runs

SUBJECT	FALL 1994		MAY 1995	
	NO. 20m LAPS	DISTANCE (m)	NO. 16m LAPS	DISTANCE (m)
1	7	140	10	160
2	3	60	8	128
3	3	60	5	80
4	9	180	23	368
5	2	40	5	80
6	3	60	4	64
7	9	180	12	192
8	3	60	1	16
9	12	240	43	688
10	7	140	8	128
11	4	80	7	112
12	9	180	25	400
13	0	0	5	80
14	10	200	13	208
15	13	260	23	368
16	10	200	10	160
17	7	140	10	160
18	3	60	7	112
19	7	140	24	384
20	14	280	22	352
21	7	140	7	112
AVERAGE	6.8	135.2	13.0	207.2

Table 5
Comparison of Late April and May Trials on the 16m PACER Run - Spring 1995

SUBJECT	LATE APRIL 1995		MAY 1995	
	NO. 16m LAPS	DISTANCE	NO. 16m LAPS	DISTANCE
1	6	96	8	128
2	3	48	5	80
3	8	128	6	96
4	10	160	12	192
5	19	304	23	368
6	3	48	5	80
7	3	48	4	64
8	17	272	18	288
9	2	32	1	16
10	9	144	12	192
11	42	672	43	688
12	7	112	8	128
13	19	304	25	400
14	17	272	23	368
15	12	192	10	160
16	8	128	10	160
17	6	96	7	112
18	26	416	24	384
19	15	240	22	352
20	8	128	7	112
AVERAGE	12.0	192.0	13.7	218.4

Reliability is also an important factor in recommending a particular test item for a particular population. To provide some insight on the reliability of the PACER, an alpha coefficient was computed using data collected during date April and May on the 16m PACER. Based on the results from 20 subjects tested during these data collection periods, an alpha correlation coefficient of .98 was calculated. Proportion (P) and Kappa (k) coefficients of 1.0 were found using 16 subjects for which data were available. Fifteen of the 20 subjects increased the total distance run in May 1995 from fall 1994.

After the completion of the project, field-testers analyzed the PACER test item and formulated comments. They emphasized that physical assistance was important in helping and motivating subjects and that the students liked holding the hands of teachers during the run. In fact, a variety of other motivational strategies were considered important. The more motivated the subjects were during the PACER, the more laps they completed. It was also found that running a small number of students during testing tended to reduce distractions.

Heart Rate Changes

The final part of the study was designed to record and analyze heart rates of selected subjects prior to and following the PACER test. This was done to gain insight on the intensity of the participant's activity as reflected by heart rate.

Heart rate was taken manually by field-testers immediately prior to each PACER test for a 10-second interval while the subject was standing. Immediately following the test (within 5 seconds), post-heart rate was taken for a ten-second interval.

Results of heart rate changes are summarized in Table 6. The average post-run heart rate ranged from 136 to 155. Individuals running fewer laps had lower post-exercise heart rates. Individuals averaging 143 to 155 beats per minute completed 17 to 25 laps and exercised very close to 70% maximum predicted heart rate. Although observations are relatively few, data suggest submaximal effort on behalf of the subjects.

Table 6
Summary of Heart Rate Changes Prior To and Following 16m PACER Test

SUBJ N0.	GENDER	AGE	NO. OF OBSERVATIONS	PRE-RUN AVERAGE HEART RATE (STANDING)	POST-RUN AVERAGE HEART RATE (STANDING)	70% OF PREDICTED MAXIMUM HEART RATE	LAPS COMPLETED DURING MAY TRIAL
1	Female	15	3	80	143	144	17
2	Female	10	4	86	136	147	8
3	Male	10	3	82	136	147	8
4	Male	13	2	99	145	145	25
5	Male	11	4	78	155	146	24

Discussion and Recommendations

The results of the study suggest that the curl-up can be learned and performed by subjects with mental retardation and mild limitations in physical fitness. Approximately 50% of this sample were able to perform the test when it was initially explained. Those not performing it correctly at the start had difficulty learning it. Only about 10% were able to reach minimal Prudential FITNESSGRAM standards for the curl-up at the end of the study. There is support for using the test item in a physical fitness test for this population. However, performance may be enhanced by greater flexibility in cadence.

The data support the contention that the sample would run a longer total distance if the length of each lap would be reduced from 20m to 16m. Caution must be used in interpreting results because factors such as the effects of growing older, number of trials, familiarity with the test, etc. were not controlled. On the other hand, more subjects appeared to run aerobically and the reduced intensity did appear to result in longer total distance run. Shortening distance to less than 16m is not recommended because of the anticipated confusion resulting from more frequent starts and stops. It may be beneficial to study the effects of decreasing the time taken for laps. For example, running 10-second laps for 2 minutes, increasing to 9-second laps for 2 minutes, etc. might be worthy of further study.

file: study3
disk:pt_3
9-9-97

New York City Public Schools
Spring, 1995

Purpose and Overview

During the spring of 1995, various schools under the administration of the New York City Board of Education participated in a field-test of tentative Project Target health-related physical fitness test items. More specifically, data were collected by physical education teachers who worked with adolescents with either visual impairments or with mental retardation and mild limitations in physical fitness. The purpose of the field-test was a) to seek reactions from testers regarding test procedures, b) to obtain test scores which could be analyzed by gender, age, and disability classification, and c) to determine FITNESSGRAM pass/fail rates (where appropriate) for adolescents with selected disabilities.

Methods and Procedures

In January 1995 the project coordinator trained nine physical education teachers who ultimately tested subjects and submitted the test results to Brockport. A total of 72 subjects with mental retardation and mild limitations in physical fitness (MR) and 51 subjects with visual impairments (VI) were tested as part of this study. Test items administered varied somewhat as a function of disability classification. The following FITNESSGRAM items, however, were given to both groups of subjects: height, weight, PACER, curl-ups, trunk lift, back saver sit and reach, and shoulder stretch. In addition, the following items were added for MR subjects: 600-yard run, isometric push-up, bench press, extended arm hang, and grip strength. The following items were added to the first list for VI subjects: mile run, push-ups, pull-ups, and flexed arm hang. Subjects with visual impairments also were asked to complete a short questionnaire regarding the frequency and intensity of their physical activity. All items were administered in accord with Project Target procedures which were operative at the time of the testing.

As part of the training session, testers were encouraged to spend sufficient time teaching their students the test items before actually collecting data on the items. Curricular pressures and varying teaching schedules prevented "standardizing" the actual learning/practice time for each student, but it is reasonable to assume that each subject was familiar with each test item prior to administration.

Results and Recommendations

This section of the report is sub-divided by disability classification. Results and recommendations for subjects with mental retardation is presented first and results and recommendations for those with visual impairment follow.

Mental Retardation

Mean values, with corresponding numbers of subjects who took a test, were determined for each test item. These values, broken down by gender and age, are presented in Tables 1 and 2.

Table 1. Mean scores (and number of subjects) obtained by girls with mental retardation on Target test items.

Test Item	Age						
	12	13	14	15	16	17	T
PACER (laps) n=	20.3 (3)	8.8 (4)	9.0 (3)	--	--	--	12.3 (10)
600-yard (s) n=	266 (3)	334 (5)	512 (3)	422 (2)	--	--	373 (13)
Curl-up (#) n=	9.2 (5)	8.0 (6)	11.0 (8)	8.8 (6)	--	--	9.4 (25)
Isometric P.U. (s) n=	14.8 (5)	0.5 (2)	7.0 (3)	--	--	--	9.6 (10)
Bench Press (#) n=	--	3.3 (3)	3.0 (3)	1.5 (2)	--	--	2.8 (8)
Extended A.H. (s) n=	6.7 (3)	6.8 (4)	8.7 (6)	22.5 (2)	--	--	9.6 (15)
Back Saver-R (cm) n=	19.3 (6)	21.2 (6)	20.1 (8)	21.5 (6)	--	--	20.5 (26)
Back Saver-L (cm) n=	20.8 (6)	21.0 (6)	19.1 (8)	18.7 (6)	--	--	19.8 (26)
BMI (ht/wt^2) n=	24.2 (6)	22.5 (6)	20.4 (8)	26.4 (6)	--	--	23.2 (26)
Right Grip (kg) n=	12.2 (6)	10.2 (6)	15.5 (8)	26.0 (6)	--	--	15.9 (26)
Left Grip (kg) n=	12.7 (6)	10.6 (6)	13.9 (8)	24.2 (6)	--	--	15.2 (26)

Table 2. Mean scores (and number of subjects) obtained by boys with mental retardation on Target test items.

Test Item	Age						
	12	**13**	**14**	**15**	**16**	**17**	**T**
PACER (laps)	15.3	13.7	11.6	10.8	--	--	12.7
n=	(4)	(6)	(7)	(4)			(21)
600-yard (s)	272	259	375	358	320	--	323
n=	(4)	(7)	(8)	(7)	(1)		(27)
Curl-up (#)	10.8	16.6	17.8	17.3	15.0	1.0	16.0
n=	(6)	(8)	(13)	(15)	(1)	(1)	(44)
Isometric P.U. (s)	29.6	27.0	5.5	15.8	--	1.0	19.3
n=	(7)	(2)	(4)	(4)		(1)	(18)
Bench Press (#)	--	6.8	13.9	10.8	13.0	--	11.4
n=		(4)	(8)	(6)	(1)		(19)
Extended A.H. (s)	30.8	13.8	21.9	14.1	19.0	--	18.7
n=	(4)	(8)	(12)	(10)	(1)		(35)
Back Saver-R (cm)	19.1	22.4	22.5	22.0	24.0	13.0	21.6
n=	(7)	(9)	(13)	(15)	(1)	(1)	(46)
Back Saver-L (cm)	18.7	22.7	21.5	21.2	27.0	13.0	21.2
n=	(7)	(9)	(13)	(15)	(1)	(1)	(46)
BMI (ht/wt^2)	25.0	20.8	23.4	23.4	20.2	18.1	23.0
n=	(7)	(9)	(13)	(15)	(1)	(1)	(46)
Right Grip (kg)	20.7	19.9	18.7	28.4	21.0	5.3	22.2
n=	(7)	(8)	(13)	(15)	(1)	(1)	(45)
Left Grip (kg)	19.8	18.3	21.7	26.2	25.7	9.3	22.0
n=	(7)	(9)	(13)	(15)	(1)	(1)	(46)

In addition to providing information on test performance, Tables 1 and 2 demonstrate that most of the 72 subjects with mental retardation were male (n=46). It is also apparent that although the age range for Project Target is 10-17, most of the subjects in this study were in the 12-15 range and, in fact, no subjects in the 10-11 range were tested.

The sample of subjects with mental retardation included youngsters both with and without Down syndrome. A breakdown of means based upon this categorization is presented in Table 3.

Table 3. Mean scores by gender and presence of Down Syndrome.

	Down Males	Non-Down Males	Down Females	Non-Down Females
Age	14.0	13.9	13.0	13.8
Total N	7	38	7	18
PACER (laps)	7.3	13.6	9.5	9.0
600-yard (s)	408	304	296	396
Curl-ups (#)	6.4	18.0	7.3	10.2
Isometic P.U. (s)	2.7	22.6	12.6	6.6
Bench Press (#)	5.3	11.1	5.0	2.0
Extended A.H. (s)	7.8	21.1	7.3	10.9
Back Saver-R (cm)	20.0	22.1	23.3	18.7
Back Saver-L (cm)	18.6	21.7	24.4	17.3
BMI (ht/wt^2)	28.9	21.6	24.5	22.6
Right Grip (kg)	16.0	23.4	12.4	17.6
Left Grip (kg)	19.1	22.7	11.4	16.9

Since only 14 of the subjects had Down syndrome, the means presented in Table 3 are categorized only by the presence of Downs and gender. Still, it is possible to get some sense of the effect of Down syndrome on test performance. In general it appears that males with Downs did not do as well on the test items as did their non-Downs counterparts. For females, however, the differences are not as great and, in fact, Downs females actually did better than non-Downs females on a few items.

In addition to determining mean scores for the test items under study, a correlation matrix which included all the test items as well as the categorical variables was generated to determine the relationships among these variables. Due to space limitations, a somewhat truncated version of this matrix appears in Table 4.

Table 4: Intercorrelations for Selected Test Items for Subjects with Mental Retardation

	Downs	Gender	Age	PACE R	600-yd	Curl	Iso. P.U.	Bench	Ext. A.H.	R. Grip	T Lift	BS-R	Shld-R	BMI
Downs	1.00	-.15	-.14	-.25	.11	-.30*	-.21	-.21	-.20	-.30*	.07	.03	.00	.05
Gender	-.15	1.00	.16	.02	-.20	.28*	.22	.51*	.21	.31*	.12	.06	.11	-.16
Age	-.14	.16	1.00	-.27	.36*	.09	-.28	.18	-.01	.32*	-.04	.05	-.16	.01
PACER	-.25	.02	-.27	1.00	-.70*	.50*	.69*	.49	.26	.37*	.39*	.31*	.27	-.26
600-yd.	.11	-.20	.36*	-.70*	1.00	-.25	-.40	-.49*	-.30*	-.19	-.38*	.01	-.13	.49*
Curl	-.30*	.28*	.09	.50*	-.25	1.00	.87*	.36*	.61*	.43*	.42*	-.02	.36*	-.21*
Iso. P.U.	-.21	.22	-.28	.69*	-.40	.87*	1.00	--	.79*	.59*	.57*	.25	.49*	-.17
Bench	-.21	.51*	.18	.49	-.49*	.36*	--	1.00	.22	.46*	.22	-.04	-.19	-.19
Ext. A.H.	-.20	.21	-.01	.26	-.30*	.61*	.79*	.22	1.00	.41*	.34*	-.10	.37*	.26*
R. Grip	-.30*	.31*	.32*	.37*	-.19	.43*	.59*	.46*	.41*	1.00	.42*	.03	.09	-.14
T. Lift	.07	.12	-.04	.39*	-.38*	.42*	.57*	.22	.34*	.42*	1.00	.28*	.41*	-.02
BS-R	.03	.06	.05	.31*	.01	-.02	.25	-.04	-.10	.03	.28*	1.00	.13	.14
Shld-R	.00	.11	-.16	.27	-.13	.36*	.49*	-.19	.37*	.09	.41*	.13	1.00	.04
BMI	.33*	-.02	-.03	-.26	.14	-.30*	-.17	.20	-.38*	.21*	.05	-.13	-.32*	1.00

* $p < .05$

The correlation coefficients in Table 4 are indicative of a number of moderate (or better) relationships between test items. The isometric push-up, in particular, has some very interesting coefficients that cut across aerobic fitness (.69 with PACER), muscular strength and endurance (.87 with curl-ups), and flexibility (.49 with shoulder stretch).

Eight of the test items taken by youngsters with mental retardation in this study were FITNESSGRAM test items. Pass/fail information relative to FITNESSGRAM standards is given in Table 5.

Table 5. Pass/fail rates on FITNESSGRAM test items for subjects with visual impairments.

Test Items	Boys		
	# Passing	# Failing	% Passing
BMI	24	22*	52
PACER	1	20	5
Curl-ups	12	32	27
Trunk Lift	21	23	48
Back Saver - R	25	21	54
Back Saver - L	26	20	57
Shoulder Stretch - R	21	25	46
Shoulder Stretch - L	17	29	37

Test Items	Girls		
	# Passing	# Failing	% Passing
BMI	15	11**	58
PACER	1	9	10
Curl-ups	4	21	16
Trunk Lift	9	16	36
Back Saver - R	7	19	27
Back Saver - L	7	19	27
Shoulder Stretch - R	9	17	35
Shoulder Stretch - L	8	18	31

* 16 failed on the "high side"
** 8 failed on the "high side"

The pass/fail data presented in Table 5 seem to suggest that as a group, youngsters with mental retardation will find the FITNESSGRAM standards very challenging, although the boys in this study were more successful than the girls in meeting the various criteria. Over 45% of the boys in this study were able to meet the standards on five of the eight items. Girls were able to achieve this criterion on only one of the eight items.

Item Analysis

Each of the test items administered to mentally retarded subjects as part of this study is discussed in this section. Comments focus on the relationships between gender, age, and presence of Down syndrome for each test item; pass/fail rates based on FITNESSGRAM standards or a discussion of the range of scores in the absence of such standards; and reactions from testers as appropriate.

PACER. Both boys and girls averaged about 12 laps for this test item. There was no apparent relationship (r=.02) between gender and PACER performance. There was, however, a low, inverse relationship (r=-.27) between age and performance indicating a slight tendency for older subjects to do worse on the test. Boys with Down syndrome did worse than their non-Downs counterparts, but girls did about the same. Overall, the relationship between Downs and PACER was low and inverse (r=-.25). Only 5% of the boys and 10% of the girls were able to achieve the FITNESSGRAM standards. Two testers noted that the concept of the PACER was difficult to teach to their students and a third teacher remarked that additional training time was probably necessary for mentally retarded students.

600-yard run. Overall, boys averaged about 5 minutes and 20 seconds for the run and girls took almost a minute longer. The correlation between gender and the 600 was -.20. As with the PACER, performance on the 600 tended to decline with age (r=.36). In fact of all the items in the study, the 600 had the strongest relationship with age, albeit inverse. The presence of Down syndrome resulted in somewhat of a "mixed bag." Boys with Downs generally did worse, and girls with Downs generally did better, than subjects without Downs. Overall, the relationship between Downs and the 600 was very low and inverse (r=.11). Mean scores for the various gender and age combinations generally were found to be below the 5th percentile when compared to the norms of the AAHPERD Youth Fitness Test. Testers indicated that they considered the 600-yard run to be appropriate for their students.

Curl-ups. Boys generally did better than girls (r=.28) on this item. Boys averaged about 16 curl-ups while girls were able to do a little more than 9. There was virtually no relationship between age and curl-up performance (r=.09). Subjects without Downs syndrome tended to do better than those with Down syndrome on this test (r=-.30). Sixteen percent of the girls and 27% of the boys made test scores above the minimum standards associated with FITNESSGRAM. Testers were apparently comfortable with the appropriateness of the curl-ups for students with mental retardation.

Isometric push-up. Boys tended to do better than girls (r=.22) on the isometric push-up. Girls averaged just under 10 seconds and boys about 19 seconds on this test. There was a tendency for performance to decline with age (r=-.28). The effect of Downs syndrome on isometric push-up performance was a function of gender; girls with Downs did better, while boys with Downs did much worse, than subjects without Downs syndrome. Overall, the relationship was low and inverse (r=-.21). In the absence of any published standards for this item it seems appropriate to note that the range of scores seems to be reasonable for the purpose of establishing standards later on. Almost all students were able to score above 0 thereby placing themselves on an "achievement continuum." Although not presented in this report, the standard deviations associated with this test are quite high. Ordinarily this is not a particularly desirable characteristic for a test, but when attempting to establish standards, the high standard deviations suggest that it should not be very difficult to establish a cut-off score that accurately describes those who "can" and those who "can't" do this test. Testers seemed to like the isometric push-up although one tester noted that this is an item that requires additional instruction and practice for youngsters with mental retardation.

Bench press. Boys did quite a bit better than girls on this test item. In fact of all the items in this study, the bench press yielded the highest Pearson r with gender (r=.51). The effect of age on performance was somewhat difficult to gauge. The overall relationship between age and the bench press was low and positive (r=.18), but female scores certainly did not improve with age. As with some other items already discussed, the effect of Down syndrome was influenced by the gender of the subject; overall the relationship was low and inverse (r=-.21). The range of test scores for boys seems to be reasonable for establishing standards, but scores obtained for girls were generally low, making the process more difficult. If this item is retained, it may be necessary to experiment with a lighter weight for mentally retarded females. Testers did not specifically comment on the appropriateness of the bench press, however, some testers were unable to administer the test (only 27 of the 72 subjects took the test) due to equipment demands. To the extent that equipment is a problem with this test, consideration of the dumbbell press as an alternative seems appropriate.

Extended arm hang. Boys averaged just under 19 seconds for the extended arm hang while girls averaged just under 10 seconds. The relationship between gender and test performance was low, but positive (r=.21). Test scores were not affected by age (r=-.01), but they were affected by the presence of Down syndrome. Subjects with Downs generally did worse than those without Downs (r=-.20). The range of scores looks promising for the establishment of standards and testers seemed to be satisfied with the item although one tester noted the need for additional training. This item appears to have good potential as an alternative test for upper body strength/endurance.

Back saver sit and reach. Interestingly, the sit and reach had virtually no relationship with any of the categorical variables in this study. Neither gender (r=.06), age (r=.05), or Down syndrome (r=.03) affected test score performance. Boys, who averaged about 21 cm, did fairly well when their scores were compared to the FITNESSGRAM standards. Depending on the side measured, 54-57% of the boys "passed" the test. Girls (with an average of about 20 cm), however, were not as successful; regardless of the side measured, only 27% met the standard. None of the testers commented specifically on the sit and reach; apparently they found this item to be appropriate for their students.

Body mass index. Both boys and girls had average BMI scores right around 23. No relationships between BMI and gender (r=-.02) or between BMI and age (r=-.03) were found. Subjects with Down syndrome had higher BMI values than those without Down syndrome, in fact of all the test items in this study BMI had the highest Pearson r with Down syndrome (r=.33). Over 50% of the subjects tested in this study were able to meet the FITNESSGRAM standards for body mass index. Of those who failed, most (73%) failed because they were too heavy for their height. None of the testers commented on the BMI test; in fact, it appears that they did not consider it an "official" test item since none even acknowledged it on their reaction forms. Presumably they felt that they were simply reporting height and weight in much the same way they reported gender and age.

Grip strength. Boys typically did better than girls on the grip strength (r=.31 for right grip and gender). Boys averaged about 22 kg to the girls 15 kg. No relationship between age and grip performance was noted (r=-.03), but grip strength was affected by the presence of Down syndrome. Youngsters with Down syndrome tended to be associated with lower grip strength scores (r=-.30). When comparing the grip strength of subjects in this study with Project UNIQUE norms, it is interesting to separate the Downs and non-Downs subjects. Using the norms for 14-year old nondisabled boys and girls as the reference, both boys and girls without Downs scored slightly above the 20th percentile for sum of grips. Subjects with Downs scored at, or slighlty below, the 10th percentile using the same reference. Testers apparently liked the grip strength test, but the "equipment issue" was raised by at least one tester.

Trunk lift and shoulder stretch. Two of the test items, trunk lift and shoulder stretch, were not scored as interval data consequently no mean data will be reported. Very limited relationships were observed between the test items and gender (.11 to .12), age (-.04 to -.16), and Down syndrome (.00 to .07). A moderate relationship, however, was found between the test items (r=.41). Forty-eight percent of the males in this study were able to meet the FITNESSGRAM standards for the trunk lift compared to 36% of the females. Depending on which arm was used, 37-46% of the males and 31-35% of the females were able to achieve the FITNESSGRAM criteria on the shoulder stretch. Regardless of gender, subjects were more successful using their right arm than their left. Testers apparently found these items to be very "do-able" with their students.

Summary and Recommendations

Based on the results obtained in this study with youngsters with mental retardation, some observations and recommendations seem warranted. Comments are arranged by categories and discussed in the following sections.

Down syndrome. Of the 72 subjects in this study, only 14 had Down syndrome so the recommendations presented here should be considered speculative. It does appear, however, that separate standards may have to be developed for at least three test items: curl-ups, BMI, and grip strength. Although not as compelling as for these three items, there is also evidence that separate norms may be necessary for the PACER, isometric push-up, bench press, and extended arm hang. The performance of other subjects with Down syndrome should be considered before any final decision on separate norms is made.

Age. Test performance either declined or remained stable with increasing age for most of the items in this study. This observation presents a dilemma when establishing standards for this population. For most test items, standards established for non-disabled adolescents get "better" with age and, in some cases, "much better" with age. The results of this study, however, suggest that such an approach might place some standards out of the reach of the older youngsters with MR covered by the project. Standards should probably "improve" with age in order to maintain incentive and to account for changes in growth and development, but perhaps the amount of improvement should be modest when compared to the standards for the non-disabled adolescents.

Gender. Differences in gender followed fairly typical patterns of development for the youngsters with MR with the exception of those items dealing with flexibility. No gender differences in flexibility were noted where the expectation was that females would score better. Consequently, males had higher pass/fail rates than females on these items. Additional research should be reviewed to see if this pattern holds for adolescents with MR for possible implications for standard setting.

Aerobic function. A fairly strong relationship (r=-.70) was noted between the PACER and 600-yard run in this study. Average performance on the PACER lasted under two minutes while 600-yard performance was over five minutes for most of the subjects. Although the external pacing of the PACER is desireable for youngsters with MR, teachers reported that the concept of the test is not easily understandable by many of these youngsters so the benefit of the pacing is lost. Consequently, it appears that the 600 might be the preferred test at this time. A modified PACER, however, could result in a longer exercise period, thereby increasing the aerobic nature of the test, and increased instruction could overcome the "concept" problems with this test. Regardless of test item, it will be necessary to modify the standards for this component of fitness. Perhaps the research at GWU will help in establishing reasonable standards, but it appears that the standards established for youngsters with MR will have to be below the 20th percentile (perhaps around the 5th percentile) established for non-disabled youngsters.

Strength and endurance. The Pearson correlation coefficients among the five strength/endurance items in this test battery are quite high, thus, lending support to the notion of selecting one test from a list of three or four items. Following the FITNESSGRAM lead, it seems reasonable to require the curl-ups of all testees and allow one additional test to be selected from among the isometric push-up, the bench press, extended arm hang, and grip strength. Each of these items seems appropriate for youngsters with MR, but establishing standards will require some additional thought. The curl-up performance of subjects in this study was especially low and will defy the 20th percentile guideline. Furthermore, there is limited data in the literature on either the isometric push-up or extended arm hang, so additional research will be necessary.

Flexibililty. Intercorrelations suggest that both the sit and reach and shoulder stretch should be given to all, rather than having a "select one" option since the relationship between the two is very low (r=.13). The Apley test could be substituted for the shoulder stretch to increase the pass/fail rate of subjects with MR. As suggested earlier, the flexibility of females with MR should be looked at more carefully before establishing sit and reach standards for this group. The trunk lift seems to be a reasonable test item (although the standards will be challenging) for youngsters with MR. Since

this item combines both strength and flexibility elements, it may not be surprising to note that this test generated significant correlation coefficients with eight of the other 10 items (including the PACER and 600).

Body composition. BMI certainly seems like a reasonable item. Intercorrelations with BMI generally were low, although significant correlations were found for the 600, curl-ups, and extended arm hang. Non-Downs youngsters will compare favorably to the FITNESSGRAM standards, but those with Down syndrome will find the standards difficult to achieve.

Visual Impairment

Mean values, with corresponding numbers of subjects who took a test, were determined for each item. These values, broken down by gender and age, are presented in Tables 6 and 7.

Table 6. Mean scores (and number of subjects) obtained by girls with visual impairments on Target test items.

Test Item	Age								
	10	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**T**
Mile (s)	700	760	1287	893	860	768	--	--	938
n=	(1)	(1)	(3)	(3)	(2)	(2)			(12)
PACER (laps)	30.7	10.6	12.0	19.0	13.5	29.0	--	--	18.0
n=	(3)	(5)	(3)	(4)	(2)	(2)			(19)
Curl-up (#)	13.0	14.6	20.0	12.3	16.0	40.0	--	--	17.5
n=	(3)	(5)	(3)	(4)	(2)	(2)			(19)
Push-ups (#)	3.0	3.4	3.7	2.3	3.5	3.5	--	--	3.2
n=	(3)	(5)	(3)	(3)	(2)	(2)			(18)
Pull-ups (#)	0	0	0	0	0	0	--	--	0
n=	(2)	(4)	(3)	(4)	(1)	(1)			(15)
Flexed A.H. (s)	1.0	16.5	0	1.3	0	0	--	--	4.9
n=	(2)	(4)	(3)	(4)	(1)	(1)			(15)
Back Saver-R (cm)	21.5	9.0	33.0	27.3	29.5	26.0	--	--	25.3
n=	(2)	(2)	(3)	(4)	(2)	(2)			(15)
Back Saver-L (cm)	22.0	19.0	33.7	28.8	27.5	28.0	--	--	27.3
n=	(2)	(2)	(3)	(4)	(2)	(2)			(15)
BMI (ht/wt^2)	24.7	18.4	22.6	22.4	21.3	20.1	--	--	21.4
n=	(3)	(5)	(3)	(4)	(2)	(2)			(19)

Table 7. Mean scores (and number of subjects) obtained by boys with visual impairments on Target test items.

Test Item	Age								
	10	11	12	13	14	15	16	17	T
Mile (s)	--	949	926	829	947	763	--	691	885
n=		(3)	(4)	(4)	(5)	(2)		(1)	(19)
PACER (laps)	9.4	29.0	15.5	26.3	22.0	29.0	--	20.0	21.5
n=	(5)	(8)	(6)	(4)	(5)	(2)		(1)	(31)
Curl-up (#)	13.2	30.7	47.3	31.8	14.8	36.5	--	10.0	28.4
n=	(5)	(9)	(6)	(4)	(5)	(2)		(1)	(32)
Push-ups (#)	2.6	9.2	6.7	5.3	3.6	11.0	--	10.0	6.5
n=	(5)	(9)	(6)	(4)	(5)	(2)		(1)	(32)
Pull-ups (#)	0.6	1.4	1.8	0.5	1.3	0	--	--	1.1
n=	(5)	(5)	(6)	(2)	(3)	(1)			(22)
Flexed A.H. (s)	14.4	13.8	4.8	9.0	1.7	0	--	4.0	8.6
n=	(5)	(5)	(6)	(2)	(3)	(1)		(1)	(23)
Back Saver-R (cm)	21.0	21.8	25.7	17.5	22.4	14.0	--	15.0	21.2
n=	(3)	(8)	(6)	(4)	(5)	(2)		(1)	(29)
Back Saver-L (cm)	20.3	23.3	24.2	14.5	23.4	17.5	--	16.0	21.3
n=	(3)	(8)	(6)	(4)	(5)	(2)		(1)	(29)
BMI (ht/wt^2)	22.4	17.3	19.5	21.4	22.4	18.3	--	22.6	20.2
n=	(5)	(8)	(6)	(4)	(5)	(2)		(1)	(31)

Of the 51 visually impaired subjects in this study, 32 were male and 19 female. For the most part, subjects generally were within the 10-15 year age range. Only one 16- or 17-year old subject was tested (a 17-year old boy).

In addition to age and gender, subjects were also classified according to the degree of visual impairment. Three classifications of blindness, according to the criteria established by the USABA, and one of partial sightedness, in accord with criteria established in IDEA, were used in this study. A breakdown of means categorized by gender and degree of visual impairment (with the three categories of blindness combined) is presented in Table 8.

Table 8. Mean scores by gender and visual impairment.

Test Items	Blind Males	Partially Sighted Males	Blind Females	Partially Sighted Females
Age	12.2	12.2	12.0	12.3
Total N	19	13	10	9
Mile (s)	876	893	725	1044
PACER (laps)	16.8	27.9	13.9	22.6
Curl-ups (#)	26.0	32.0	17.7	17.3
Push-ups (#)	6.0	7.2	3.2	3.2
Pull-ups (#)	1.1	1.4	0	0
Flexed Arm Hang (s)	6.3	12.8	6.4	3.1
Back Saver - R (cm)	18.2	25.5	20.9	29.3
Back Saver - L (cm)	19.2	24.3	24.3	29.9
BMI (ht/wt_2)	20.4	19.8	22.2	20.5

An inspection of the means presented in Table 8 suggests that the degree of visual impairment may have influenced the performance of subjects on at least some of the test items. Partially sighted subjects, however, did not always do better than their blind counterparts.

In addition to determining mean scores for the test items under study, a correlation matrix which included all the test items as well as the categorical variables was generated to determine the relationships among these variables. Due to space limitations, a somewhat truncated version of this matrix appears in Table 9.

A number of moderate (and significant) correlation coefficients appear in Table 9, but probably none that would be considered high. One of the interesting (but not necessarily unexpected) findings is that BMI had a significant inverse relationship with six test items, including the mile run, PACER, curl-ups, push-ups, pull-ups, and flexed arm hang.

Twelve of the test items taken by youngsters with visual impairments were FITNESSGRAM test items. Pass/fail information relative to FITNESSGRAM standards is given in Table 10.

Table 9: Intercorrelations for Selected Test Items for Subjects with Visual Impairments.

	Class	Gender	Age	Mile	PACER	Curl	Push	Pull	FAH	T Lift	BS-R	Shld-R	BMI	HPA
Class	1.00	.03	-.13	.27	.20	.15	-.02	.14	.03	.13	.37*	-.04	-.03	.12
Gender	.03	1.00	.01	-.11	.09	.25*	.31*	.53*	.16	.09	-.21	-.25	-.13	.31*
Age	-.13	.01	1.00	-.22	.08	.07	.07	-.03	-.34*	.12	.00	.27*	.05	.27*
Mile	.27	-.11	-.22	1.00	-.30*	-.10	-.22	.02	-.25	-.26	.40*	.04	.31*	-.21
PACER	.20	.09	.08	-.30*	1.00	.24*	.37*	.26	.11	.17	-.04	.06	-.27*	.16
Curl	.15	.25*	.07	-.10	.24*	1.00	.36*	.68*	.03	.44*	.09	.39*	-.34*	.31*
Push	-.02	.31*	.07	-.22	.37*	.36*	1.00	.37*	.27	.42*	-.08	.10	-.29*	.27*
Pull	.14	.53*	-.03	.02	.26	.68*	.37*	1.00	-.01	.24	.19	.26	-.30*	.25
FAH	.03	.16	-.34*	-.25	.11	.03	.27	-.01	1.00	.05	-.07	-.10	-.30*	-.22
T Lift	.13	.09	.12	-.26	.17	.44*	.42*	.24	.05	1.00	.00	.06	-.13	.47*
BS-R	.37*	-.21	.00	.40*	-.04	.09	-.08	.19	-.07	.00	1.00	.20	-.08	.04
Shld-R	-.04	-.25*	.27*	.04	.06	.39	.10	.26	-.10	.06	.20	1.00	-.04	-.05
BMI	-.03	-.13	.05	.31*	-.27*	-.34*	-.29*	-.30*	-.30*	-.13	-.08	-.04	1.00	.02
HPA	.12	.31*	.27*	-.21	.16	.31*	.27*	.25	-.22	.47*	.04	-.05	.02	1.00

* $p < .05$

Table 10. Pass/fail rates on FITNESSGRAM test items for subjects with visual impairments.

Test Items	Boys		
	# Passing	# Failing	% Passing
BMI	20	11*	65
Mile	2	17	11
PACER	5	26	16
Curl-ups	16	16	50
Push-ups	10	22	31
Pull-ups	12	10	55
Flexed A.H.	7	16	30
Trunk Lift	27	4	87
Back Saver-R	14	15	48
Back Saver-L	15	14	52
Shld. Str.-R	15	16	48
Shld. Str.-L	11	20	36

Test Items	Girls		
	# Passing	# Failing	% Passing
BMI	13	6**	68
Mile	2	10	17
PACER	11	8	58
Curl-ups	8	11	42
Push-ups	3	15	17
Pull-ups	0	15	0
Flexed A.H.	3	12	20
Trunk Lift	13	5	74
Back Saver-R	9	6	60
Back Saver-L	10	5	67
Shld. Str.-R	14	5	74
Shld. Str.-L	9	10	47

* 6 failed on the "high side"
** 4 failed on the "high side"

The data in Table 10 suggest that youngsters with visual impairments should be able to attain FITNESSGRAM standards on at least some of the items. At least 45% of both boys and girls in this study were able to achieve the standards on seven of the 12 items. Girls actually had more success; at least 60% of the girls were able to pass five of the 12 items.

Item Analysis

Each of the tests administered to visually impaired subjects is discussed in this section. Comments focus on the relationships between gender, age, and degree of visual impairment for each test item; pass/fail rates based on FITNESSGRAM standards; and reactions from testers where appropriate.

Mile run. On the average, boys (14:45) ran the mile slightly faster than the girls (15:37); the relationship between gender and performance, however, was low (r=-.11). Some improvement with age is apparent based on a visual inspection of the means (r=-.22). The relationship between class (i.e., degree of visual impairment) and mile run performance (r=.27) is surprising. The relationship indicates a tendency for those with more serious visual impairments to do better on the mile run than those with less serious visual impairments. FITNESSGRAM standards for the mile will be difficult to achieve by VI youngsters based on the results of this study. Only 11% of the males and 17% of the females were able to achieve the criterion score. Two of the testers indicated that they felt that the test was too long for students they tested.

PACER. Results indicate that boys did a little better than girls on the PACER. The relationship between gender and performance was very low (r=.09), but that relationship may have been clouded by the exceptional performance of one 10-year old girl. Age did not seem to be a significant factor (r=.08) in PACER performance (although the score of the 10-year old may have reduced this coefficient). Unlike the mile run, higher classes (i.e., less visually impaired) tended to do better on the PACER than those from the lower classes (r=.20). FITNESSGRAM pass/fail rates were a bit curious. Fifty-eight percent of the girls in the study were able to meet the standards, but this could be accomplished by only 16% of the boys. None of the testers commented specifically on the feasibility of the PACER; presumably they felt that it was an appropriate item for their students.

Curl-ups. Boys did quite a bit better than the girls on the curl-ups (r=.25). Performance, however, did not change appreciably with age (r=.07). Subjects from the higher classes tended to do a bit better than those from the lower classes, but the relationship was still low (r=.15). The pass/fail rates for FITNESSGRAM standards are somewhat encouraging; 50% of the boys and 42% of the girls met or exceeded the criterion score. There were no tester comments pertaining to the curl-ups.

Push-ups. Boys, again, seemed to outperform their female counterparts. In fact, the relationship between gender and performance was on the low side of moderate (r=.31). The range of scores made by girls on push-ups was very restricted. One teacher commented that he thought that the test was too difficult for VI girls. Scores appeared to change only slightly with increasing age (r=.07) and class was no factor at all (r=-.02) in push-up performance. Thirty-one percent of the

boys and 17% of the girls achieved the FITNESSGRAM standards.

Pull-ups. The relationship between gender and test performance was higher for the pull-ups (r=.53) than for any other item. This was apparently due to the fact that no female was able to do a single pull-up. As with the push-ups, age did not seem to have an impact (r=-.03) on pull-up performance. The relationship between performance and class was low (r=.14) indicating that those from the higher classes tended to do a little better than those from the lower. Since none of the girls could do a pull-up, none were able to meet the FITNESSGRAM standards. Fifty-five percent of the boys, however, could meet the FITNESSGRAM criterion.

Flexed arm hang. On the average, boys tended to do better than girls on flexed arm hang (r=.16). Interestingly, age had an inverse relationship with performance (r=-.34); older subjects tended to make lower scores than younger subjects. In fact, the relationship between age and flexed arm hang performance is stronger than exists between age and any other item in the battery. Degree of visual impairment was not related to performance in any way (r=.03). Thirty percent of the boys and 20% of the girls met the FITNESSGRAM standards. One teacher remarked that the flexed arm hang was difficult to teach to VI youngsters and difficult for them to do.

Back saver sit and reach. As would be expected, girls tended to score better than boys on the sit and reach (r=-.21). Age was no factor in sit and reach performance (r=.00), but class certainly was. Subjects from higher classes tended to do better (r=.37) and this relationship was higher than the relationship between class and any of the other items. The pass/fail rates for FITNESSGRAM are encouraging. Depending on the leg tested, 48-52% of the boys met the standards and 60-67% of the girls "passed" the test.

Trunk lift and shoulder stretch. Neither gender (r=.09), age (r=.12), or class (r=.13) had a significant effect on trunk lift performance for subjects in this study. The pass/fail rates for FITNESSGRAM were quite high; 87% of the boys and 74% of the girls met the criteria for the trunk lift. Shoulder stretch performance, while being unrelated to class (r=-.04), was influenced by the gender (r=-.25) and age (r=.27) of the subject. Girls tended to do better than boys and older subjects tended to do better than younger subjects on this test. FITNESSGRAM-related success was not as great for the shoulder stretch as it was for the trunk lift. Depending on the arm tested, 36-48% of the boys and 47-74% of the girls met the standards. The rather large pass/fail ranges (especially for the girls) is due to the fact that, for whatever reason, the right arm was "easier" than the left for these subjects.

Body mass index. Girls tended to have slightly larger scores than boys (r=-.13) on BMI. Neither age (r=.05) or class (r=-.03) were related to BMI scores. Sixty-five percent of the boys and 68% of the girls were able to achieve the FITNESSGRAM standards.

Habitual physical activity. VI subjects completed a physical activity questionnaire to estimate the extent of the intensity and frequency of their exercise habits. In correlating an habitual physical activity score with other variables in the study, significant coefficients were found for gender (r=.31), age (r=.27), curl-ups (r=.31), push-ups (r=.27), and trunk lift (r=.47).

Summary and Recommendations

Findings for VI subjects tested in New York City are summarized in the following sections. Recommendations based on these summary statements follow. Results are discussed both in terms of categorical variables (class, age, and gender) and fitness components (aerobic funtioning, strength and endurance, flexibility, and body composition).

Class. Generally speaking, the relationship between degree of visual impairment and test performance was low. In only three instances did a Pearson r between class and a test item equal or exceed .20 and in only one case (sit and reach) was the coefficient significant. As might be expected, two of the three items that generated a Pearson r of at least .20 were the running items (mile run and PACER). The influence of class, however, had the opposite effect on the performance of these items (the more seriously VI subjects did better on the mile run, but worse on the PACER). The correlational data seem to suggest, therefore, that it is not necessary to provide separate standards for blind and partially sighted subjects. On the other hand, an inspection of the means in Table 8 suggests that such a conclusion might be hasty since some of those differences appear large. Although not presented in the body of this paper, an additional analysis that looked at FITNESSGRAM pass/fail rates as a function of classification was performed. Of the 12 relevant items, partially sighted boys had higher pass rates than blind boys on 10 of the items. The average pass rate for partially sighted boys was 53% compared to an average of 38% for blind boys. The differences were not quite as dramatic for the girls. Partially sighted girls had higher pass rates on six of the items, while blind girls had higher rates on five (there was no difference on one). The average pass rate for partially sighted girls was 50% compared to 40% for blind girls. These results should be considered in concert with the findings of other studies (most notably the Michigan study) before a decision on separate standards is reached.

Age. It is generally accepted, and expected, that fitness test performance of children and adolescents with visual impairments improves with age. In fact, there is some data to indicate that VI youngsters tend to "catch-up" to their sighted peers with increasing age. This was not the case in the NYC study. Age did not have much of an influence on test item performance for these subjects. Although not presented in the body of this report, a separate analysis of the data was performed to determine if FITNESSGRAM pass/fail rates fluctuated significantly as a function of age. Subjects were divided into two age groups, 10-12 and 13-15, as well as by gender, and pass/fail rates were calculated for each group. The most interesting result of this analysis was that pass/fail rates declined with increasing age for both boys and girls on the four tests of muscular strength and endurance (curl-ups, push-ups, pull-ups, and flexed arm hang). (Actually, the pull-up performance of girls remained the same for both age groups, 0%.) The performance of the VI subjects in this study generally did not improve with age (especially on strength and endurance items), consequently older subjects were less successful than younger subjects on FITNESSGRAM standards since those standards tend to increase with age. Taken independently, therefore, these findings suggest that Target standards should reflect only minimal gains with increasing age. As with class, however, other studies should be considered before this decision is finalized.

Gender. The influence of gender on the test performance of VI subjects was fairly typical and predictable. Boys tended to do better than girls on the measures of muscular strength and

endurance and girls tended to do better than boys on measures of flexibility. The differences observed on the measures of aerobic functioning favored the boys, but these differences seem to be less dramatic than those observed for the MS/E items.

Aerobic functioning. The Pearson r between the mile run and the PACER was significant, but only moderate (r=.30). For whatever reason, these two tests were not measuring the same element of fitness. It may be that the high times recorded by subjects on the mile run suggest frequent walking on the part of these participants. If this was the case, it may be that the mile run was not really tapping an aerobic element in these subjects. Regardless, the PACER seems to be the better measure of aerobic functioning for VI younsters based upon better pass/fail rates (especially for girls), shorter administration time for teachers, and easier administrative control in that the mile run could require guide runners for some subjects while the PACER can be conducted with guide ropes instead.

Strength and endurance. The four MS/E items are intuitively appropriate for VI youngsters. Each is done "in place" and none requires vision for successful execution. The push-up test and flexed arm hang test do, however, require an element of body image and some VI subjects have difficulty getting into the correct body position for these items. As mentioned earlier, the pass/fail rates for younger (10-12) subjects were better than those of the older (13-15) subjects for these four items. The average pass rate for younger boys was 53% compared to the 20% average obtained by the older boys. Younger girls averaged a 24% pass rate on the MS/E items compared to 13% for the older girls. The 53% pass rate for younger boys seems to suggest that these items and standards may well be appropriate for VI youngsters and that the lower rates obtained by other groups is more a reflection of poor fitness than inappropriate tests/standards. Still, the pull-up test for females should be looked at more closely before its inclusion is finalized since it failed to place youngsters on an achievement continuum in this study.

Flexibility. The best pass/fail rates for VI subjects were associated with the flexibility items (sit and reach, shoulder stretch, and trunk lift). These items appear to be quite appropriate for youngsters with visual impairments and the standards are clearly within reach.

Body composition. As with the flexibility items, body mass index also appears to be an appropriate test item for VI adolescents. Over 60% of both boys and girls were able to achieve the FITNESSGRAM standards. Of those who failed to stay in the "healthy fitness zone," most were too heavy for their height.

Disk: F.X. Short
File: nyc.sum

Validation of Cardiovascular Fitness Field Tests for Children with Mental Retardation.
A final report for Project Target

Bo Fernhall, The George Washington University and Kenneth Pitetti, Wichita State University.

INTRODUCTION

Background: Cardiovascular fitness is one of the most important components of physical fitness because of its relationship to health and disease. Cardiovascular fitness can be accurately assessed in the laboratory by measuring oxygen uptake during an incremental exercise test. However, because this procedure is costly, time consuming and not widely available, field tests have been developed that can predict aerobic capacity with reasonable accuracy. Among the most common field tests used are timed distance run/walks, submaximal cycle ergometer tests, timed distance swims, and step tests. For children without disabilities, timed distance runs are the most common and the most accurate (valid) tests. The most common distance used for children is the one mile run.

Problem: To date, only the ½ mile walk-run has been validated as a field test for measuring cardiovascular fitness for use with children with mental retardation. However, this test had low-moderate concurrent validity, and a large SEE. The most commonly used test, the 300 yard run, is not a valid test of cardiovascular fitness in adults with mental retardation or in children without disabilities. Thus, there is a need to find a valid field test of cardiovascular fitness for children with mental retardation. However, because of problems with motivation and attention span, it is unlikely that children with mental retardation with mild limitations can as a group perform adequately on the one mile run. Instead, it has been suggested that the 600-yard run/walk or a paced shuttle run be used. The 600-yard run/walk already has norms available for children with mental retardation, and there is a large data base for comparison with children without disabilities. However, it is unknown if this test is a valid indicator of cardiovascular fitness in children with mental retardation with mild limitations.

Objective: To validate the 600-yard run/walk and the 20m PACER for use as tests of cardiovascular fitness in children with mental retardation with mild limitations. Furthermore, a modification of the 20m PACER, with the distance changed to 16 meters, was also validated.

METHODS

Subjects: Thirty-four (34) children with MR with mild limitations in physical fitness (22 boys and 12 girls) between the ages of 10-18 years participated in the study. The subjects were recruited from a summer school program, and both subject and parental informed consent were obtained prior to testing. Individual IQ scores were not released by the school, but all children had to be classified with either mild or moderate mental retardation with mild limitations in physical fitness to be eligible for the summer school program.

Testing Protocol: The laboratory testing for VO2peak followed the procedures described by Fernhall and Tymeson (1987) and Pitetti, Rimmer, Fernhall (1993) for subjects with MR. Briefly, subjects were familiarized with the laboratory setting, treadmill walking, and use of the headgear prior to any data collection. The number of familiarization sessions varied from 2-4 times depending on the subject.

A treadmill protocol was used with the speed individualized to the ability of the subjects. The subjects started at 0% grade, and the grade was increased 4% every 2 minutes until 12% grade was achieved. At this point the speed was increased 0.5 mph until the subjects could no longer continue. These procedures have been shown to produce valid and reliable data in adolescents and adults with MR.

Heart rates were collected each minute during the exercise test using a Polar heart monitor. Subjects breathed into a Hans-Rudolph valve, and expired air was collected and analyzed by a Quinton Q-Plex metabolic system. Metabolic data were calculated and displayed in one minute averages. The metabolic system was calibrated prior to each test, and all tests were conducted by the same testers. The highest VO2 attained was recorded as VO2peak.

The 600-yard run/walk was conducted on an outdoor grass field, marked with cones. A circular path was established and 1.5 laps constituted 600 yards. Subjects received instruction and pacing training several times before an actual test was conducted. To help control for motivation, each subject completed the run-walk with a tester, who was instructed to run-walk about 2-3 feet in front of the subject, and to give verbal encouragement, but never to physically touch the subject during the test. The time of completion was recorded to the nearest second. Each subject completed two tests, separated by 2-5 days. If the time of completion was not within 45 seconds, a third test was conducted. The two tests times within 45 seconds of each other were used in the data analysis.

The 20m and 16m PACER runs were completed on a tennis court surface with the distances marked by a painted stripe on the ground and by cones. The subjects were instructed to run the distance between the cones in the time allotted. Following several practice sessions, subjects completed 2 tests of both the 20 and 16 meter PACER tests. Each test was separated by 2-5 days. Subjects received verbal encouragement and pacing throughout the test in a manner similar to the 600-yard test. The number of laps completed in the allotted time was recorded.

Data Treatment: A test-retest Pearson coefficient was calculated for the 600-yard run/walk completion time and for the number of laps completed on the 20 and 16 meter PACER. The relationships between VO2peak and the field tests were determined by Pearson product-moment correlation. Since age, height, weight and BMI can influence the relationship between run-walk time and VO2peak, a stepwise multiple regression was conducted to evaluate the contribution of these variables, in conjunction with field test performance, to VO2peak. Descriptive characteristics were calculated for all variables.

RESULTS

The descriptive characteristics are presented in Table 1 below. There was no difference between any of the field test trials at the .05 level.

Table 1.

Descriptive characteristics

Variable		**N=34**
Age (years)		14.3 ± 2.34
Height (cm)		157.7 ± 14.0
Weight (kg)		58.4 ± 18.5
BMI		22.9 ± 6.3
VO2peak ($ml \cdot kg^{-1} \cdot Min^{-1}$)		36.6 ± 9.1
Vepeak ($L \cdot min^{-1}$)		78.7 ± 26.7
Rpeak		1.08 ± 0.08
Peak Heart Rate (bpm)		186 ± 10
20m PACER (laps)	T1	15.5 ± 15.8
	T2	15.2 ± 16.0
16m PACER (laps)	T1	23.1 ± 20.3
	T2	24.5 ± 22.5
600-yard run/walk time (min)	T1	4.23 ± 1.1
	T2	4.21 ± 1.2

Alpha test-retest reliability coefficients for the 16m and 20m PACER were of particular interest for the purpose of Project Target. A test-retest a = .96 was reported for the 16m, PACER and an a = .97 score was reported using the 20m PACER. Thirty-four (34) males and females were used for alpha calculations. Pearson r coefficients were 0.97 ($p>0.001$) for the 600-yard run/walk, 0.95 ($p<0.001$) for the 20m PACER and 0.92 ($p<0.001$) for the 16m PACER. The relationship between VO2max, the field tests and some of the descriptive variables are presented in Table 2 below.

Table 2
Correlation Matrix

	VO2Peak	20m PACER	16m PACER	600-yard run/walk
Age	-0.05	-0.02	-0.03	0.03
Height	0.30	0.18	0.19	-0.43*
Weight	-0.29	-0.04	-0.07	0.11
BMI	-0.50*	-0.12	-0.17	0.36
20m PACER	0.74**	1.00	0.94**	-0.62**
16m PACER	0.77**	0.94**	1.00**	-0.74**
600-yard run/walk	-0.80**	-0.62**	-0.64**	1.00

* $p<0.05$ ** $p<0.01$

To evaluate the impact of age, gender and BMI in conjunction with field test performance, on VO2peak we performed a stepwise multiple regression. Only field test performance, gender, and BMI were significant predictors of VO2peak. Thus, from the three field tests, different regression equations were developed. These are listed in Table 3. Scatter plots of predicted and measured VO2 peak are shown in Figures 1, 2, 3.

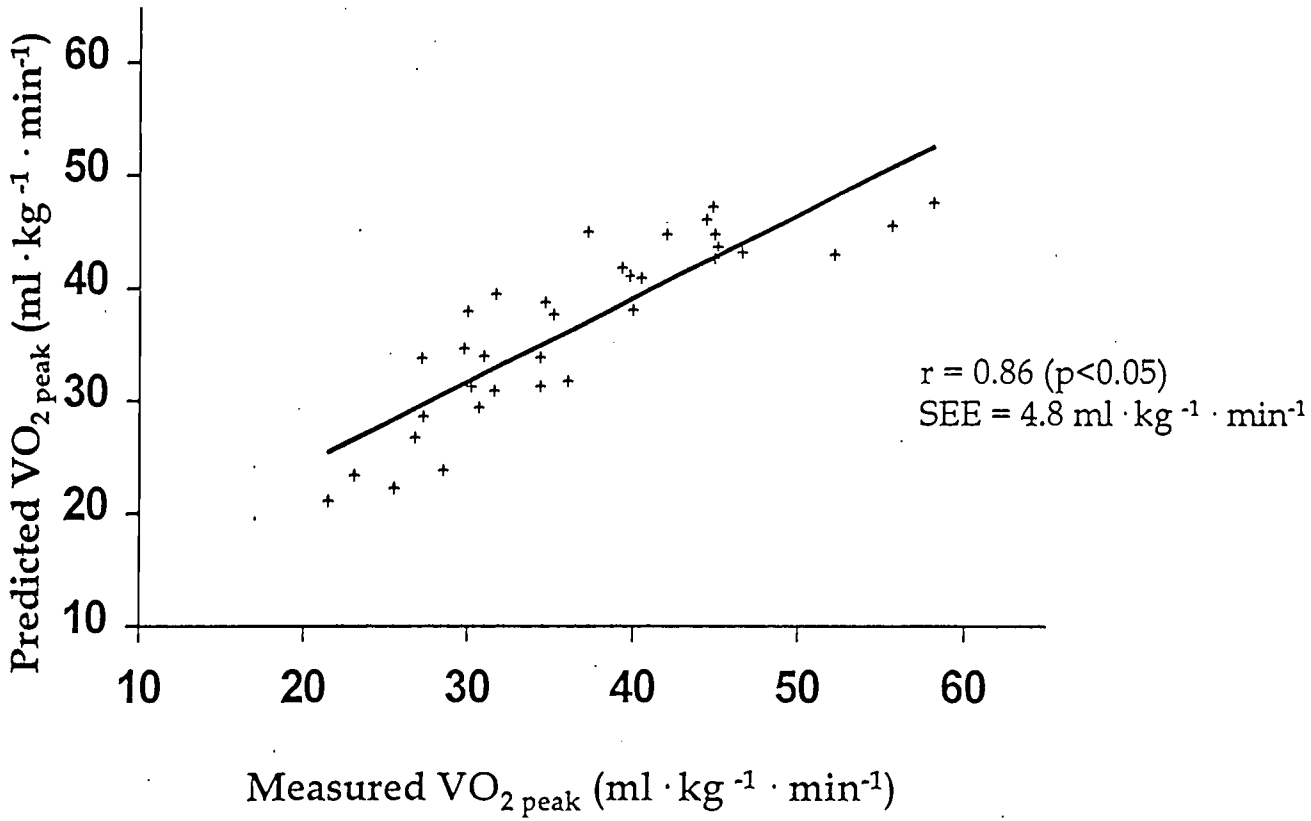

Figure 1: Relationship between measured and predicted VO2peak using 600-yard run/walk time, BMI and gender. The regression equation is presented in table 3.

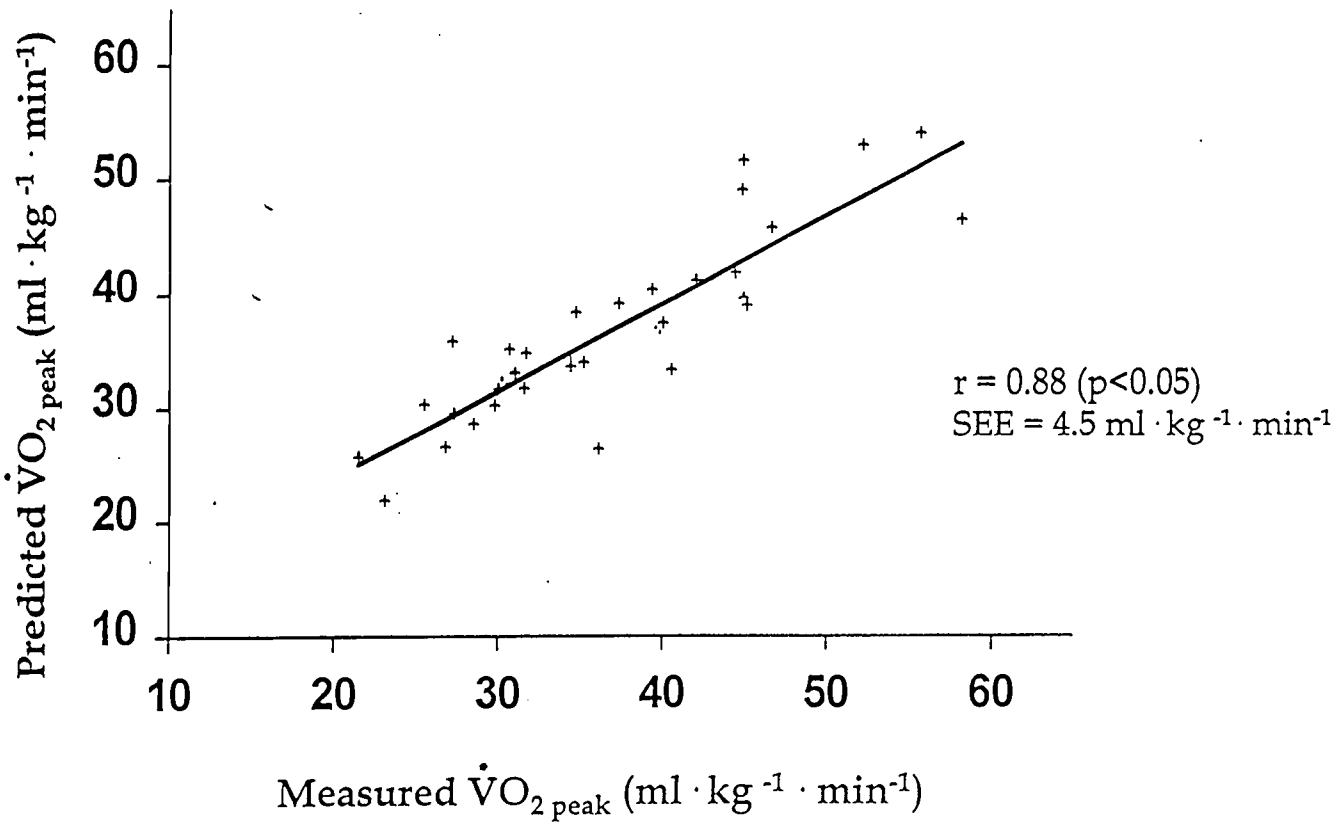

Figure 2: Relationship between measured and predicted VO2peak using the number of 20m laps completed, BMI and gender. The regression equation is presented in table 3.

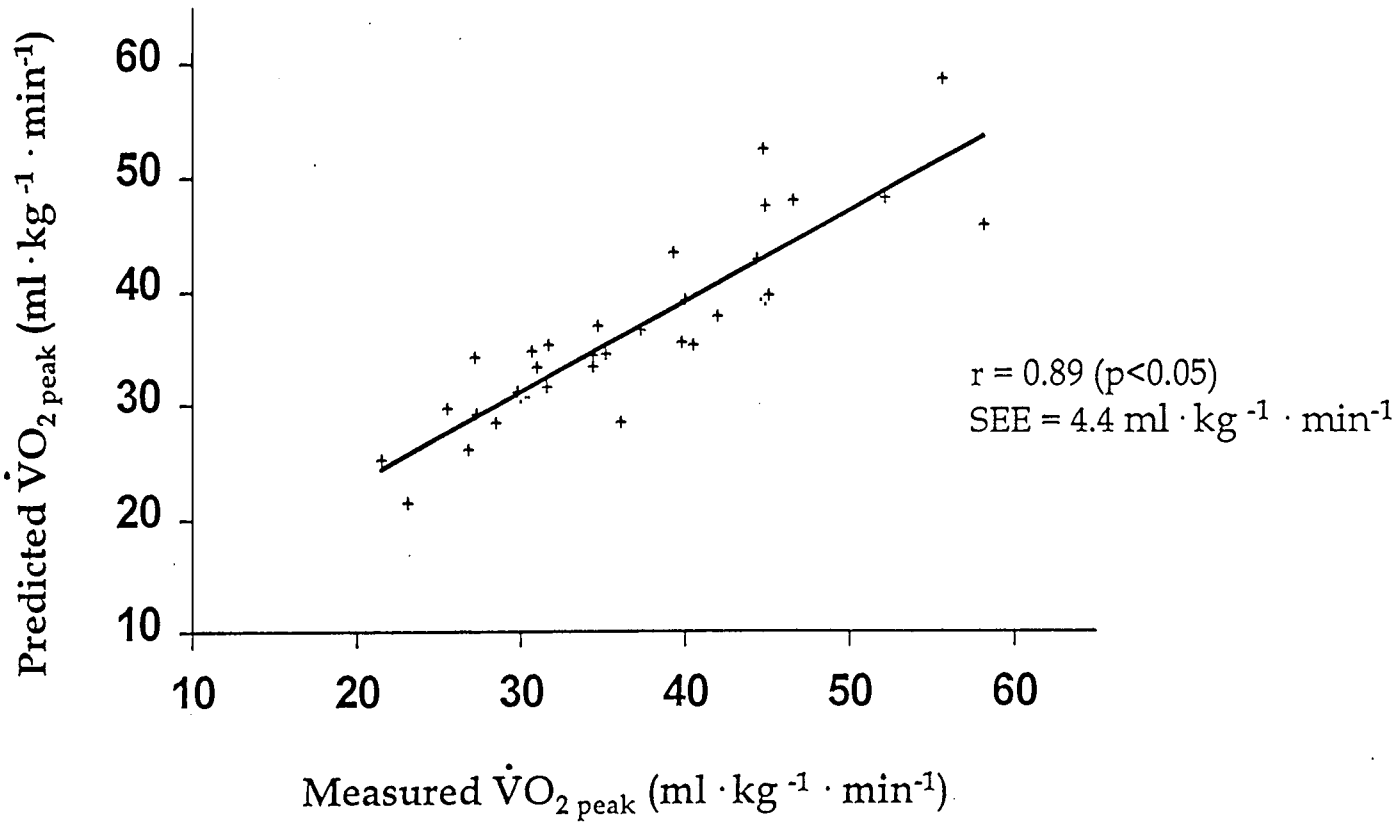

Figure 3: Relationship between measured and predicted VO2peak using the number of 16m laps completed, BMI and gender. The regression equation is presented in table 3.

Table 3
Prediction Equations Using Field Test and Descriptive Variables to Estimate VO2peak

Equation 1: 600-yard run/walk time

VO2peak ($ml \cdot kg^{-1} \cdot min^{-1}$) = -5.24(600 yard run/walk time) - 0.37 (BMI) - 4.61 (sex) + 73.64
$R = 0.86$; $R^2 = 0.74$ ($p<0.0001$) SEE = 4.8

Equation 2: 20m PACER

VO2peak ($ml \cdot kg^{-1} \cdot min^{-1}$) = 0.35 (# of laps, 20m PACER) - 0.59 (BMI) - 4.5 (sex) + 50.8
$R = 0.88$; $R^2 = 0.77$ ($p<0.0001$) SEE = 4.5

Equation 3: 16m PACER

VO2peak ($ml \cdot kg^{-1} \cdot min^{-1}$) = 0.28 (# of laps, 16m PACER) - 0.54 (BMI) - 4.4 (sex) + 48.6
$R = 0.89$; $R^2 = 0.79$ ($p<0.0001$) SEE = 4.4

For sex: Use 1 for boys and 2 for girls.

Conclusions:

1. All three field tests were highly reliable.

2. All three field test were valid predictors of VO2peak. Adding BMI and gender in a stepwise multiple regression, generated an increased R and explained approximately 80% of he variance for each test. The SEE was also manageable for each test, at just under 5 $ml \cdot kg^{-1} \cdot min^{-1}$, which is comparable to non-disabled populations. The concurrent validity appears to be equal or higher than that found for non-disabled populations.

3. All three field test were valid as a measurement of cardiorespiratory fitness in these adolescents with MR with mild limitations.

Disk: DC-Study
File: validation

References

Fernhall, B., A. L. Millar, G.T. Tymeson, and L.N. Burkett. Maximal exercise testing of mentally retarded adults: reliability study. Arch. Phys. Med.Rehabil., 71:1065-1068, 1991.

Pitetti, K.H., J.H. Rimmer, and B. Fernhall. Physical fitness and adults with mental retardation: an overview of current research and future directions. Sports Med., 16:23-56, 1993

Houston Independent School District
Houston, Texas
Spring, 1996

Purpose and Overview

During the spring semester of 1996, a study was conducted at the Houston Independent School District (HISD) in Texas to field-test selected items on a tentative health-related criterion-referenced physical fitness test for youngsters with disabilities. Field testing was conducted to: 1) provide descriptive data related to the test items administered; 2) determine whether subjects were able to learn how to perform test items in the time period permitted; 3) determine whether subjects were able to score on test items to the extent needed to fall within a measurable achievement continuum; 4) determine the percentage of subjects reaching healthy fitness zones tests established by Project Target.

Disabilities included in the study were mental retardation (MR), Cerebral Palsy (CP), Congenital Anomaly/Amputee (CA/A), and spinal cord injury (SCI). Data were collected on 38 individuals between the ages of 10 and 17. Nine subjects were omitted from the sample because of incomplete data provided by testers. The sample included 29 youngsters including 11 males and 9 females with mental retardation and mild limitations in physical fitness and five males and four females with cerebral palsy. Subjects with cerebral palsy were subclassified in accord with The Cerebral Palsy International Sport and Recreation Association (CP-ISRA) classification system. Seven subclassifications were used.

Methods and Procedures

Test items field analyzed for subjects identified as mentally retarded included 1) body mass index; 2) the 16m Progressive Aerobic Cardiovascular Endurance Run (PACER); 3) the 600-yard run/walk; 4) Level I of the Target Aerobic Movement Test; 5) dominant grip strength; 6) the back saver; 7) the trunk lift; and 8) the shoulder stretch. Test items field analyzed for youngsters with cerebral palsy included 1) body mass index; 2) Level I of the Target Aerobic Movement Test; 3) and the seated push-up. Additional test items were administered to both groups, however, the lack of subject numbers prohibited meaningful analysis.

The field test was conducted by four teachers of physical education at HISD following training sessions on the test administration by the project director. Test items were administered in accordance with the 1995 Project Target Test Manual.

The subjects in the study were enrolled in various classes in the school district. The frequency of physical education each week varied amongst the subjects as did the amount of time used for learning and performing test items by different subjects. Field-testers were instructed to introduce test items to subjects and teach students how to perform tests until the students were able to perform tests correctly. Once field-testers were convinced that the subjects understood

how to perform the test items, data were collected and recorded.

Table 1

Mean Scores of Subjects with Mental Retardation on Selected Physical Fitness Test Items

	Group	
Variable	**Females**	**Males**
Number	9	11
Age	15.3(n=20)	14.8(n=20)
BMI	31.3(n=7)	23.5(n=10)
16M Pacer laps (#)	9(n=5)	21(n=3)
600yd Run(sec)	833(n=20)	441(n=20)
Dominant		
Grip Strength (kg)	21(n=5)	27(n=3)
Back Saver Sit & Reach(cm)		
Right Leg	25(n=9)	42(n=11)
Left Leg	22(n=9)	38(n=11)

Table 2

Number and Percent of Subjects with Mental Retardation Reaching Minimal Prudential FITNESSGRAM Standards

	Group		
Variable	**Female**	**Male**	**Total**
BMI	2/7 (29%)	5/10 (50%)	7/17 (41%)
Back Saver			
Right	3/9 (33%)	8/11 (73%)	11/20 (55%)
Left	2/9 (22%)	8/11 (73%)	10/20 (50%)
Trunk Lift	5/9 (56%)	10/11 (91%)	15/20 (75%)
Shoulder Stretch			
Right	3/5 (60%)	3/3 (100%)	6/8 (75%)
Left	3/5 (60%)	1/3 (33%)	4/8 (50%)

Table 3

Number of Subjects with Mental Retardation Reaching Project Target Standards

	Group		
Variable	**Female**	**Male**	**Total**
TAMT-Level I	7/9 (78%)	10/10 (100%)	17/19 (89%)
Dominant Strength	4/5 (80%)	1/3 (33%)	5/8 (63%)

Table 4

Number of Subjects with Mental Retardation Scoring Zero on Test Items

Variable	No. Taking Test	No. Scoring Zero	Percent Scoring Zero
PACER (16m)	8	0	0
Grip Strength			
Right	8	0	0
Left	8	0	0
Back saver			
Right	20	0	0
Left	20	0	0
Trunk Lift	20	0	0
Extended Arm Hang	5	2	40%

Results

Test Item Analysis-Mental Retardation:

Body Mass Index (BMI): Results of the study appear in Tables 1, 2, 3, and 4. Criterion used for the evaluation of performance on selected test items appears in Appendix A. The body measurement of height and weight of subjects to determine BMI appeared to be reliable and economical. A total of 41% (7 of 17) of the subjects were within healthy fitness zones established by the Prudential FITNESSGRAM. Standards range from (15.3-21 to 18.8-27). A total of 90% (9 of 10) who failed with scores outside the healthy fitness zone had BMI scores over rather than below the acceptable zone. The percentage of nondisabled adolescents who "passed" the FITNESSGRAM criteria from the National Children and Youth Fitness Study (NCYFS) was at least 73% for females and 82% for males throughout the ages of 10-17 (Looney & Plowman, 1980).

16M Pacer: Based on observation of the field testers, it was clear that the 16m PACER was a test item understood by the subjects and which they performed correctly with physical assistance. The mean score in laps for females (N=5) was nine and the mean score for males (N=3) was 21. The performance of this sample was analyzed using minimum standards tentatively identified in Project Target, which appear in Appendix A. Only 25% (2 of 8) subjects met the minimum adjusted standards. Two of five females met the standards and none of the males reached the standard.

TAMT: The TAMT assesses the ability of youngsters to exercise within a recommended targeted rate zone (THRZ) for a sustained period of time. A total of 89% (17 of 19) of the subjects passed the 15 minute test. This included 100% (10 of 10) of males and 78% (7 of 9) of females.

Dominant Grip Strength: Although the concept of squeezing or gripping to exert maximum force requires learning, subjects were able to perform appropriately and clearly obtain scores within an achievement continuum. When comparing values in this study with 20th percentile values set for nondisabled subjects using Project Target minimum standards presented in Appendix A (Winnick & Short, 1995), 62% (5 of 8) male and female subjects met or exceeded the minimum standard. This criteria for dominant grip strength was met by 80% (4 of 5) females and 33% (1 of 3) males

Back Saver: Out of 20 subjects, a total of 55% (11 of 20) reached or surpassed the Prudential FITNESSGRAM standards for the right leg. This includes 33% (3 of 9) females and 73% (8 of 11) males. A total of 50% (10 of 20) reached or surpassed the Prudential FITNESSGRAM standards for the left leg. This includes 22% (2 of 9) females and 73% (8 of 11) males. Criteria for males was 8 inches and for females the criteria ranged from 9-12 inches depending upon the age of the subject.

Trunk Lift: The results in this study indicate that this sample performed the trunk lift successfully. A total of 75% (15-20) subjects met the Prudential FITNESSGRAM criterion-referenced standard, which was 9-12 inches for both males and females. A total of 56% (5 of 9) females passed and 91% (10 of 11) males met the standard. For the present study it was acceptable for testers to hold the feet down during the trunk lift.

Shoulder Stretch: Unfortunately there were only a small number of males and females who were tested on this item. In the total sample, 75% (6 of 8) passed the right shoulder stretch, and 50% (4 of 8) passed the left shoulder stretch. Of the male subjects, 100% (3 of 3) passed the right shoulder stretch and 33% (1 of 3) passed the left shoulder stretch. Of the female subjects, 60% (3 of 5) passed the right shoulder and left shoulder.

600-yard run/walk: The 600-yard run/walk was administered to 20 subjects in this study and group results appear in Table 1. Further analysis was not conducted since this test item will not be recommended as a test item for individuals with mental retardation on the health-related test being developed by Project Target.

Test Item Analysis-Cerebral Palsy:

In view of the small number of subjects included in the study (n=9), and the fact that this sample was further subclassified, meaningful analysis and interpretation of the data is limited. Raw data including subject subclassification is presented in Appendix C.

Body Mass Index (BMI): In regard to BMI, only 25% (2 of 8) of the subjects fell within recommended healthy fitness zones. Interestingly, 5 of 6 (including 4 of 4 males) outside of acceptable levels were below minimal acceptable levels (underweight).

TAMT: The TAMT was administered to all nine subjects. A total of 66% (6 of 9) passed the test. Interestingly, all three failing were class I subjects with cerebral palsy. A total of 25% (1 of 4) class I subjects passed the TAMT.

Seated Push-up: The seated push-up was administered to five subjects in the study for which this test item is recommended. A total of 80% (4 of 5) met both of 5-second and 10-second standard for success.

Summary and Conclusion

In regard to the objectives of this study, the project was successful in providing descriptive data reflecting the performance of subjects. Further field testing supported the contention that subjects were able to learn how to perform test items and obtain scores clearly falling on the achievement continuum being used for each test item. In reviewing the results of the study it appears that youngsters with mental retardation in this sample were generally unfit. This is partially noticable by their BMI and 16m PACER scores.

Although subject numbers were low in some instances, results generally supported a reasonable attainability by youngsters with mental retardation, except in the 16m run. In regard to performance items, at least 50% of the subjects met standards used in the back saver (right and left), trunk lift, shoulder stretch, the TAMT, and dominant grip strength. A total of 89% (17 of 19) successfully passed Level I of the TAMT. Only 25% (2 of 8) subjects met minimum adjusted standards for performance in the 16m run. In regard to the nine subjects with cerebral palsy, 66% (6 of 9) passed Level I of the TAMT. A total of 82% (23 of 28) of the subjects in the study passed the TAMT. This total is 92% (22 of 24) if class I youngsters with cerebral palsy are not included in the sample.

Disk: Project Target#3
File: houston

Subject snd Gender	Age (yrs)	Height (cm)	Weight (kg)	BMI	16M Pacer laps	600yd Run (sec)	TAMT	Dominant Grip Strength(kg)	Back Saver Sit & Reach R	Back Saver Sit & Reach L	Trunk Lift (inches)	Shoulder Stretch R	Shoulder Stretch L	Curl-up
1. (RA) M	12	165	51	18.8	——	184	P							——
2. (H) M	15	182	56	16.9	——	211	P							
3. (RZ) M	16	188	99	28.0	——	282	P							
4. (WJ) M	15	170	101	34.9	——	357	P							
5. (GD) M	13	142	46	22.8	——	240	P							
6. (TB) M	13	165	50	18.4	——	300	P							
7. (AM) M	13	——	——	——	14	466	——							
8. (T) M	17	173	69	23.1	——	798	P							
9. (LC) M	17	173	61	20.4	——	782	P							
10.(HN) M	17	165	92	33.9	15	735	P							
11. (ZJ) F	15	166	47	17.0	35	502	P							
12. (BS) F	17	163	54	20.4	——	198	P							
13.(MM) F	12	——	——	——	18	330	P							
14. (LA) F	16	——	——	——	11	490	F							
15. (MH) F	15	152	69	29.9	——	948	P							
16. (LO) F	16	157	70	28.5	——	995	P							
17. (SD) F	15	155	42	17.5	——	892	P							
18. (VS) F	16	137	68	36.4	6	1117	P							
19. (GD) F	17	160	155	60.5	1	1579	F							
20. (WS) F	14	168	71	25.2	11	950	P							

Study #4
School of the Holy Childhood
Rochester, New York
Spring 1996

Purpose and Overview

During the spring semester of 1996, a study was conducted at the School of the Holy Childhood in Rochester, NY pertaining to the physical fitness of individuals with mental retardation and mild limitations in fitness. The study was conducted to: 1) to provide descriptive data relative to several test items: Target Aerobic Movement Test (TAMT), Progressive Aerobic Cardiovascular Endurance Run (PACER), curl-up, isometric push-up, 10-lb. dumbbell press (ages, 10-12), 15-lb. dumbbell press (ages, 13-17); 2) determine whether subjects were able to learn to perform test items in the time period permitted; 3) determine pass/fail rates in regard to selected standards. 4) determine test-retest reliability for the isometric push-up and 15-lb. dumbbell press (ages, 13-17). 5) determine the relationship between the number of laps in the 16m PACER and TAMT, and compare the mean number of laps of those passing and failing the TAMT with performance on the 16m PACER; 6) determine intensity of activity as reflected by heart rate at the termination of the PACER test.

Subjects in the study included 42 youngsters with mental retardation and mild limitations in physical fitness. The subject sample was obtained from 6 classes of 51 subjects. All subjects in the classes were tested except for those 1) who failed to submit a consent form, 2) were reporting to weight training as a part of the their physical education program, 3) were out of the age range designated for the study, or 4)did not fit the definition of the subjects to be included in this study. The sample included 19 males and 23 females.

Methods and Procedures

Test items which were field tested included the Target Aerobic Movement Test (TAMT), The Progressive Aerobic Cardiovascular Endurance Run (PACER); the isometric push-up, 10-lb. dumbbell press (ages, 10-12), 15-lb. dumbbell press (ages, 13-17). In order to gain some additional information related to reactions of youngsters to the PACER test, additional measures were collected. These included resting heart rate, heart rate at the conclusion of the PACER run and 2 minute recovery heart rate.

Data were collected by two teachers of physical education at the school of the Holy Childhood and Project Target staff under direction of the project director. Test items were administered in accordance with the procedure associated with the 1995 Project Target test manual.

In this study subjects were tested both on the TAMT and the PACER. Youngsters were tested on an alternating basis on these two test items with at least 1 day in between tests. One-half of the group was tested first on the TAMT and one-half of the group was tested first on the PACER. In addition to the procedures in the 1995 Project Target test manual a resting heart rate was obtained prior to administration of the PACER test and a 2 minute recovery heart rate was taken following completion of the test for each student. Also, during the PACER test, heart rate was recorded at the completion of each 16m lap and the end of the exercise session for each

student. During the TAMT, heart rate after each minute and the number of minutes completed in the target heart rate zones were recorded.

Test/retests were administered for the curl-up, isometric push-up, and the dumbbell press. At least one day and no more than one week elapsed between test and retest.

The youngsters in the study were enrolled in various classes in the school. The frequency of physical education each week varied amongst the subjects as did the amount of time used for learning and performing test items. The field testers were instructed to introduce test items to subjects and teach students how to perform test items.

Results

The results of the study will be presented in three sections: test item analysis, test item reliability, and PACER and TAMT relationships.

Test Item Analysis

Mean scores on the 16m PACER and isometric push-up are presented in Table 1. Because this sample exhibited considerable difficulty with the curl-up and insufficient number of subjects completed the dumbbell test, descriptive data on these items are not provided.

Because few appeared to be able to complete the curl-up appropriately, it was decided by the investigators to follow-up the study with a modified version of the curl-up. In the modified version the starting position for the arms is to place them together over the anterior portion of the body and rest the fingertips of each hand on the corresponding thighs. From this position the participant is instructed to curl-up slowly sliding the fingertips approximately four inches (approximately one inch beyond the patella). Thus, the fingertips slide across the patella rather than across a measuring strip. The arms are held over the anterior portion of the body instead of straight and adjacent to the trunk as in the unmodified curl-up. A total of 14 youngsters who appeared to be of higher fitness were informally taught using this modified procedure. Subjects appeared to be much more successful using the modified procedure and rate of learning appeared to be enhanced substantially. A total of 9 of 14 (64%) met minimal specific standards using the modified version. However, some subjects exhibited some difficulty in following the cadence and/or keeping the feet flat on the floor during the curl-up.

Relative to the 16m PACER, subjects appeared to know how to perform the test. Their performance was compared to minimal specific standards for aerobic capacity established by Project Target. These standards are different for each gender and age. Table 2 presents the number and percent of subjects with mental retardation reaching 16m minimal specific standards for youngsters with MR. A total of 70% of females (7 of 10) and 0% of males (0 of 12) obtained or exceeded these standards.

Table 1
Mean Score of Subject on Selected Physical Fitness Item

Variable	Group		
	Males	Females	Total
16m PACER Laps(No.)	12 N=12	15 N=14	13.5 N=26
Isometric Push-up(Sec.)	6 N=18	9 N=20	8 N=38
PACER Terminal Heart Rate (No.)	178 N=12	169 N=14	173 N=26
15-minute Terminal Heart Rate on the TAMT	162 N=10	155 N=14	158 N=24

Thus, 32% (7 of 22) reached the set standard. Interestingly, the mean number of laps completed by females exceeded those completed by males.

Youngsters in the sample appeared to be able to learn to properly perform the isometric push-up. Their performance was compared to minimal specific standards for their test item established by Project Target (See Table 2). According to standards established by Project Target, males, ages 10-12, are expected to hold the isometric position for 20 seconds and females are expected to hold for 18 seconds to meet minimum specific standards. None of the males and only 30% (3 of 14) females were able to reach those standards. Five of seven males were unable to hold for one second. The number of subjects tested on this item was very small. Only seven males and 10 females were tested.

Table 2
Number of Percent of Subjects with Mental Retardation Reaching Minimal Standards

	Group		
Variable	**Males**	**Females**	**Total**
Isometric Push-up (Ages 10-12)	0/7 (00%)	3/10(30%)	3/19(16%)
Target Aerobic Movement Test(TAMT)	10/13(78%)	14/14(100%)	24/27(89%)
16m PACER	1/12(08%)	7/10(70%)	7/22(32%)

Test Item Reliability

The isometric push-up and the dumbbell press are relatively uncommon test items and the authors found no related literature on the reliability of these test items for youngsters with mental retardation. In this study, 38 youngsters (18 males, 20 females) were given a test-retest on the isometric push-up. Alpha coefficients of correlation were a = .83 for the entire sample (N= 38). The proportion (P) coefficient was .94 and the kappa (k) coefficient was .77.

A total of 12 youngsters were given three trials on the 15-lb dumbbell press. The alpha coefficient between trial 1 and trial 2 was .98 (P=.91, k=.61). The three-tail alpha was also .98. All males were able to perform at least 1 repetition on one of three trials. However, five of seven females tested were unable to complete even a single repetition of the dumbbell press.

The Relationship Between the PACER and TAMT

Students in this sample were, relatively speaking, very successful in participating in and passing the TAMT. A total of 89% (24 of 27) of the subjects passed the item. All 14 female subjects taking the test successfully reached the passing standard. The activities performed during the test included a fast walk, playing tag, and running. The average heart rate on the subjects at the 15-minute conclusion of the test was 158. This represents approximately 75-80% maximum predicted heart rate (MPHR) for this sample.

As presented earlier, only 41% of the sample obtained standards set for the 16m PACER. The average terminal heart rate (heart rate recorded when subjects stopped running) was 173 which is approximately 85% of MPHR.

Table 3 presents the number and percentage of subjects at or exceeding selected heart rates at the end of the PACER. These data show that the subjects continued activity if the intensity was at a moderate level but discontinued as running became more vigorous. This, of course, is expected. What is less likely is the continuance of activity at rather vigorous levels. A total of 92% (24 of 26) were still active at heart rates of 160 or above. However, this percentage declined to 69% and 23%, respectively, at heart rates of 170 and 180. The latter intensities are 85% to 90% MPHR. Also interesting are the results that 18 of 26 subjects continued activity at an intensity of 170 and 4 subjects performed at a near maximum predicted heart rate (190 or above). These data suggest that subjects gave "good" effort and that poor results on the PACER can be attributed, at least in part, to poor aerobic capacity. This counters the often made contention that subjects with mental retardation discontinue exercise primarily because of poor motivation and unwillingness to endure discomfort.

Three subjects failed the TAMT. Their performance on the 16M PACER was composed with subjects passing the TAMT. The mean number of laps completed by the three failing the TAMT was 13 and the mean number of laps by the 24 subjects passing the TAMT was 11.

In analysis results, it is clear that the 16m PACER which is designed to measure aerobic capacity, is more demanding than the TAMT which is designed to measure aerobic behavior. Their results support the belief that this sample is generally able to sustain moderate physical activity but, reflects a rather poor level of aerobic capacity.

Table 3
Number and Percentage of Subjects at or Exceeding Selected Heart Rates at the End of the PACER

Heart Rate	Number and Percentage		
	Males	Females	Total
140	12/12 (100%)	14/14 (100%)	26/26 (100%)
150	12/12 (100%)	13/14 (93%)	25/26 (96%)
160	12/12 (100%)	12/14 (86%)	24/26 (92%)
170	10/12 (83%)	8/14 (57%)	18/26 (69%)
180	4/12 (33%)	2/14 (14%)	6/26 (23%)
190	2/12 (17%)	1/14 (4%)	3/26 (12%)
200	1/12 (8%)	0/14(0%)	1/26(4%)

Summary and Conclusion

Generally, the test items selected for this study appeared to be suitable for this sample. Subjects were able to learn and correctly perform test items. Based on observation and test results, the curl-up was the one test item which was questioned seriously in regard to appropriateness. In view of this, a modified version was established and piloted. This modified version involves the subjects reaching across the patella rather than reaching for a strip along the sides of the individual. This modification enhanced learning and performance, however, some subjects exhibited difficulty in adhering to the cadence and holding the feet flat while performing the curl-up.

Although the test items generally appeared to be suitable for this sample, participants were generally not highly successful in reaching established standards except for the TAMT. Both males and females performed poorly on the isometric push-up and males scored poorly on the 16m PACER. Interestingly, subjects successfully completed the TAMT indicating the ability to sustain moderate physical activity for a sustained period of time.

In regard to test-retest reliability it was found that the 15-lb dumbbell press exhibited very high (very dependable) reliability. The alpha coefficient of correlation was a= .98. The test-retest reliability of the isometric push-up was determined to be very acceptable with an alpha coefficient of correlation of .83.

Subjects in this sample were successful in performing the TAMT but exhibited poor aerobic capacity as measured in the 16m PACER. A total of 89% (24 of 27) passed the TAMT and only 32% (7 of 22) reached acceptable standards for the 16m PACER. Although the percentage of subjects passing the PACER was only 32%, data collected support the belief that subjects discontinue running the PACER because of the level of intensity required is vigorous and clearly exceed that required in the TAMT. The Target standard for males appears to be less attainable for them than is the case with females.

Although this study sheds light in regard to the original objectives, additional testing of other subjects with mental retardation and mild limitations would be beneficial. This is particularly important in regard to the muscle strength/endurance test items. This sample appears to be particularly of low fitness in muscular strength and endurance.

File: hc-mr-4
Disk: Project Target
9-9-97

Study #5
School of the Holy Childhood
Rochester, New York
Spring 1997

Purpose and Overview

During the Spring semester of 1997, a study was conducted at the School of the Holy Childhood in Rochester, New York pertaining to the physical fitness of individuals with mental retardation and mild limitations in physical fitness. The study was conducted to: (1) provide descriptive data relative to several test items: the 16m PACER, the 20m PACER, dominant grip strength, modified curl-ups, extended arm hang, flexed arm hang, 35-lb bench press, and the isometric push-up; (2) determine whether subjects were able to learn to perform test items in the test period permitted; (3) determine pass/fail rates in regard to selected standards; (4) determine test-retest reliability for several test items administered; (5) determine relationships between the 16m PACER and the 20m PACER.

Subjects in the study included 38 youngsters with mental retardation and mild limitations in physical fitness. The subject sample consisted of 38 youngsters obtained from six classes at the school. All subjects in the classes were tested except for those (1) who failed to submit a consent form, (2) were out of the age range designated for the study, or (3) did not fit the definition of the subjects to be included in this study.

Methods and Procedures

The study involved two parts. The first part involved teaching youngsters to be able to perform test items appropriately. Teaching was conducted by three physical educators employed at the School of the Holy Childhood. The teaching was conducted in March of 1997 during the regular physical education classes of the pupils involved in the study. The physical education instructors taught youngsters using procedures described in the Project Target Test manual under overall direction provided by the Project Target director and coordinator. In January of 1997, a meeting was held at the School of the Holy Childhood in order for the project staff to convey to the physical education instructors the purposes of the study, the time line, and to review the test item procedures to be used by the teachers at the Holy Childhood School. Test items were administered in accordance with the procedures associated with the November 1996 Project Target test manual.

file: overview.hc
disk: mr_5

The first formal testing day was April 7, 1997. Re-testing for reliability purposes was conducted on April 21, 1997. On these two test dates, the following items were tested: the 16m PACER, the modified curl-up, the isometric push-up, the 35-lb. bench press, the flexed arm hang, the extended arm hang, and grip strength. On May 19, 1997, subjects were tested on the 20m PACER in order to compare results of that testing with performance of the 16m PACER on April 21, 1997. All formal testing was done by a testing team trained and organized by the Project Target director and coordinator. The testing team included the project director and coordinator, five graduate assistants, one undergraduate student, and three physical education teachers employed at the School of the Holy Childhood.

Procedures followed on test days were consistent. The first step was to introduce the staff to the students being tested and to motivate the students to perform as well as possible. Subjects were informed that they would receive a Project Target shirt if they performed to the best of their abilities. All subjects involved in the study were provided with a free shirt upon completion of the second day of testing.

Following the orientation, a brief warm-up was provided for each of the subjects. The warm-up was conducted for approximately two minutes during which time subjects typically performed knee bends, toe touches, curl-ups, isometric push-ups, trunk movements, and two laps of the 16m PACER at a very easy pace. Following the warm-up, students were assigned to testers who took each of the youngsters to a particular station in order to begin the testing procedures. Once test items were administered to a student at a station, they student moved to a different station until all tests were completed. All test items were administered in a single day in random order with the exception of the 16m PACER. In each case, the PACER was administered as the last test item in each testing session. All subjects performed the 16m PACER, the dominant grip strength test, and the modified curl-up. Individuals ages, 13-17. were also administered the 35-lb. bench press and the flexed arm hang. Individuals ages, 10-12, were administered the extended arm hang and the isometric push-up instead of the 35-lb. bench press and the flexed arm hang. A brief rest period was given to each of the subjects between the performance of each of the test items.

The youngsters in the study were enrolled in various classes in the school. The frequency of physical education each week and the amount of time used for learning and performing test items varied amongst the subjects. In recording the results, test administrators were instructed to record results of subjects who appeared to know how to execute the test item correctly. In cases where test items could not be executed properly, a comment on the test recording form indicating this was written. Subjects generally ran the 16m PACER with a partner. In some instances the partner ran with physical contact with the student. However, in no case was a student given any physical benefit from running the PACER with a partner (i.e., pulled along, etc.).

Results

The results of the study will be presented in three sections: learning and passing test items, test item reliability, and relationships between 16m and 20m PACER laps.

Learning and Passing Test Items

Mean scores of subjects on physical fitness test items appear in Table 1. In order to enhance comparison of performance between test 1 and test 2, only the data of subjects participating in both tests are presented.

With possibly one exception, subjects in the study were clearly able to perform test items correctly. The one possible exception to this is related to the performance on the modified curl-up. When the curl-up was administered during test 1, 10 of the 37 subjects tested did not correctly perform the test item (See Table 2). In the second test administration, only 2 of 35 did not perform the test item appropriately and correctly. This result influenced both the reliability coefficient between test 1 and 2 and the passing percentage rate associated with test 1 and test 2 (See Tables 2 and 3). Fewer subjects failed the curl-up on test 2 because they learned to perform the test item correctly. This contributed to a higher pass rate for test 2. Subjects may have learned the test items within the two week period between test 1 and 2 or the test administrator may have been more lenient in interpreting correct performance during the retest.

As a part of data analysis, the number and percentage of subjects reaching the Project Target minimum specific standards for youngsters with mental retardation and mild limitations in physical fitness were recorded (See Tables 2, 3, 4, and 5). In regard to the 16m PACER, between 30% (test 2) and 34% (test 1) of the subjects were able to meet the minimum specific standards. Females were much more successful in reaching the specific standards than were males (See Table 2). Youngsters were more successful in attaining standards associated with dominant grip. A total of 53% were able to attain minimum specific standards with similarity in performance for both boys and girls.

In regard to the modified curl-up, 49% of the subjects attained the minimum specific standards on test 1 and 74% attained the standard on test 2. This was primarily due to the fact that a greater percentage of subjects had learned to perform the curl-up for test 2. Since not learning that item constituted a failure for the analysis of data, this ability to perform the test correctly during test 2 had a positive impact on the passing rates associated with the modified curl-up on test 2.

In regard to the extended arm hang, only 12 subjects participated. Of the 12, five were males and none of these males successfully reached the minimum specific standards. Females on the other hand, although few in number, (n=7) were much more successful in reaching the minimum specific standards for females with mental retardation (See Table 2).

Table 1
Mean Scores of Subjects on Selected Physical Fitness Test Items [1]

Variable	Males	Females	Total
PACER (16m)			
Test 1	21.3 (n=13)	20.0 (n=22)	20.5 (n=35)
Test 2	22.6 (n=13)	20.0 (n=22)	21.0 (n=35)
Dominant Grip			
Test 1	22.0 kg. (n=14)	16.5 kg. (n=22)	18.6 kg. (n=36)
Test 2	24.1 kg. (n=14)	17.0 kg. (n=22)	19.8 kg. (n=36)
Isometric Push-up			
Test 1	13.0 sec. (n=1)	35.3 sec. (n=7)	32.5 sec. (n=8)
Test 2	40.0 sec. (n=1)	35.6 sec. (n=7)	36.1 sec. (n=8)
35-lb. Bench Press			
Test 1	21.2 (n=9)	7.9 (n=14)	13.1 (n=23)
Test 2	22.0 (n=9)	8.2 (n=14)	13.6 (n=23)
Extended Arm Hang			
Test 1	11.5 sec. (n=4)	25.6 sec. (n=7)	20.5 sec. (n=11)
Test 2	5.5 sec. (n=4)	28.0 sec. (n=7)	19.8 sec. (n=11)
Flexed Arm Hang			
Test 1	5.1 sec. (n=8)	3.0 sec. (n=9)	4.0 sec. (n=17)
Test 2	4.0 sec. (n=8)	3.7 sec. (n=9)	3.8 sec. (n=17)
Curl-up (modified)			
Test 1	20.5 (n=8)	22.3 (n=17)	21.4 (n=25)
Test 2	19.9 (n=8)	23.4 (n=17)	22.3 (n=25)

[1] Includes subjects available for both Test 1 and Test 2.

Table 2
Number and Percent of Subjects Reaching Project Target Minimal Specific Standards for Youngsters with Mental Retardation and Mild Limitations in Physical Fitness [2]

Test Item	Males	Females	Total
16m PACER			
Test 1	02/15 (13%)	09/17 (53%)	11/32 (34%)
Test 2	02/13 (15%)	07/14 (41%)	09/30 (30%)
Dominant Grip			
Test 1	07/15 (47%)	13/23 (56%)	20/38 (53%)
Test 2	07/14 (50%)	12/22 (56%)	19/36 (53%)
Curl-ups (modified)			
Test 1	05/14 (36%)	13/23 (56%)	18/37 (49%)
Test 2	10/14 (71%)	16/21 (76%)	26/35 (74%)
Extended Arm Hang (ages 10-12)			
Test 1	00/05 (00%)	06/07 (86%)	06/12 (50%)
Test 2	00/05 (00%)	05/07 (71%)	05/12 (42%)
Flexed Arm Hang (ages 13-17)			
Test 1	02/10 (20%)	04/15 (27%)	06/17 (35%)
Test 2	02/09 (22%)	06/15 (40%)	08/24 (33%)
35-lb. Bench Press (ages 13-17)			
Test 1	04/10 (40%)	07/15 (47%)	11/25 (44%)
Test 2	07/09 (67%)	07/15 (47%)	13/24 (54%)
Isometric Push-up (ages 10-12)			
Test 1	00/04 (00%)	06/07 (86%)	06/11 (55%)
Test 2	01/03 (33%)	07/07 (100%)	08/10 (80%)

[2] The number and percent of subjects passing include all those who reach minimal specific standards. The number and percent of subjects failing include those not meeting the minimal specific standards or learning the test item.

Table 3
Test-Retest Reliability of Males and Females with Mental Retardation and Mild Limitations in Physical Fitness Performing Selected Test Items

Test Item	n	Alpha (a)	Proportion of Agreement (P)
PACER (16m)	35[1]	.98	.93
Dominant Grip Strength	36	.96	.92
35-lb. Bench Press	23	.91	.82
Extended Arm Hang	11	.85	.72
Flexed Arm Hang	17	.93	.82
Curl-up (modified)	25	.82	.72

[1] The P coefficient was computed on 30 subjects.

A greater number of individuals completed the flexed arm hang (See Table 2). Approximately 1 out of 3 were able to reach the minimal specific standards for youngsters with mental retardation. Again, females appeared to have more success in attaining minimum standards than did males. The results associated with the flexed or extended arm hang suggest that this sample has difficulty with upper body strength and endurance. In regard to attainability, youngsters in the sample were more successful with the 35-lb. bench press. Approximately ½ of the youngsters reached the minimal standards established for youngsters with mental retardation.

Results, in regard to the isometric push-up, are presented in Table 2. The number of subjects involved in performing this test item were few. An interesting finding, however, was the successful ability of females to attain minimal specific standards on the isometric push-up.

Test-Retest Reliability

As indicated previously, youngsters in this particular sample were tested on two different occasions on each of the test items associated with the study. Table 3 presents the alpha coefficients attained when comparing subjects who took both test 1 and test 2. Reliability coefficients range from .82 (curl-up) to .98 (16m PACER). The very acceptable values for the PACER (a=.98), dominant grip strength (a=.96), flexed arm hang (a=.93), and 35 lb. bench press (a=.91) are very encouraging. Although the alphas associated with the extended arm hang (a=.85) and the isometric push-up (a=.83) are a bit lower, they are also very acceptable[1].

In addition to computing alpha coefficients between test-retest groups, a proportion of agreement (P) coefficient was computed to estimate reliability. The P score indicates consistency of classification between test and retest. In this study, individuals were compared to determine consistency in reaching minimal criterion-referenced specific standards (Tables 4, 5, and 6) for their age and gender. Results summarized in Table 4 reflect proportions ranging from .72 to .93. Using a P of .70 or greater as an acceptable standard, all test items were found to consistently classify youngsters. Thus, consistency of classification ranges from 72 to 93% with the highest consistency found in the 16m PACER. A z test for significance of difference between two proportions revealed no significant differences ($p > .05$) in the proportion of subjects who passed the test or retest.

Relationship Between 16m PACER and 20m PACER Laps

As a part of the study, subjects ran the 16m PACER on two occasions and the 20m PACER on a third occasion. The results of the second administration of the 16m PACER and the results attained in running the 20m PACER were compared (See Table 7). The time elapsing between these two test administrations was 27 days with the 20m PACER being administered on the second occasion. The Pearson r attained between the two tests was $r=.95$ ($p < .01$). The mean number of laps for the 16m PACER was 21.8 and for the 20m PACER was 10.3. Interestingly, the Project Target assumed relationship between laps is that 20m lap values are approximately 63% of 16m lap scores. In this study, the lap values for the 16m PACER were 47% of those attained on the 20m PACER.

---

[1] A correlation of less than .19 indicates no relationship; .20-.39 is a low correlation; .40-.59 is a moderate correlation; .60-.79 is a moderately high correlation; and a .80-1.00 is a very high correlation. For the purposes of Project Target, Pearson, alpha, and P coefficients of .70-.79 are considered to be minimally acceptable and .80 and above to be very acceptable for test items measuring physical fitness. This is consistent with recommendations by Safrit and Wood (1995) regarding norm-referenced reliability coefficients.

Table 4
VO_{2max} and PACER
Specific Standards for Youngsters with Mental Retardation

Males						
Age	VO_{2max} (ml/kg/min.)		PACER (20 m) # laps[1]		PACER[1] (16 m) # laps	
	M	P	**M**	P	**M**	P
10	**38**	42	**4**	17	**9**	25
11	**38**	42	**10**	23	**16**	33
12	**38**	42	**16**	29	**24**	40
13	**38**	42	**21**	35	**30**	48
14	**38**	42	**27**	41	**38**	55
15	**38**	42	**33**	46	**45**	61
16	**38**	42	**38**	52	**57**	69
17	**38**	42	**44**	57	**59**	75
Females						
10	**35**	39	-	7	-	13
11	**34**	38	-	9	-	15
12	**33**	37	**1**	13	**5**	20
13	**32**	36	**4**	15	**9**	23
14	**31**	35	**6**	18	**11**	26
15	**31**	35	**12**	23	**19**	33
16	**31**	35	**17**	28	**25**	39
17	**31**	35	**22**	34	**31**	46

M - Minimal Standards associated with a 10% downward adjustment of VO_{2max} from minimal general standards
P - Preferred Standards are associated with the minimal general standards

[1] Laps for the 16m are based upon estimates from 20m PACER lap scores.

[2] 16m laps = 1.25 (20m laps) + 3.8, S.E. = 7.4. 20m laps = .71 (16m laps) -.87, S.E. = 5.5. 20m lap values are approximately 63% of 16m lap scores.

Table 5
Isometric Push-up, Bench Press, Extended Arm Hang, Flexed Arm Hang, Grip Strength
Specific Standards for Youngsters with Mental Retardation

Age	Isometric Push-up[1] (sec.)		Bench Press[1] # Completed		Extended Arm Hang[2] (sec.)		Flexed Arm Hang[1] (sec.)		Dominant Grip Strength[3] (kg.)	
	M	P	**M**	P	**M**	P	**M**	P	**M**	P
Males										
10	**20**	40			**23**	30			**12**	18
11	**20**	40			**23**	30			**14**	21
12	**20**	40			**23**	30			**16**	25
13			**10**	20			**6**	12	**19**	29
14			**16**	33			**8**	15	**22**	33
15			**20**	40			**8**	15	**24**	37
16			**23**	47			**8**	15	**28**	43
17			**25**	50			**8**	15	**32**	49
Females										
10	**13**	25			**15**	20			**11**	17
11	**13**	25			**15**	20			**12**	19
12	**13**	25			**15**	20			**14**	22
13			**5**	10			**4**	8	**16**	24
14			**6**	13			**4**	8	**17**	26
15			**7**	14			**4**	8	**19**	29
16			**7**	14			**4**	8	**19**	29
17			**8**	15			**4**	8	**19**	29

1 Minimal specific standards reflect a 50% adjustment to minimal general standards. Preferred specific standards are equal to the minimal general standards.

2 Minimal specific standards are 75% minimal general standards. Preferred specific standards are equal to the minimal general standards.

3 Minimal specific standards are 65% minimal general standards. Preferred specific standards are equal to minimal general standards.

Table 6
General Standards for Trunk Lift and Curl-ups
Specific Standards on the Modified Curl-up
for Youngsters with Mental Retardation

Age	General Standards						Specific Standards	
	Trunk Lift[2]				Curl-ups and Modified Curl-ups # Completed		Modified Curl-ups[1] # Completed	
	L (in.)	L (cm.)	U (in.)	U (cm.)	**M**	P	**M**	P
Males								
10	9	23	12	30	**12**	24	**7**	12
11	9	23	12	30	**15**	28	**9**	15
12	9	23	12	30	**18**	36	**11**	18
13	9	23	12	30	**21**	40	**13**	21
14	9	23	12	30	**24**	45	**14**	24
15	9	23	12	30	**24**	47	**14**	24
16	9	23	12	30	**24**	47	**14**	24
17	9	23	12	30	**24**	47	**14**	24
Females								
10	9	23	12	30	**12**	26	**7**	12
11	9	23	12	30	**15**	29	**9**	15
12	9	23	12	30	**18**	32	**11**	18
13	9	23	12	30	**18**	32	**11**	18
14	9	23	12	30	**18**	32	**11**	18
15	9	23	12	30	**18**	35	**11**	18
16	9	23	12	30	**18**	35	**11**	18
17	9	23	12	30	**18**	35	**11**	18

1 Minimal specific standard is 60% of minimal general standard for curl-ups. Preferred specific standard is equal to minimal general standard for curl-ups.

2 Scores higher than 12 inches or 30 cm should not be encouraged.

Table 7
Relationship Between the 16m PACER (Test 2) and the 20m PACER (n=31)

Student	Gender	Age	Laps 16m PACER	Laps 20m PACER
1	M	17	6	3
2	M	17	13	6
3	M	16	28	16
4	M	15	92	48
5	M	15	18	8
6	M	15	36	14
8	M	13	37	13
9	M	13	13	6
11	M	12	19	8
12	M	12	7	6
13	M	12	5	0
14	M	11	9	5
17	F	17	17	6
18	F	17	21	14
19	F	16	23	11
20	F	16	23	14
21	F	16	4	1
22	F	16	41	15
23	F	16	5	0
25	F	15	17	7
26	F	15	7	2
28	F	14	8	2
29	F	14	28	17
30	F	14	45	18
31	F	13	59	23
32	F	12	16	7
33	F	12	12	8
34	F	11	8	7
35	F	11	15	8
37	F	11	21	8
38	F	11	24	17
Correlation (Pearson)			.95 ($p < .01$)	
Mean			21.8	10.3
Stand. Dev.			18.5	9.1

Discussion

Results of this study indicate that youngsters with mental retardation and mild limitations in physical fitness are able to learn in a reasonable period of time the test items that were administered to them. The test item which appears to have been most difficult for youngsters to learn is the modified curl-up. It is clear that much care needs to be given to teaching youngsters to correctly perform the test and to develop clear and objective criteria for determining correct movements so that on different administrations of the test the same standards for determining a correct performance exist.

In regard to the number and percent of subjects reaching Project Target minimum specific standards for youngsters with mental retardation, successful attainment ranged from 30% on the 16m PACER to an 80% success rate on test 2 for the isometric push-up. Subjects revealed at least a 40% passing rate for dominant grip, the modified curl-up, the extended arm hang, the 35 lb. bench press, and the isometric push-up. Subjects in this sample had particular difficulty with the flexed arm hang and the 16m PACER. Although the success rate of subjects was below 40% on these two test items, it is believed that this was due to the lack of physical fitness rather than to an overly stringent specific standard.

A key dimension of this study was the determination of reliability for several test items. The 16m PACER item appears to be highly reliable since an a=.98 and a P of .93 were found between test 1 and test 2 performance. This result is consistent with a test-retest r=.89 reported by Leger, Mercier, Gadoury, and Lambert (1988) with 188 normal subjects between the ages of 8-19 on the 20m shuttle run. In earlier data collected as a part of Project Target an alpha coefficient of .98 was found in a test-retest of the 16m PACER with 20 males and females, ages 10-17, in 1995. In 1996, another study was conducted as a part of Project Target in which test-retest data were collected on 34 males and females with mental retardation on both the 16m and 20m PACER. A test-retest a=.96 was reported for the 16m PACER and a=.97 was reported on the 20m PACER. Subject ages ranged from 10 to 18.

In regard to dominant grip strength, the results of the study indicated a test-retest alpha coefficient of .96 and a P of .92 (See Table 3). The .96 relationship found in the present study is in agreement with a study by Daquilla (1982). Daquilla reported alpha coefficients ranging from .97 to .99 in a reliability study of youngsters designated as normal (n=50), visually impaired (n=50), auditory impaired (n=50), and orthopedically impaired (n=50). Subjects in the Daquilla study included males and females, ages 10-17. Rarick and Dobbins (1972) found correlations ranging from .90 to .96 for nonretarded boys and girls and those designated as educable mentally retarded (EMR) on left grip strength and from .88 to .98 for boys and girls on right grip strength. Rarick and Dobbins (1972) used EMR subjects ages 6-13 and normal subjects aged 6-9.

Much less reliability data is available on the 35 lb. bench press. In this study an alpha coefficient of .91 and a P of .82 were attained. In an earlier study, a test-retest correlation of a=.92 using 64 nondisabled males and females between the ages of 11-13 was found.

Several reliability studies have been conducted using field tests associated with abdominal strength/endurance. In the present study, an a=.82 and a P of .72 were found between test and retest trials on the modified curl-up. This is a very acceptable correlation coefficient. Pizzaro (1990) found a test-retest coefficient of r=.83 using subjects designated as EMR (n=44)

and r=.94 using subjects labeled as (N=37) trainable mentally retarded (TMR) on the modified sit-up test. Subjects for the study included 81 males and females between the ages of 12-15. Johnson and Londeree (1976) in testing 1105 male and female individuals with moderate mental retardation between the ages of 6-21 on the 30-second sit-up, found test-retest coefficients that ranged from .78 to .99. As indicated earlier, the reliability coefficient is a very acceptable one that can be improved if subjects learn the test and testers identify objective criteria and follow these in administering the test.

The test-retest reliability coefficient associated with the flexed arm hang in the present study was a=.93 and the P found was .82. The correlation coefficient is very acceptable and quite consistent with previous studies testing individuals with disabilities on the flexed arm hang. For example, Johnson and Londeree (1976) found a test-retest value ranging from .78 to .99 using 1105 males and females, ages 6-21, on the flexed arm hang. Daquilla (1982) found the following values using same day test-retest procedures on the flexed arm hang for a variety of groups: an alpha coefficient of .93 with normal males and females, ages 10-17; an alpha coefficient of .84 with 50 males and females with visual impairments, ages 10-17; an alpha coefficient of .92 with 50 males and females, ages 10-17, who exhibited auditory impairments; and an alpha coefficient of .96 using males and females, ages 10-17, with orthopedic impairments.

The present study tested only 11 subjects to analyze reliability in regard to the extended arm hang. An a=.85 and a P =.72 were obtained using the 11 subjects in the study. This item, however, was investigated in a previous study conducted as a part of Project Target. In that study, 38 males and females, ages 10-12, with mental retardation and mild limitations in physical fitness were tested on the isometric push-up and a test-retest alpha value of .83 was attained using the subjects involved in that study.

References

Daquilla, G.A. (1982). Reliability of selected health and performance related test items from the Project UNIQUE Physical Fitness Test Battery. Unpublished master's thesis, SUNY College at Brockport.

Johnson, L., & Londeree, B. (1976). Motor Fitness Testing Manual of the Moderately Mentally Retarded. Washington, DC: AAHPER.

Pizzaro, D.C. (1990). Reliability of the health related fitness test for mainstreamed educable and trainable mentally handicapped adolescents. Adapted Physical Activity Quarterly, 7, 240-248.

Rikli, R.E., Petray, C., & Baumgartner, T.A. (1992). The reliability of distance run tests for children in grades K-4. Research Quarterly for Exercise and Sport, 63, 270-276.

Safrit, M.J., & Wood, T.M. (1995). Introduction to Measurement in Physical Education and Exercise Science (3rd edition). St Louis, MO: Mosby-Year Book, Inc.

Study #6
School of the Holy Childhood
Rochester, New York
Spring 1998

Purpose and Overview

During the Spring semester of 1998, a study was conducted at the School of the Holy Childhood in Rochester, New York pertaining to the physical fitness of individuals with mental retardation and mild limitations in physical fitness. Specifically the study focused on three test items: back saver sit and reach, shoulder stretch, and trunk lift. The study was conducted to: (1) provide descriptive data relative to the back saver sit and each; (2) determine whether subjects were able to perform each test item; (3) determine pass/fail rates in regard to selected standards; and (4) determine test-retest reliability for all three test items administered.

Subjects in the study included 40 youngsters with mental retardation and mild limitations in physical fitness. The subject sample consisted of four classes at the school. All subjects in the classes were tested except those (1) who failed to submit a consent form, (2) were out of the age range designated for the study, or (3) did not fit the definition of the subjects to be included in this study.

Methods and Procedures

The study involved two parts. The first part involved teaching youngsters to be able to perform test items appropriately. Teaching was conducted by three physical educators employed at the School of the Holy Childhood. The teaching was conducted in late January and early February 1998 during the regular physical education classes of the pupils involved in the study. The physical education instructors taught youngsters using procedures described in the Project Target Test manual under overall direction provided by the Project Target director and a graduate assistant. In January 1998, a meeting was held at the School of the Holy Childhood in order for the Project Target staff to convey to the physical education instructors the purposes of the study, the time line, and to review the test item procedures to be used by the teachers at Holy Childhood. Test items were administered in accordance with the procedures associated with the January 1998 Project Target Test manual.

The first formal testing date was February 12, 1998. Re-testing for reliability purposes was conducted on February 26, 1998. All formal testing was done by a testing team trained and organized by the research graduate assistant and the Project Target director. The testing team included five graduate assistants who were monitored by the Project Target director and three physical education teachers employed at the School of the Holy Childhood.

Procedures followed on test days were consistent. The first step was to introduce the staff to the students being tested and to motivate the students to perform as well as possible. Subjects were informed that they would receive a lollypop on the first day of testing and a Brockport trinket, such as a pencil or a key chain, on the second day of testing if they performed to the best of their abilities. All subjects involved in the study were provided with both a lollypop on the first day of testing and a Brockport trinket on the second day of testing.

Following the orientation, a brief warm-up was provided for each of the subjects. Each of the three items tested was given a specific warm-up. Prior to the warm-up, students were randomly assigned to one of the three testing stations. Directly before performing the back saver sit and reach each individual jogged one full length of the gymnasium and then performed sitting toe touches three times alternating legs and holding for six seconds each time. Participants would then perform the trunk twist one time for six seconds. Finally, directly prior to testing the trunk lift, students has one practice trial using procedures described in the Project Target Test Manual. Once the warm-up was completed, the students began to test procedures. Upon completing the testing procedures at a station, students moved to different stations until all tests were completed. All test items were administered in a single day in random order. A brief rest period was given to each of the subjects between the performance of each of the test items.

Since the youngsters were enrolled in various classes in the school, the frequency of physical education each week and the amount of time used for learning and performing test items varied somewhat among the subjects. In recording the results, test administrators were instructed to record results of subjects who appeared to know how to execute the test item correctly. In cases where test items could not be executed properly, a comment on the test recording form indicating this was written.

Results

The results of the study will be presented in two sections: subject performance (including passing rates) and test item reliability.

Subject Performance

According to the data recording instruments, subjects in this study were clearly able to perform the test items correctly, with the exception of one subject who on one occasion appeared to have improper body positioning on the back saver sit and reach. For this reason these data were not used for the reliability information.

Mean scores and standard deviations of the back saver sit and reach appear in Table 3. This test item is the only one for which there is continuous data; for this reason both the means and standard deviations are provided as descriptive data. In order to enhance comparison of performance between test 1 and 2, only data of subjects participating in both tests are presented.

As part of data analysis, the number and percentage of subjects reaching the Project Target general standards were recorded (See Table 4). General standards were used to assess this population due to indications that these individuals are capable of fitness levels consistent with positive health, requiring intermittent or limited support in learning and/or performing test items. After testing the sit and reach (right), 68% of the participants successfully met the general standard in test 1 and 64% met the standard in test 2. For the sit and reach (left), 73% of the subjects attained the general standard in test 1 and 61% met the standard in test 2. For this test, the males were slightly more successful in meeting the standards for both sides of the body tested.

With reference to the trunk lift, between 98% (test 1) and 92% (test 2) were able to meet the general standards. Both the females and the males were highly successful in reaching the general standards.

In regard to the shoulder stretch, right and left, 62% of the subjects attained the general standard on test 1 and 69% attained the standard on test 2, and 46% met the general standard in test 1 and 64% attained the standard on test 2 respectively. The males were slightly more successful with the shoulder stretch on both sides of the body, particularly on the left side.

Test-Retest Reliability

As indicated previously, youngsters in this particular sample were tested on two different occasions on each of the test items associated with the study. Table 5 presents the alpha coefficients attained when comparing subjects who took both test 1 and test 2. (Due to the distribution of scores obtained for the trunk lift, neither alpha nor kappa coefficients could be computed.) Reliability coefficients range from .83 (shoulder stretch L) to .96, (back saver sit and reach R). The very acceptable values for the should stretch R (a=.94), back saver sit and reach R (a=.96), and the back saver sit and reach L (a=.95) are very encouraging. Although the alpha associated with shoulder stretch L (a=.83) is a bit lower, it is also very acceptable.[1]

[1]A correlation of less than .19 indicates no relationship; .20-.39 is a low correlation; .40-.59 is a moderate correlation; .60-.79 is a moderately high correlation; and a .80-1.00 is a very high correlation. For the purposes of Project Target, Pearson and alpha coefficients of .70-.79 are considered to be minimally acceptable and .80 and above to be very acceptable for test items measuring physical fitness. This is consistent with recommendations by Safrit and Wood (1995) regarding norm-referenced reliability coefficients. Proportion of agreement (P) coefficients greater than .70 are acceptable and greater than .95 are high (Rikli, Petray, & Baumgartner, 1992).

Table 1
General Standards for Trunk Lift and Curl-ups and Specific Standards on the Modified Curl-up for Youngsters with Mental Retardation

Age	General Standards						Specific Standards
	Trunk Lift[2]				Curl-ups and Modified Curl-ups # Completed		Modified Curl-ups[1] # Completed
	L (in.)	L (cm.)	U (in.)	U (cm.)	**M**	P	
Males							
10	9	23	12	30	**12**	24	**7**
11	9	23	12	30	**15**	28	**9**
12	9	23	12	30	**18**	36	**11**
13	9	23	12	30	**21**	40	**13**
14	9	23	12	30	**24**	45	**14**
15	9	23	12	30	**24**	47	**14**
16	9	23	12	30	**24**	47	**14**
17	9	23	12	30	**24**	47	**14**
Females							
10	9	23	12	30	**12**	26	**7**
11	9	23	12	30	**15**	29	**9**
12	9	23	12	30	**18**	32	**11**
13	9	23	12	30	**18**	32	**11**
14	9	23	12	30	**18**	32	**11**
15	9	23	12	30	**18**	35	**11**
16	9	23	12	30	**18**	35	**11**
17	9	23	12	30	**18**	35	**11**

L = Lower boundary of acceptable range
U = Upper boundary of acceptable range

1 Minimal specific standard is 60% of minimal general standard for curl-ups. Preferred specific standard is equal to minimal general standard for curl-ups.
2 Scores higher than 12 inches or 30 cm should not be encouraged.

Table 2
Shoulder Stretch, Apley, Thomas, Back Saver Sit and Reach, TST
General Standards

Males							
Age	Shoulder stretch (P/F)	Apley Test (modified)	Thomas Test (modified)	Back Saver Sit & Reach (in.)	(cm.)	Target Stretch Test M	P
10	Pass	**3**	**3**	**8**	20	**1**	2
11	Pass	**3**	**3**	**8**	20	**1**	2
12	Pass	**3**	**3**	**8**	20	**1**	2
13	Pass	**3**	**3**	**8**	20	**1**	2
14	Pass	**3**	**3**	**8**	20	**1**	2
15	Pass	**3**	**3**	**8**	20	**1**	2
16	Pass	**3**	**3**	**8**	20	**1**	2
17	Pass	**3**	**3**	**8**	20	**1**	2
Females							
10	Pass	**3**	**3**	**9**	23	**1**	2
11	Pass	**3**	**3**	**10**	25	**1**	2
12	Pass	**3**	**3**	**10**	25	**1**	2
13	Pass	**3**	**3**	**10**	25	**1**	2
14	Pass	**3**	**3**	**10**	25	**1**	2
15	Pass	**3**	**3**	**12**	30	**1**	2
16	Pass	**3**	**3**	**12**	30	**1**	2
17	Pass	**3**	**3**	**12**	30	**1**	2

Table 3
Mean Scores of Subjects on the Back Saver Sit and Reach[1]

Variable	Males	SD	Females	SD	Total	SD
Sit and Reach (R)						
Test 1	25.9 (n=15)	8.83	30.3 (n=23)	9.21	28.1 (n=38)	9.16
Test 2	25.5 (n=14)	6.12	27.5 (n=19)	10.12	26.5 (n=33)	8.87
Sit and Reach (L)						
Test 1	26.5 (n=15)	5.45	30.9 (n=23)	9.29	28.7 (n=38)	8.20
Test 2	25.4 (n=14)	7.22	28.8 (n=19)	9.87	27.1 (n=33)	8.88

[1] Includes subjects available for both Test 1 and Test 2

Table 4
Number and Percent of Subjects Reaching Project Target General Standards[2]

Test Item	Males	Females	Total
Trunk Lift			
Test 1	17/17 (100%)	16/17 (94%)	39/40 (98%)
Test 2	15/17 (88%)	18/19 (95%)	33/36 (92%)
Shoulder Stretch Right			
Test 1	11/16 (69%)	13/23 (57%)	24/39 (62%)
Test 2	13/17 (76%)	12/19 (63%)	25/36 (69%)
Shoulder Stretch Left			
Test 1	11/16 (69%)	07/23 (30%)	18/39 (47%)
Test 2	13/17 (76%)	10/19 (53%)	23/36 (64%)
Sit and Reach Right			
Test 1	12/17 (71%)	15/23 (65%)	27/40 (68%)
Test 2	12/17 (71%)	11/19 (58%)	23/36 (64%)
Sit and Reach Left			
Test 1	14/17 (82%)	15/23 (65%)	29/40 (73%)
Test 2	11/17 (65%)	11/19 (58%)	22/36 (61%)

[2] The number and percent of subjects passing include all those who reach general standards. The number and percent failing include those not meeting the general standards or learning the test item.

Table 5
Test-Retest Reliability of Males and Females with Mental Retardation and Mild Limitations in Physical Fitness Performing Selected Test Items

Test Item	n	Alpha (a)	Proportion of Agreement (P)	Kappa (k)
Trunk Lift	36		.87	
Shoulder Stretch (R)	35	.94	.94	.87
Shoulder Stretch (L)	35	.83	.83	.67
Sit and Reach (R)	33	.96	.92	.83
Sit and Reach(L)	33	.95	.89	.76

In addition to computing alpha coefficients, a proportion of agreement (P) coefficient was completed to estimate reliability. The P score indicates consistency of classification (pass vs. fail) between test and retest. In this study, individuals were compared to the general standards to determine consistency in reaching pass vs. fail decisions. Results summarized in Table 5 reflect proportions ranging from .83 to .94. Using a P of .70 or greater as an acceptable standard, all test items were found to consistently classify youngsters. A criticism of the proportion of agreement is that it does not account for the classification agreements that may occur by chance. Kappa values, therefore, were calculated for all test items to determine what degree of agreement was obtained beyond the expectation of chance. Kappa values ranged from .67 (shoulder stretch L) to .87 (shoulder stretch R). Although kappa is a more rigorous test of criterion-referenced reliability than P, three of the four k values exceeded .70.

Discussion

Results of this study indicate that youngsters with mental retardation with mild limitations in physical fitness are able to learn, in a reasonable period of time, the back saver sit and reach, shoulder stretch, and trunk lift. It is clear that much care needs to be given to teaching youngsters to correctly perform the test and to develop clear and objective criteria for determining correct performance.

In regard to the number and percent of subjects reaching Project Target general standards, successful attainment ranged from 61% on the sit-and-reach L to 92% success rate on test 2 for the trunk lift. Subjects revealed at least a 45% passing rate on all of the items tested, with most of the rates of success at the 60% or better range.

A key dimension of this study was the determination of reliability for all items tested. The shoulder stretch (R,L) item appears to be highly reliable since a's ranging from .83 to .94, P's from .83 to .94, and k's from .67 to .87 were found between test 1 and test 2. Although these numbers suggest good reliability, no other reliability data are available in the literature for the shoulder stretch to confirm these findings.

In regard to the trunk lift, the results of the study indicated a P of .87 (See Table 5). Meaningful alpha and kappa values were unable to be obtained due to an uneven distribution among the cells. Although little information is available on the reliability of any version of trunk extension, a comparable study was done by Wear in 1963. This study reported that prone back extension (height unlimited) was found to be a highly reliable test of trunk flexibility (Wear, 1963).

The test-retest reliability coefficients associated with the sit and reach (R, L) in the present study were a= .96 and .95, the P values were .92 and .89, and the kappa values attained were .83 and .76. These results are consistent with the study done by Patterson, Wiksten, Ray, Flanders, and Sanphy (1996), who tested 88 nondisabled boys and girls on the back saver sit and reach. The study reported an R= .99 for both boys and girls on both sides of the body, which indicates that the back saver sit and reach test is highly reliable. Although very little specific data are available on the back saver sit and reach test variation, it is unlikely that the results will be very different from better established form of the sit-and-reach test (Morrow, Falls, & Kohl, 1994).

References

Patterson, P., Wiksten, D.L., Ray, L., Flanders, C., & Sanphy, D. (1996). The validity and reliability of the back saver sit and reach test in middle school girls and boys. Research Quarterly for Exercise and Science, 67, 448-451.

Plowman, S.A., & Corbin, C.B. (1994). Muscular strength, endurance, and flexibility. In Morrow, J.R., Falls, H.B., & Kohl, H.W. (Eds.) (1994). The Prudential FITNESSGRAM Technical Reference Manual. Dallas: The Cooper Institute for Aerobics Research.

Rikli, R.E., Petray, C., & Baumgartner, T.A. (1992). The reliability of distance run test for children in grades K-4. Research Quarterly for Exercise and Science, 63, 270-276.

Safrit, M.J., & Wood, T.M. (1995). Introduction to Measurement in Physical Education and Exercise Science (3rd edition). St. Louis, MO: Mosby-Year Book, Inc.

Wear, C.L. (1963). Relationship of flexibility measurements to length of body segments. Research Quarterly, 34, 234-238.

Appendix A

Raw Scores of Subjects on Selected Test Items

Sub #	Age	Trunk Lift		Shoulder Stretch (R)		Shoulder Stretch (L)	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
1	15	P	P	P	P	P	P
2	13	P	P	F	F	F	F
3	13	P	P	F	F	F	F
4	16	P	P	P	P	P	P
5	16	P	ABSENT	F	ABS	F	ABS
6	13	P	P	F	P	F	F
7	11	P	F	NT	P	NT	P
8	15	P	P	P	P	P	P
9	12	P	P	P	P	F	P
10	14	P	P	F	F	F	F
11	16	P	P	P	P	P	P
12	17	P	P	F	F	F	P
13	17	P	P	P	P	P	P
14	16	P	P	F	F	F	F
15	12	P	F	F	F	F	F
16	17	P	ABSENT	P	ABS	P	ABS
17	15	P	P	P	P	P	P
18	12	P	P	P	P	P	P
19	14	P	P	P	P	F	F
20	12	P	P	F	P	F	F
21	14	P	P	P	P	F	F

Sub #	Age	Trunk Lift		Shoulder Stretch (R)		Shoulder Stretch (L)	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
22	13	P	P	P	P	P	P
23	17	P	P	F	F	F	F
24	15	P	P	P	P	P	P
25	16	P	P	P	P	P	P
26	13	F	P	P	P	F	F
27	13	P	P	P	P	P	P
28	15	P	ABSENT	P	ABS	P	ABS
29	12	P	P	F	F	F	F
30	12	P	P	P	P	P	P
31	14	P	P	F	F	F	F
32	15	P	P	P	P	P	P
33	13	P	P	F	F	F	P
34	10	P	P	P	P	P	P
35	12	P	F	F	F	P	P
36	15	P	P	P	P	F	P
37	15	P	P	P	P	F	P
38	16	P	P	P	P	P	P
39	14	P	ABSENT	F	ABS	F	ABS
40	17	P	P	P	P	F	P

Appendix A

Raw Scores on Selected Test Items (Continued)

Sub #	Age	Back Saver Sit-and-Reach (R)		Back Saver Sit-and-Reach (L)	
		Test 1 (cm.)	Test 2 (cm.)	Test 1 (cm.)	Test 2 (cm.)
1	15	29	26	29	27
2	13	UNT -OW	UNT-OW	UNT-OW	UNT-OW
3	13	17	13	22	15
4	16	46	39	38	40
5	16	42	ABSENT	41	ABSENT
6	13	23	23	21	21
7	11	18	24	21	28
8	15	39	39	37	31
9	12	34	28	35	29
10	14	23	22	24	24
11	16	14	12	19	16
12	17	22	27	30	29
13	17	20	24	25	25
14	16	UNT OW	UNT-OW	UNT-OW	UNT-OW
15	12	27	23	22	20
16	17	21	ABSENT	20	ABSENT
17	15	28	24	32	26
18	12	23	22	22	25
19	14	26	29	28	30
20	12	18	14	19	16
21	14	27	25	25	26
22	13	29	30	29	28

23	17	31	33	31	31
24	15	23	29	26	24
25	16	43	UNT	29	UNT
26	13	37	34	38	37
27	13	15	17	25	14
28	15	35	ABSENT	32	ABSENT
29	12	37	42	37	34
30	12	27	32	28	37
31	14	18	17	19	15
32	15	26	25	33	34
33	13	25	20	20	15
34	10	25	27	30	29
35	12	28	27	26	26
36	15	41	43	45	46
37	15	19	15	17	18
38	16	44	45	49	48
39	14	41	ABSENT	40	ABSENT
40	17	44	41	44	40

Appendix B

Raw Score and Test-Retest Alpha Coefficient, Proportion of Agreement, and kappa Coefficient on the Shoulder Stretch (R) (N=35)

	Test 1	Test 2
N=35	P	P
	F	F
	F	F
	P	P
	F	P
	P	P
	P	P
	F	F
	P	P
	F	F
	P	P
	F	F
	F	F
	P	P
	P	P
	P	P
	F	P
	P	P
	P	P
	F	F
	P	P
	P	P
	P	P
	P	P
	F	F
	P	P
	F	F
	P	P
	F	F
	P	P
	F	F
	P	P
	P	P
	P	P
a=.94	P= .94	k= .87

Appendix B

Raw Score and Test-Retest Alpha Coefficient, Proportion of Agreement, and kappa Coefficient on the Shoulder Stretch (L) (N=35)

	Test 1	Test 2	
N=35	P	P	
	F	F	
	F	F	
	P	P	
	F	F	
	P	P	
	F	P	
	F	F	
	P	P	
	F	P	
	P	P	
	F	F	
	F	F	
	P	P	
	P	P	
	F	F	
	F	F	
	F	F	
	P	P	
	F	F	
	P	P	
	P	P	
	F	F	
	P	P	
	F	F	
	P	P	
	F	F	
	P	P	
	F	P	
	P	P	
	P	P	
	F	P	
	F	P	
	P	P	
	F	P	
a=.83		P= .83	k= .67

Appendix B

Raw Score and Test-Retest Alpha Coefficient, Proportion of Agreement, and kappa Coefficient on the Sit and Reach (R) (N=33)

	Test 1	Test 2
(N= 33)	29	26
	17	13
	46	39
	23	23
	18	24
	39	39
	34	28
	23	22
	14	12
	22	27
	20	24
	27	23
	28	24
	23	22
	26	29
	18	14
	27	25
	29	30
	31	33
	37	34
	15	17
	37	42
	27	32
	18	17
	26	25
	25	20
	25	27
	28	27
	41	43
	19	15
	44	45
	44	41
	a=.96	P= .92 k= .83

Raw Score and Test-Retest Alpha Coefficient, Proportion of Agreement, and kappa Coefficient on the Sit and Reach (L) (N=33)

	Test 1	Test 2
(N= 33)	29	27
	22	15
	38	40
	21	21
	21	28
	37	31
	35	29
	24	24
	19	16
	30	29
	25	25
	22	20
	32	26
	22	25
	28	30
	19	16
	25	26
	29	28
	31	31
	26	24
	38	37
	25	14
	37	34
	28	37
	19	15
	33	34
	20	15
	30	29
	26	26
	45	46
	17	18
	49	48
	44	40
	a= .95	P=.89 k= .76

MICHIGAN SPORTS CAMP STUDY
SPRING 1995
October 16, 1997

Purpose and Overview

During the spring of 1995, a study was conducted in conjunction with a sports camp, sponsored by the Michigan Association for Blind Athletics in Lansing and Kalamazoo, Michigan. The study was designed to 1) field-test selected test items on the Prudential FITNESSGRAM test, 2) field-test the Canadian Aerobic Fitness Test, and 3) field-test the 600-yard run/walk test. The study began on April 30 and ended on May 4, 1995.

Subjects for the study included 53 adolescents with blindness (27 males and 26 females between the ages of 10 and 17). Data were collected as a part of and during a sports camp program. Two groups were involved in the study. One group (N=19), consisting primarily of younger subjects, was tested in Lansing, Michigan. The second group (N=34), consisting primarily of older subjects, was tested in Kalamazoo, Michigan.

Table 1
Gender and Ages of Subjects

Age	Females	Males
10	4	5
11	2	3
12	2	3
13	5	4
14	9	3
15	1	4
16	3	4
17	0	1
Total	**26**	**27**

Methods and Procedures

Field-testing was conducted by the project director, project coordinator, and eight camp staff, with support from several volunteers. Data were collected under the overall direction of Project Target staff.

Several test items from different sources were administered to the subjects. Items from the Prudential FITNESSGRAM (Cooper Institute for Aerobic Research, 1992) that were administered

included age, height, weight, triceps and calf skinfold, curl-up, push-up, pull-up, flexed arm hang, trunk lift, shoulder stretch, back saver and the Progressive Aerobic Cardiovascular Endurance Run (PACER). The only modification in the test items for the sample was that the PACER could be run with a sighted partner. From the scores attained in these test items, maximum oxygen intake, percent body fat, and body mass index could be calculated.

The Canadian Aerobic Fitness Test (CAFT) developed by the Ministers of Fitness and Amateur Sport in Canada (1986) was also administered to subjects, ages 15-17. The CAFT consists of a series of stepping sequences performed on 20.3 cm (8 inch steps to a six-count cadence which increases in tempo with each stepping sequence). In the test the participant performs up to three successive three-minute exercises. A 10-second post-exercise heart rate is taken following the completion of each stepping sequence. From age, weight, and post-exercise heart rate, oxygen uptake (VO_2) may be calculated.

The final item which was administered to all subjects was the 600-yard run/walk. In this test, subjects run a 600-yard distance, if possible. Walking is permitted if necessary, but the purpose is to complete the distance in the fastest possible time. Subjects could run alone or with a sighted partner.

Not all subjects took every item. The number and ages of subjects performing test items is presented in tables presented in this manuscript.

Results

The results in this study are presented within the following categories: aerobic capacity/endurance, body composition and musculoskeletal functioning. In analyzing data, emphasis was given to percent of subjects reaching Prudential FITNESSGRAM standards or other appropriate reference data representative of sighted peers. In addition the appropriateness of test items was subjectively evaluated in regard to procedural feasibility, i.e., ability of subjects to learn test items and perform them appropriately. Results of the study for this analysis appear on Tables 3, 4, 5, and 6.

Aerobic Capacity/Endurance

The Canadian Aerobic Fitness Test (CAFT) was used to measure the maximum oxygen intake (VO_{2max}) of 9 males and 7 females, ages 15-17. The sample of subjects appeared able to perform the test correctly following a brief period of time designed to learn how to perform. Although the sample size was small, 7 of 9 males (77.8%) met Prudential FITNESSGRAM standards for VO_{2max} recommended for nondisabled peers. None of the 4 female subjects reached these standards. Mean values for males but not females were within the Prudential FITNESSGRAM healthy fitness zone for VO_{2max}.

The PACER test item was administered to 39 subjects in the 10-14 age group. Subjects were able to learn and perform the test item after a very brief explanation and trial run. Only 11 of 39 subjects (28.2%) were able to reach Prudential FITNESSGRAM standards. A total of 10 of 22

(45.5%) females but only 1 of 17 (5.9%) males attained the minimum Prudential FITNESSGRAM standard although the average number of laps attained by males (12.6) slightly exceeded that of females (11.1).

The 600-yard run/walk was administered to 52 subjects in the 10-17 age group. Subjects were able to learn and perform the test item with a brief explanation. The average scores of both male and female subjects improved with age with the greatest improvement found in boys. The average group scores of younger males and females were similar (see Table 3). The results of this sample were compared with 600-yard run/walk normative data presented on the AAHPERD Youth Fitness Test (1976). A total of 19 of 52 subjects reached or exceeded the 20th percentile and a total of 25 of 52 subjects (48.1%) reached or exceeded the 10th percentile of normally sighted peers on the AAHPERD test. Interesting, more than half of the older subjects (ages 15-17) reached these standards (see Tables 6 and 7). When individual results were analyzed it was clear that "fit" subjects not only met the standards but often far exceeded the standards and "unfit" subjects often did not only not meet but did not come close to meeting the 10th or 20th percentile standards.

Body Composition

In this study percent body fat was estimated from triceps and calf skinfold and body mass index (BMI) was based upon the height and weight of each subject. Calculations were made using procedures specified by the Prudential FITNESSGRAM program. BMI was calculated for subjects throughout the 10-17 age range. Since skinfold tests were only administered to the 10-14 age range, percent body fat is only available for that age range. A total of 64.7% (11 of 17) of subjects who were administered the skinfold tests scored within the Prudential FITNESSGRAM healthy fitness zone. A slightly higher percentage of females (71.4%) than males (60.0%) met the recommended standards. The average group scores of both males and females were within the recommended health fitness zone for each gender.

In regard to BMI, a total of 68.6% (35 of 51) subjects met the Prudential FITNESSGRAM standard with a higher percentage of males (70.4%) than females (66.7%) meeting standards. The average scores of both males and females were generally well within recommended health fitness zones.

Musculoskeletal Functioning

Three items measuring strength or endurance of the upper extremity included pull-ups, the flexed arm hang, and push-ups. None of the 16 subjects, ages 10-14, were able to perform a single pull-up. In regard to the flexed arm hang, the group average for subjects was 1.4 seconds with only 21.0% of the group attaining minimum Prudential FITNESSGRAM standards. In the push-up test, results, were more positive, however, only 36.7% of the group reached the minimal healthy fitness zone value. In regard to these items, it was clear that the subjects need much more conditioning to reach the standard but youngsters may undoubtedly meet minimal standards.

Table 2

Mean and Standard Deviation Results for Adolescents with Visual Impairments - Michigan Study[1]

Test Item	Females	Males	Total	Ages
Max VO_2	32.5 (.9) N=4	45.6 (6.6) N=9	41.6 (8.3) N=13	15-17
PACER (laps)	11.1 (5.4) N=22	12.6 (9.6) N=17	11.8 (7.5) N=39	10-14
600-yard run/walk (sec.)	225 (58) N=26	198 (72) N=26	211 (66) N=52	10-17
Percent body fat (%)	18.9 (5.2) N=7	18.1 (7.4) N=10	18.4 (6.4) N=17	10-14
Body mass index (BMI)	22.4 (4.8) N=24	22.7 (4.3) N=27	22.5 (4.5) N=51	10-17
Curl-up (no.)	21.6 (19.6) N=26	35.3 (23.3) N=27	28.5 (22.5) N=53	10-17
Push-up (no.)	6.8 (7.4) N=24	14.1 (14.2) N=25	10.5 (11.8) N=49	10-17
Pull-ups (no.)	0 (0) N=9	0 (0) N=7	0 (0) N=16	10-14
Flexed arm hang (sec.)	1.3 (2.2) N=9	1.5 (2.3) N=10	1.4 (2.2) N=19	10-14
Trunk lift (inches)	10.8 (1.8) N=26	10.9 (2.5) N=27	10.8 (2.2) N=53	10-17
Back saver sit and reach (R) (cm)	12.0 (3.2) N=25	11.7 (2.6) N=27	11.9 (2.9) N=52	10-17
Back saver sit and reach (L) (cm)	11.7 (3.4) N=25	12.1 (2.4) N=27	11.9 (2.9) N=52	10-17

[1] Standard deviations are within parenthesis ( ).

Table 3
Test Item Results Distinguished by 10-14 and 15-17 Age Groups

Test Item and No.	Females	Males	Ages
600-yard run/walk (sec.) N=52	226 (22)	229 (17)	10-14
	220 (4)	138 (9)	15-17
Body Mass Index (BMI) N=51	22.0 (20)	24.4 (4)	10-14
	22.1 (18)	23.7 (9)	15-17
Curl-up N=53	18.0 (22)	25.9 (18)	10-14
	41.5 (4)	53.9 (9)	15-17
Push-up N=49	7.3 (20)	10.9 (17)	10-14
	4.3 (4)	20.9 (8)	15-17
Trunk Lift N=53	10.7 (22)	10.7 (18)	10-14
	11.3 (4)	11.2 (9)	15-17
Back Saver (R) N=52	12.1 (21)	11.2 (18)	10-14
	11.4 (4)	12.7 (9)	15-17
Back Saver (L) N=52	11.8 (21)	11.9 (18)	10-14
	11.4 (4)	12.6 (9)	15-17

Table 4

Percent of Adolescents with Visual Impairments
Passing Prudential FITNESSGRAM Standards

Test Items	Ages	No. Tested		Total Tested	Males Passed	Females Passed	Total Passed	% Males Passed	% Females Passed	% Total Passed
		Males	Females							
Max VO^2	15-17	9	4	13	7	0	7	77.8	0	53.8
PACER	10-14	17	22	39	1	10	11	5.9	45.5	28.2
Percent body fat	10-14	10	7	17	6	5	11	60.0	71.4	64.7
Body mass index	10-17	27	24	51	19	16	35	70.4	66.7	68.6
Curl-up	10-17	27	26	53	19	14	33	70.4	53.8	62.3
Push-ups	10-17	25	24	49	9	9	18	36.0	37.5	36.7
Pull-ups	10-14	7	9	16	0	0	0	0.0	0.0	0.0
Flexed arm hang	10-14	10	9	19	2	2	4	20.0	22.2	21.0
Trunk lift	10-17	27	26	53	24	23	47	88.9	88.5	88.7
Back saver (R)	10-17	27	25	52	24	20	44	88.9	80.0	84.6
Back saver (L)	10-17	27	25	52	25	18	43	92.6	72.0	82.7
Shoulder stretch (R)	10-17	27	26	53	20	17	37	74.1	65.4	69.8
Shoulder stretch (L)	10-17	27	26	53	13	17	30	48.1	65.4	56.6

Table 5
Percent of Adolescents with Visual Impairments
Passing Prudential FITNESSGRAM Standards
X Age Groups

Ages 15-17

Test Items	Total Tested	No. Tested		Total Passed	Males Passed	Females Passed	% Males Passed	% Females Passed	% Total Passed
		Males	Females						
Max VO_2	13	9	4	7	7	0	77.8	0	53.8
Body Mass Index	13	9	4	10	8	2	88.9	50.0	76.9
Curl-up	13	9	4	13	9	4	100.0	100.0	100.0
Push-up	12	8	4	4	3	1	37.5	25.0	33.0
Trunk Lift	13	9	4	13	9	4	100.0	100.0	100.0
Back Saver (R)	13	9	4	10	8	2	80.0	50.0	76.9
Back Saver (L)	13	9	4	10	8	2	88.9	50.0	76.9
Shoulder Stretch (R)	13	9	4	11	7	4	77.8	100.0	84.6
Shoulder Stretch (L)	13	9	4	9	5	4	55.6	100.0	69.2

Ages 10-14

PACER	39	17	22	26	16	10	94.1	45.5	66.7
Percent Body Fat	17	10	7	11	6	5	60.0	71.4	64.7
Body Mass Index	38	18	20	25	11	14	61.1	70.0	65.8
Pull-ups	16	7	9	0	0	0	00.0	00.0	00.0
Flexed Arm Hang	19	10	9	4	2	2	20.0	22.2	21.0
Curl-up	40	18	22	20	10	10	55.6	45.5	50.0
Push-up	37	17	20	14	6	8	35.3	40.0	37.8
Trunk Lift	40	18	22	34	15	19	83.3	86.4	85.0
Back Saver (R)	39	18	21	34	16	18	88.9	85.7	87.1
Back Saver (L)	39	18	21	33	17	16	94.4	76.2	84.6
Shoulder Stretch (R)	40	18	22	26	13	13	72.2	59.1	65.0
Shoulder Stretch (L)	40	18	22	21	8	13	44.4	59.1	52.5

Abdominal strength was reflected by scores on the curl-up test in this study. Of the 53 subjects performing this test item 33 (62.3%) attained scores in the Prudential FITNESSGRAM healthy fitness zone. Results appeared to be influenced by age and opportunity since 100% of 13 subjects, age 15-17, met the Prudential FITNESSGRAM healthy fitness zone values.

The trunk lift, back saver sit and reach, and the shoulder stretch were administered to measure flexibility/range of motion. In regard to the trunk lift and back saver, subjects performed relatively well when compared with the Prudential FITNESSGRAM standards. The total sample passing ranged from 82 to 89%. The subgroup with the lowest passing rates were females ages 15-17. A total of only 2 of 4 passed (50%). The discrepancy from other results may possibly be explained by the relatively small number of subjects in this subgroup.

The shoulder stretch was the final test item administered. Results indicate that 37 of 53 (69.8%) passed the right shoulder stretch and that 30 of 53 (56.6%) passed the left shoulder stretch. Based on percent passing, older subjects were more successful than the younger subjects.

Intercorrelations

A review of intercorrelations in this sample supports the contention that different test items are measuring different components of physical fitness.

In regard to aerobic functioning, a coefficient of correlation of r= .72 was found between VO_{2max} and the 600-yd. run/walk. A coefficient of .75 was found between the 600-yd. run/walk and laps completed on the PACER.

Summary and Conclusions

The results in this study support the use of the following test items on a criterion-referenced health-related test of physical fitness for adolescents with visual impairments: maximum VO_{2max} (ages 15-17), percent body fat, body mass index, trunk lift, back saver sit and reach, and the shoulder stretch. For each of these test items at least 50% of individuals who are blind met or exceeded the Prudential FITNESSGRAM standards. Also, each test item could be learned and performed by this sample in a reasonable amount of time.

This sample performed relatively poorly on the following test items: push-ups, pull-ups, and flexed arm hang. The investigators attribute this simply to lack of physical fitness rather than to any condition inherit in blindness.

Although the performance of subjects in the PACER test was relatively poor, the PACER was learned quickly by the subjects. The poor results are attributed to the poor physical fitness of the sample. Only youngsters aged 10-14 took this test item.

The 600-yd. run/walk is believed to be an acceptable test item for adolescents, ages 10-17.

The CAFT is believed to be an acceptable item for youngsters, ages 15-17. Subjects were able to learn and perform these test items with minimal modifications after a brief explanation. Based on the observation of the subjects performing in this study, a one-mile run/walk would not be a recommended test item in a health-related physical fitness test for younger subjects unless participants could run the entire distance. It may be most advisable to recommend the 600-yd. run/walk for younger subjects and the one-mile run for older subjects.

If the 600-yd. run/walk is used as a test of aerobic functioning, it would be necessary to identify an acceptable standard of performance. Tables 6 and 7 provide information on the number and percent of adolescents who are blind reaching the 20th and 10th percentile on the 600-yd. run/walk on the AAHPERD Youth Fitness Test (1976). When data representing groups are analyzed, it appears that a standard even at the 10th percentile would be challenging for adolescents who are blind. The data suggest that the percentage of subjects reaching the 10th percentile (48.1%) or the 20th percentile (36.5%) is not dramatically different (See Tables 6 and 7). From a total of 52 subjects only 6 of 33 subjects who failed meeting the 20th percentile standard met the 10th percentile standard. At ages 15-17, only 1 of 6 subjects not meeting the 20th percentile standard met the 10th percentile standard. A total of 8 of 13 or 61.5% met the 10th percentile and 7 of 13 or 53.8% met the 20th percentile standard. The data suggest that (1) "fit" subjects will meet standards and "unfit" subjects will not meet and often do not come close to meeting the 10th or 20th percentile standard, and (2) that the performance gap between nondisabled youngsters and adolescents who are blind decreases with age. In view if these observations, serious consideration should be given to using the 10th percentile as the criterion for youngsters ages 10-14 and the 20th percentile as the criterion for youngsters, ages 15-17.

Table 6

Number and Percent of Adolescents Who are Blind
Reaching the 20th Percentile on the
600-yd. Run/Walk on the AAHPERD Youth Fitness Test (1976)

Number by Group	No. of males	No. of males reaching standard	% of males reaching standard	No. of females	No. of females reaching standards	% of females reaching standards	Total meeting standards	Total % reaching standards
Total Group N=52	26	8	30.8	26	11	42.3	19	36.5
Ages 15-17 N=13	9	5	55.6	4	2	50.0	7	53.8
Ages 10-14 N=39	17	3	17.6	22	9	40.9	12	30.8

Table 7

Number and Percent of Adolescents who are Blind
Reaching the 10th Percentile on the 600-yd.
Run/Walk on the AAHPERD Youth Fitness Test (1976)

Number by Group	No. of males	No. of males reaching standard	% of males reaching standard	No. of females	No. of females reaching standards	% of females reaching standards	Total No. meeting standards	Total % reaching standards
Total Group N=52	26	11	42.3	26	14	53.8	25	48.1
Ages 15-17 N=13	9	6	66.7	4	2	50.0	8	61.5
Ages 10-14 N=39	17	5	29.4	22	12	54.5	17	43.6

Disk: Project Target#2
File: vimichig.std

NEW YORK STATE GAMES
FOR THE PHYSICALLY CHALLENGED

Brockport, New York
October 1994

Purpose and Overview

The New York State Games for the Physically Challenged (NYSGPC) is a state-supported program that provides athletic competition for children and adolescents with a variety of physical impairments. Each October regional competition is conducted on the campus of SUNY Brockport. On October 14-15, 1994 a Project Target testing team administered a number of test items to NYSGPC youngsters with cerebral palsy (CP) and spinal cord injuries (SCI). The primary purpose of the field test was to gain some insight into the feasibility of some of the tentative test items which had been identified as appropriate for youngsters with CP or SCI. Although the administration of test items was somewhat dependent upon a subject's sport classification, the items included for youngsters with CP were as follows: elbow range of motion, shoulder range of motion, knee range of motion, Apley test, shoulder external rotation test, shoulder stretch, trunk lift, back saver sit and reach, seated push-ups, and dumbbell press. Test items for youngsters with SCI included: trunk rotation, elbow range of motion, shoulder range of motion, shoulder stretch, seated push-ups, dumbbell press, and bench press. All tests were administered in accord with the procedures described in the October 1994 Project Target Test Manual.

A total of 24 youngsters were tested. Eighteen of the subjects had cerebral palsy and six had spinal cord injuries. A gender and age breakdown of the subjects tested is presented in Table 1.

Table 1
Age, gender, and disability breakdown of NYSGPC subjects.

Classification		Age 10-12	13-14	15-17
SCI	Males	0	0	1
	Females	2	3	0
CP	Males	1	3	5
	Females	3	4	2

In addition to age and gender, each subject was also classified according to NYSGPC sport classifications. Youngsters with CP were assigned to one of eight classes (C1-C8) consistent with the USCPAA classification system and youngsters with SCI were assigned to one of two classes (J2-J3) consistent with the NWAA classification system for juniors.

Observations and Results

Each of the test items administered is discussed below. Some items were administered to both groups of subjects and other items to only one group of subjects.

Range of Motion Tests (Goniometry)

Elbow, shoulder, and knee ROM tests were administered to CP subjects. SCI subjects were tested on elbow and shoulder ROM only. At the time of the testing the following observations were made:

- Testers need to make sure that the trunk does not rotate when assessing shoulder extension.
- The subject's arm should be relaxed and at the side at the start of shoulder extension and he/she should move the arm straight back when being tested.
- Testers can manipulate the arm so that subjects get the "feel" of the movement. When the shoulder is extended, it should not be abducted as well; if wheels, arm supports, or back supports restrict this movement the subject should be re-positioned.
- Testers should encourage subjects to "sit up straight" prior to assessing shoulder function to facilitate aligning the goniometer with the midaxillary line of the trunk.
- Some youngsters cannot supinate the forearm for elbow testing, but the joint angle can still be measured. The inability to supinate should be noted under "comments."
- Knee ROM might best be measured on a mat (rather than in a wheelchair).

Results of the ROM tests by sport classification are given in Table 2.

Table 2
Mean ROM scores (in degrees) for CP and SCI subjects.

	Sport Classifications							
	C1	**C3**	**C4**	**C5**	**C6**	**C7**	**J2**	**J3**
Right Elbow ROM (optimum ROM = 150)	125.5 (n=4)	117.0 (n=2)	121.0 (n=1)	130.0 (n=5)	134.3 (n=3)	115.5 (n=2)	115.3 (n=3)	147.0 (n=1)
Left Elbow ROM (optimum ROM = 150)	110.8 (n=4)	117.0 (n=2)	123.0 (n=1)	129.0 (n=5)	124.3 (n=3)	133.0 (n=1)	110.0 (n=3)	154.0 (n=1)
Right Shoulder ROM (optimum ROM = 230	184.0 (n=3)	201.5 (n=2)	198.0 (n=1)	215.8 (n=5)	181.0 (n=3)	239.5 (n=2)	233.7 (n=3)	281.0 (n=1)
Left Shoulder ROM (optimum ROM = 230)	177.0 (n=3)	207.0 (n=2)	204.0 (n=1)	216.0 (n=5)	171.0 (n=3)	245.0 (n=1)	232.3 (n=3)	270.0 (n=1)
Right Knee ROM (optimum ROM = 135)	56.3 (n=4)	74.5 (n=2)	48.0 (n=1)	85.4 (n=5)	104.7 (n=3)	113.5 (n=2)	-	-
Left Knee ROM (optimum ROM = 135)	68.0 (n=4)	75.5 (n=2)	67.0 (n=1)	88.8 (n=5)	100.0 (n=3)	108.5 (n=2)	-	-

Although the number of subjects in each class is limited, the data in Table 2 support the popular observation that youngsters with CP (classes C1 - C7) have restricted range of motion. None of the elbow or knee ROM means are within 90% of the optimum ROM values. The shoulder scores are better with five of the 12 mean scores achieving at least 90% of the optimum values. The data for most of the subjects with SCI are actually quite comparable to the optimum values. This is disguised somewhat in the table, however, because one of the three J2 youngsters recorded scores quite a bit lower than the other two which served to suppress (perhaps artificially) the mean. The shoulder ROM scores of the J3 subject are interesting because they exceed the optimum value. This might raise questions with regard to the stability of the shoulder joint and is an issue that should be considered in establishing standards for SCI youngsters.

At this time it appears that it would be appropriate to adopt general health related standards (GHRS) for the SCI group, although more subjects should be tested. The preliminary data here suggests that it may not be necessary to also present disability health related standards (DHRS) as an optional approach. For youngsters with CP, however, it appears that they will have a great deal of difficulty achieving GHRS on ROM tests and, consequently, it would be desirable to provide DHRS which reflect reasonable expectations. A basis for DHRS might be found in an earlier version of the CP sport classification system. For each of the eight classes in the 1984 NASCP/CP-ISRA system an approximate range of motion was provided expressed in percentage form. The percent values for each class in the 1984 system are provided in Table 3.

Also provided in Table 3 are the percent-adjusted optimum ROM values in degrees for each of the joints under consideration.

Table 3
Approximate Percent ROMs by CP Class and Corresponding Adjusted ROM Values for the Three ROM Tests.

	CP Sport Classes							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Percent ROM	25%	40%	60%	70%	80%	70%	90%	100%
Elbow (in degrees)	38	60	90	105	120	105	135	150
Shoulder (in degrees)	58	92	138	161	184	161	207	230
Knee (in degrees)	34	54	81	95	108	95	122	135

Using the adjusted ROM values as the DHRS for the CP subjects results in the pass/fail rates shown in Table 4.

Table 4
Pass/Fail Rates of CP Subjects Using Proposed DHRS.

		Sport Classifications						
		C1	C3	C4	C5	C6	C7	Total
Right Elbow	**Pass**	4	2	1	4	3	0	14
	Fail	0	0	0	1	0	2	3
Left Elbow	**Pass**	4	2	1	4	3	0	14
	Fail	0	0	0	1	0	2	3
Right Shoulder	**Pass**	3	2	1	4	2	2	14
	Fail	0	0	0	1	1	0	2
Left Shoulder	**Pass**	3	2	1	5	2	2	15
	Fail	0	0	0	0	1	0	1
Right Knee	**Pass**	4	1	0	1	2	0	8
	Fail	0	1	1	4	1	2	9
Left Knee	**Pass**	4	1	0	1	2	1	9
	Fail	0	1	1	4	1	1	8

Subjects with CP compared very favorably to the proposed DHRS on elbow and shoulder ROM. Approximately 86% of all arm measures would result in "pass" judgments. The proposed DHRS for the knee was a more difficult criterion, especially for class 5. Still, approximately 50% of the NYSGPC CP subjects met the standard.

External Rotation Test

The external rotation test was administered only to subjects with cerebral palsy. The subjective nature of this test was noted during field testing as was the necessity, in some cases, of "anchoring" a subject's elbow against the side during the test. Pass/fail rates are given in Table 5.

Table 5
Pass/Fail Rates for External Rotation Test

		Sport Classification						
		C1	C3	C4	C5	C6	C7	Total
Right Arm	**Pass**	1	2	1	4	3	3	14
	Fail	3	0	0	0	0	0	3
Left Arm	**Pass**	1	2	1	4	3	2	13
	Fail	3	0	0	0	0	1	4

The data in Table 5 suggest that this is an easy test for youngsters in classes 3 and higher. The only failure among classes C3-C7 was recorded by a C7 with hemiplegia affecting the left side. Despite concerns about its subjective nature, this test could still have some utility for classes 1 and 2, but more testing is necessary.

Shoulder Stretch

The shoulder stretch was administered to both SCI and CP subjects. Of the five SCI subjects who were tested on this item, only one could pass this test (she did so with both arms). None of the remaining four subjects could pass with either arm. Of the 18 CP subjects tested, only one subject (a class 5) could pass the test using either arm (he passed on both arms). The performance of the CP subjects was anticipated by project staff and the Apley test was included as an optional test for that group. The difficulty of the item for the SCI group, however, was somewhat unanticipated. It will have to be determined whether this ability is critical for this group (so that the item will be retained despite low preliminary results) or whether a modified item (such as the Apley) could be included as an option.

Apley Test

Only youngsters with CP took the Apley. Success rates are given in Table 6.

Table 6
Pass/Fail Rates on the Apley Test.

		Sport Classification						
		C1	C3	C4	C5	C6	C7	Total
Right Arm	**Pass**	1	0	1	4	2	2	10
	Fail	4	2	0	1	1	0	8
Left Arm	**Pass**	0	1	1	5	1	1	9
	Fail	5	1	0	0	2	1	9

At least based upon pass/fail rates, the Apley test appears to be superior to the shoulder stretch as a practical measure of shoulder function, especially for classes 4 and beyond. Performance on the Apley appears to be highly related to shoulder ROM. Of the 25 shoulders that had at least 170 degrees of ROM, 19 passed the Apley. Of the six shoulders that had less than 170 degrees of ROM, all failed the Apley.

Trunk Lift

NYSGPC field test procedures called for the trunk lift to be administered only to subjects with CP in classes 4 through 8. Results indicated that this item might be reasonable for classes 6-8 (three out of five subjects in classes 6 and 7 passed the test), but is probably too difficult for classes 4 and 5 (only one out of six subjects was able to lift the chin 12 inches from the mat). It is somewhat interesting to note, however, that if the criterion for success was lowered to 8 inches, nine of the 11 subjects in classes 4 through 7 would have passed the test.

Back Saver Sit and Reach

Seven subjects with CP (classes 5-7) took the sit and reach test. The mean score with right leg extended was 19.1 cm and 13.4 cm with left leg extended. Two subjects met FITNESSGRAM standards with the right leg extended and two achieved the standard with the left leg extended; only one of the seven subjects met the standard for both legs. Even if standards were lowered to improve the success rate for youngsters with CP (i.e., DHRS), there would still be a question with regard to the appropriateness of this item. Most notably many youngsters with CP have difficulty extending their legs completely when attempting this activity.

The inability to extend contaminates the results because the hands are closer to the box with bent knees and there is no objective way to adjust scores based on the degree of knee flexion that might be present.

Seated Push-up

Four youngsters with CP and six with SCI attempted the seated push-up. Subjects from the C3 class had difficulty with this item. It appeared to the testers that this difficulty was due in part to an inability to completely extend the elbows which placed greater strain on the muscles. Other classes had more success although that success was not universal. Using 10 seconds as a criterion score, two subjects with CP passed and four with SCI passed. Three of the six youngsters with SCI were able to hold the position for 50 seconds which was the maximum score during the field test.

The seated push-up appears to be a good "functional" item for subjects who use wheelchairs and also has some health ramifications related to decubitus ulcers. Still there are a few problems with this item. First, success on this item is a function of body weight. As with pull-ups or flexed arm hang, a heavier youngster will find this item more difficult. Second, youngsters who cannot completely extend their elbows, as noted above, will be at a mechanical disadvantage when performing this item. Third, and last, not all wheelchairs have arm rests, consequently it may be necessary to transfer youngsters prior to testing them.

Dominant Dumbbell Press

Sixteen subjects (10 CP, 6 SCI) took the dumbbell press and performance was quite variable. The mean number of repetitions was about 10 with a group standard deviation of almost 12. Using five repetitions as an arbitrary criterion also speaks to the variability of performance on this test. Among classes C5-C7, five youngsters passed the five repetition criterion and five failed. Among classes J2-J3, three youngsters passed the criterion and three failed. The mean age of the subjects who passed the test was 14.6 and the mean age of youngsters who failed the item was 13.4. It may be that a lighter weight would be more appropriate for the younger subjects, perhaps 10 lbs. for 10-12 year-olds and 15 lbs. for 13-17 year olds.

Bench Press

Five subjects with SCI took the test. As with the dumbbell press, performance was variable. With a maximum number of repetitions set at 30, the scores for the five youngsters were 30, 30, 0, 0, 15. Interestingly, but not necessarily surprisingly, the bench press and the

dumbbell press were highly related. A Pearson r of .87 existed between these two items based on the performance of five subjects. Consequently one consideration would be to select only one of these items for the final battery. If so, preference might be given to the dumbbell press because it is more easily administered (and spotted) and it eliminates problems caused by hemiplegia.

Trunk Rotation

The trunk rotation test was administered only to J3 subjects. Of the three J3 youngsters, two passed the test and the other failed. One of the subjects who passed did not use a wheelchair for ADLs and was tested in a regular chair. (The test was also given to two J2 subjects; one passed, one failed.)

Date: 09/19/97
File: nysgpc.rep
Disk: FX Short

PROJECT TARGET

NEW YORK STATE GAMES for the PHYSICALLY CHALLENGED

OCTOBER 13-14 1995 SUNY BROCKPORT

Purpose and Overview

For the second time, the Project Target staff collected data on young athletes who participated in the New York State Games for the Physically Challenged. There were five primary objectives for the 1995 testing: 1) field test the aerobic movement test (AMT); 2) investigate the inter-rater reliability of the active range of motion test items (AROM); 3) investigate the concurrent validity of one of the AROM items (elbow extension) with measures of goniometry; 4) identify possible relationships between AROM scores and measures of activities of daily living (ADL); and 5) identify possible relationships between muscular strength and endurance scores and ADL measures.

Subjects

Twenty-five subjects were tested on at least one test item over the course of the weekend. Eleven female subjects ranged in age from 10 to 18 with an average age of 13.4 years. The 14 male subjects were in the 10 to 17 age range with a mean age of 12.7 years. Five of the subjects had a spinal cord injury (SCI); 11 had cerebral palsy (CP); eight were classified as les autres (LA); and one had a congenital anomaly or amputation (CA/A).

Methods and Procedures

All data were collected by graduate students associated with the adapted physical education program at SUNY Brockport. Each student who served as a tester had been trained to administer the item or items for which he or she was responsible. Depending on classification subjects took the following test items: AMT, seated push-up, dumbbell press (15 lb.), grip strength, shoulder stretch or Apley test, elbow extension (goniometry), and AROM. In addition SCI subjects were scheduled to take the trunk rotation test and CP subjects were scheduled to take the Thomas test. Test protocols as described in the October 1995 Project Target test manual were followed during test administration. Throughout the testing top priority was given to the AMT, the AROM, and elbow extension. As many subjects as possible were scheduled to take these three items. Additional items were administered as time permitted. Operative guidelines for the AROM test items are attached to this report.

In addition to the test items listed above, parents, teachers, or other adults familiar with the subjects, were asked to complete a questionnaire designed to evaluate each subject's level of independence on ADLs. The questionnaire was based on the Functional Independence Measure and asked respondents to rate the following ADLs: eating, dressing upper body, dressing lower body, toileting, transfers (bed, chair, wheelchair), transfers (toilet), walking, wheelchair propulsion, and stair climbing. Criteria for level of independence was provided and ranged from "total assistance" (a score of 1) to "complete independence" (a score of 7). A copy of the "Project Target Functional Independence Survey" is attached to this report.

Results and Recommendations

This section of the report is sub-divided by test items. The first section will discuss the field test results of the AMT. The second section will pertain to the AROM tests. Specifically, inter-rater reliability, concurrent validity, and relationship to ADLs will be addressed in the AROM section. Finally, relationships found between measures of muscular strength and endurance and ADLs will be discussed in the final section.

Aerobic Movement Test

Twenty of the 25 subjects attempted the AMT. Essentially the AMT requires testees to exercise within their target heart rate zone (THRZ) for 15 minutes. In this version of the AMT subjects could either use an arm crank ergometer or run/walk as the exercise modality. Of the 20 subjects who took the test, 15 passed. The 15 successful subjects included two with SCI (HLP and LLP), seven with CP (C4 through C8), and six with LA conditions. Nine of the 15 subjects used arm ergometry to meet the test criteria, while the other six ran. Of the five subjects who could not meet the test criteria, four were unable to achieve their THRZ. Two of the four unsuccessful subjects were youngsters with Class 1 CP, a third was classified C7, and the fourth was a T4 SCI (HLP). The fifth unsuccessful subject (double above knee amputee) complained of dizziness a minute or two into the test and the test was terminated at that time.

Results of the AMT were encouraging. It appears that many disabled youngsters with a variety of conditions should be able to reach the test standards. No quadriplegic SCI youngsters were tested, however, and both Class 1 CP youngsters who attempted the test failed to reach the test criteria. A question remains, therefore, relative to the appropriateness of the test for the most severely disabled SCI and CP youngsters. Based on clinical observation of CP subjects who participated in Project UNIQUE, a mat crawl might be an exercise mode that could be suggested for Class 1 youngsters for the purpose of penetrating their THRZ.

Active Range of Motion Tests

Inter-rater reliability. Two testers independently scored subjects on four AROM test items as recommended in the October 1995 test manual. Scores ranged from 1 to 5 for each item based upon the visual criteria provided. Percent of agreement between the two testers broken down by AROM test item is given in Table 1.

TABLE 1. Percent of Agreement on AROM Scoring by Specific Test Item

AROM Item	Number of Identical Ratings	Total Number of Ratings	Percent Agreement
Wrist Extension	22	24	92
R	10	12	83
L	12	12	100
Elbow Extension	44	46	96
R	22	23	96
L	22	23	96
Forearm Supination	20	25	80
R	9	13	69
L	11	12	92
Shoulder Flexion	8	8	100
R	4	4	100
L	4	4	100
Shoulder Extension	13	15	87
R	7	8	88
L	6	7	86
Shoulder Abduction	8	8	100
R	4	4	100
L	4	4	100
Shoulder Horizontal Abd.	4	6	67
R	1	3	33
L	3	3	100
Shoulder External Rotation	16	22	73
R	7	11	64
L	9	11	82
Hip Extension	10	12	83
R	5	6	83
L	5	6	83
Knee Extension	21	28	75
R	11	14	79
L	10	14	71
Total	166	194	86

The data in Table 1 indicate that there was a fairly high rate of agreement between the two testers on most of the AROM items. There were three items, however, that fell below 80% agreement: shoulder horizontal abduction, shoulder external rotation, and knee extension. Both shoulder items are designed to be viewed "from above" the subject (i.e., looking down over the subject's head). These are the only two items scored from that perspective and it apparently caused some problem for at least one of the testers. Modifications in the scoring procedure, or possibly more specific instructions for scoring, may help to improve the objectivity of these tests. The low percent of agreement on the knee may be due to a problem in administering that item. It appears that it is possible for a youngster to completely extend the knee (which is the objective of the test), but not have the lower leg pointing at 3:00 (or 9:00) (which is one of the visual criteria used in the scoring). Positioning the subject on a hard desk or table so that the relaxed lower leg contacts the edge of the desk just below the knee joint should help to alleviate this problem.

Concurrent validity. To determine if the AROM approach had potential for accurately reflecting the range of motion in a joint, AROM scores were compared to goniometry values for elbow extension. Goniometry values were used to establish "objective" AROM scores and the testers values served as "subjective" AROM scores. (AROM scores range from 1 to 5 where each number represents 20% of the normal range of motion for that joint. A score of 5 indicates that the individual has between 80-100% normal range of motion in the joint. When elbow extension is assessed through goniometry, 180 degrees is considered to be the normal end point for that movement. By starting at 30 degrees and moving to 180 degrees, elbow extension can move the lower arm through a 150 degree arc. Twenty percent of that arc is 30 degrees. Goniometry, therefore, could be used to establish "objective" AROM scores by assigning a 5 for any measure between 150 and 180 degrees; a 4 for any measure between 120 and 150 degrees; etc.) Table 2 provides information on the percent of agreement between the testers "subjective" AROM scores and the "objective" AROM scores established through goniometry.

TABLE 2. Percent of Agreement Between Subjective AROM and Goniometry - determined AROM for Elbow Extension

	Number of Identical Ratings	Total Number of Ratings	Percent Agreement
Right Elbow Extension	39	42	93
Tester 1	20	21	95
Tester 2	19	21	91
Left Elbow Extension	37	40	93
Tester 1	18	20	90
Tester 2	19	20	95
Total	76	82	93
Tester 1	38	41	93
Tester 2	38	41	93

The data in Table 2 seem to indicate that testers, using AROM procedures, can subjectively classify the range of motion in the elbow joint with a fairly high degree of accuracy. One obvious limitation to this study is that only one joint motion was assessed. Future work should look at the ability of testers to accurately classify other joint movements. Still, the results of this investigation suggest that the AROM has good potential for accurately classifying (within 20%) range of motion.

Relationship to ADLs. One approach being considered for establishing standards for AROM test items as well as for the functional tests of flexibility (i.e., shoulder stretch, Apley test, Thomas test, trunk rotation) is to tie test scores into a measure or measures of functional independence. The Functional Independence Measure (FIM) evaluates activities of daily living on a 7-point scale. A score of 7 represents complete independence for that task while a score of 6 indicates that the individual can complete the task independently but needs an assistive device or additional time, or there may be safety or risk factors associated with completing the task. Scores of 1 through 5 require the presence (or assistance) of another individual to complete the task. Nine ADLs which were thought to have a possible relationship to fitness were selected from the FIM to comprise the Project Target Functional Independence Survey.

The first step in determining the relationship between AROM measures and ADLs was to run a Pearson correlation matrix including the 20 AROM measures and the nine ADLs. The Pearson r's generated in the matrix were generally low, but this was not surprising because the range of scores for both the AROM and ADL measures was very limited which served to suppress the Pearson r's. The ultimate goal in assessing the relationship between AROMs and ADLs was to generate contigency tables and phi coefficients for those relationships which seemed to hold the greatest promise. Arbitrarily, therefore, contigency tables and phi coefficients were constructed for AROM/ADL combinations which had either a Pearson r of .50 or greater, or a Pearson r that was significant at the .05 level as given in the correlation matrix. The contingency tables were

constructed in a 2x2 matrix. For the AROMs, a score of 5 served as one level and scores from 1 to 4 served as the other level. For the ADLs, scores of 6 and 7 on the FIM were placed in one category and scores 1 through 5 in the second category. In essence the contingency tables were designed to assess the ability of the AROM scores to accurately classify one as "independent" or "dependent" on certain ADLs. In this case, a functional range of motion of 80% of normal served as the criterion for classification.

Table 3 depicts the relationship between AROMs and ADLs as expressed via the phi coefficient. Where no coefficient is presented the relationship was not analyzed because it did not meet the arbitrary criteria established for the correlation matrix. A dash appears in situations where the Pearson r from the matrix met the criteria for contigency table analysis, but for which a phi could not be computed because of empty cells in the table.

TABLE 3. Phi Coefficients for AROM and ADL Measures

	Eat	Dress Upper	Dress Lower	Toilet Hygiene	Transfer Chair	Transfer Toilet	Walk	WheelChair	Stairs
R Elbow Ext				.31			-		
L Elbow Ext	.46	.69	.54	.35	.69	.68	-		
R Forearm Sup			.66						
L Forearm Sup				.66					
R Shld Ext	.75	.75	.75	.75	.75	.75			
R Ext Rot					-		-	1.00	.66
R Hip Ext					-		-	.50	
R Knee Ext							-	1.00	
L Knee Ext							-	1.00	.37

Table 4 presents phi coefficient information on the relationship between the flexibility tests and the ADLs. (Only one subject took both the trunk rotation test and ADL survey, so this combination was not analyzed.) Since each of the flexibility tests is scored dichotomously, those scores were simply classified as "pass" vs. "fail." The ADL scores were once again classified as "independent" vs. "dependent." As with the AROM/ADL analysis described previously, phi coefficients were calculated only for those flexibility test/ADL combinations that had Pearson r's of at least .50 or Pearson r's that were significant at the .05 level.

TABLE 4. Phi Coefficients for Flexibility Tests and ADL Measures

	Eat	Dress Upper	Dress Lower	Toilet Hygiene	Transfer Chair	Transfer Toilet	Walk	WheelChair	Stairs
R Shld St			-	-			-		
L Shld St			-	-			-		.50
L Apley	.82	.43	.63	.63					
R Thomas	.30	.65	.65	.30	.65	.65	-		-

The phi coefficients presented in Tables 3 and 4 are not "earth shaking;" for the most part, the relationships might be considered moderate or slightly better than moderate. At first glance these values might appear to be too small to have any predictive value in classifying subjects as "independent" or "dependent." These modest relationships, however, might be more a function of small sample size than in predictive power. As an example, consider the relationship between left elbow extension and wheelchair transfer. The phi value is only .69, but upon inspecting the contingency table one notes that of the 17 subjects who scored 5 on the AROM, 16 were classified as "independent" and the only subject who scored less than a 5 was classified as "dependent." So, 80% left elbow AROM accurately classified 17 out of 18 subjects (94%) as either "independent" or "dependent." The rather modest phi might be a function of low overall n or a function of the fact that three of the four cells in the contingency table contained one or zero frequencies. In discussing chi square (a statistic related to phi), Ferguson (1971) has written, "If an actual difference exists between observed and expected values, this difference will tend to increase as sample size increases," and when discussing phi, "N should, clearly, not be too small."

The phi values, therefore, might be considered encouraging as part of pilot work. Hopefully ROM/ADL relationships can be studied more extensively in the future. In addition to trying to generate larger phi's via larger samples, it might also be possible to increase phi by using different classification criteria, such as requiring a FIM score of 7 (not 6 or 7) for one to be considered "independent."

Muscular Strength/Endurance

Only four tests of muscular strength and endurance were administered to subjects in this investigation: seated push-up, bench press, dumbbell press, and grip strength. Only two subjects took the bench press so no analysis was performed on this item. The other three items were correlated with scores from the ADL survey to determine if this approach might have some potential for standard setting with the MS/E items. Correlations between the MS/Es and the ADLs, however, generally were low. Grip strength, in particular, showed little relationship to the ADLs; only one correlation coefficient even reached .20 for these combinations. Correlations between the dumbbell press and the ADLs generally ran in the .20 to .30 range, but none was significant at the .05 level. The seated push-up was the only MS/E item that showed a significant relationship with ADL values. A Pearson r of .58 (p=.05) was generated for seated push-up and dressing upper body and a 1.00

coefficient was calculated between seated push-up and wheelchair propulsion. (The coefficient between seated push-up and transfers to chair, bed, and wheelchair was .41, p=.134.)

In examining more closely the relationship between the seated push-up and the three ADLs discussed above (the transfer ADL was included as a matter of interest), three different criteria for the seated push-up were investigated: 5, 10, and 20 seconds. The number of subjects for each analysis ranged from six to nine, so conclusions are tentative at best. The best criterion, however, appeared to be 5 seconds. It accurately classified eight out of nine subjects as either "independent" or "dependent" for dressing upper body and transfers, and six out of six subjects for wheelchair propulsion. The other two criteria were only slightly less accurate with correct classifications for seven out of nine subjects on dressing and transfers, and the same six out of six for wheelchair propulsion. It may be that the 5-second criterion would be sufficient and could be based upon both the recommendation for pressure sore relief and functional independence for a few ADLs.

Summary

The results of the present investigation are summarized below.

1) The AMT is a promising item for adolescents with physical disabilities. Data collected here suggest that the standards for the test are attainable by youngsters with a variety of disabilities, although some additional investigation with more seriously disabled subjects seems necessary.

2) The inter-rater reliability of the AROM tests is generally good. Three items, however, should be looked at with an eye toward improving objectivity.

3) AROM procedures (at least for elbow extension) can be used by testers to accurately categorize a subject's range of motion into 20% increments. Such an ability might be important in distinguishing ROM for the purpose of assessing a criterion-referenced standard. For instance, if an ROM of 80% is deemed necessary for "functional independence," testers will be able to make that determination subjectively with a high degree of accuracy.

4) It appears that certain active range of motion tests have relationships with certain activities of daily living and that independence on these ADLs has the potential for providing standards for the AROM tests. It appears, for instance, that 80% ROM of certain movements can be used to distinguish independent from dependent subjects on certain ADLs.

5) Unlike the AROM tests, tests of muscular strength and endurance have little relationship to ADL independence. The exception to this finding is the seated push-up. The seated push-up had a good enough relationship with two or three of the ADLs to classify subjects as independent or dependent in a fairly accurate way.

Disk: Target(short)
File: nysgpc.95

Feasibility of the Target Aerobic Movement Test in Children and Adolescents With Spina Bifida

James H. Rimmer
Northern Illinois University

Fiona Connor-Kuntz
Cleveland Heights-University Heights City School District

Joseph P. Winnick and Francis X. Short
State University of New York College at Brockport

The purpose of this study was to determine the feasibility of the Target Aerobic Movement Test (TAMT)[1] in a group of children and adolescents with spina bifida ($n = 32$). Thirty-two children (11 subjects-thoracic lesion, 21 subjects-lumbar lesion) volunteered for the study. Results indicated there were no significant differences in the proportion of subjects who passed Test 1 or Test 2 ($p > .05$). Twenty-seven out of 28 eligible subjects (96%) on Test 1 and 25 of 27 eligible subjects (93%) on Test 2 met the criteria for successful completion of the TAMT. The TAMT appears to be a reliable and feasible test for measuring aerobic behavior in children and adolescents with spina bifida. Future research should focus on studying the feasibility of the TAMT with other populations with disabilities and to also determine if the test can become a more refined discriminator of aerobic behavior and aerobic capacity.

There is a growing body of evidence that developing good cardiorespiratory fitness early in life and maintaining it throughout adulthood can reduce the incidence of morbidity and mortality from all causes of disease, particularly heart disease (Consensus Development Conference, 1995; Pate et al., 1995). In a population-based study of children and adolescents ages 10 to 14 years, investigators found that a higher level of cardiorespiratory fitness was significantly related to

---

James H. Rimmer is with the Department of Physical Education at Northern Illinois University, DeKalb, IL 60115. Fiona Connor-Kuntz is with the Cleveland Heights-University Heights City School District, University Heights, OH 44118-3397. Joseph P. Winnick and Francis X. Short both are with the Department of Physical Education and Sport at the State University of New York College at Brockport, Brockport, NY 14420.

lower body weight, body fat, and systolic and diastolic blood pressure, and a higher high-density lipoprotein cholesterol (HDL-C) to total cholesterol ratio (Tell & Vellar, 1988). Much of the available data on the important benefits of cardiovascular fitness for children and adolescents led experts in the field to create physical fitness objectives for the United States, which are cited in the *Healthy People 2000* report (Public Health Service, 1991).

The promotion of cardiorespiratory fitness in children and adults with disabilities is also very important. An improvement in fitness can lead to a higher level of independence by reducing the need for assistance to perform activities of daily living (Santiago, Coyle, & Kinney, 1993). Janssen, Van Oers, Van Der Woude, and Hollander (1994) found that poor physical fitness levels resulted in much greater difficulty in persons with spinal cord injuries performing activities of daily living, particularly in performing transfers, entering or leaving a car, and negotiating environmental barriers. Noreau and Shephard (1992) noted that low levels of physical fitness among persons with a spinal cord injury made physical barriers very difficult to negotiate, and that persons who are wheelchair users needed to strengthen their shoulder muscles in order to facilitate ambulation.

Despite the overwhelming importance of maintaining good cardiorespiratory fitness in terms of overall health, there are relatively few field-based tests that can reliably and accurately measure cardiorespiratory fitness in children and adolescents with physical disabilities (Rimmer, 1994). In a recent paper by Rimmer, Braddock, and Pitetti (1996), it was recommended that researchers become more focused on developing accurate field-based measuring instruments for assessing cardiovascular fitness in persons with disabilities.

It is generally agreed that aerobic capacity is a preferred indicator of cardiorespiratory fitness because it reflects the ability to carry out prolonged strenuous exercise (American College of Sports Medicine, 1995). However, it is a measure that is very difficult and sometimes impossible to attain in certain individuals with disabilities (Rimmer, 1994). Part of this difficulty can be attributed to poor physiological compliance, low motivational levels, and a low threshold for fatigue, which prevent the individual from attaining a "true" measure of maximal aerobic capacity (Rimmer, 1994).

Recently, Winnick and Short (1995) have addressed this issue by developing a health-related cardiovascular fitness test for children and adolescents with disabilities. The investigators recommended that in children and adolescents with disabilities, a viable option would be to measure *aerobic behavior*. Aerobic behavior reflects the ability to perform an aerobic activity at a selected duration and intensity, but ordinarily does not require the individual to perform an all-out effort.

Using the guidelines established by the American College of Sports Medicine (1991, 1995)—which recommends that physical activity should elevate heart rate to between 60 and 90% of maximum heart rate for a period of 15 to 60 minutes, and should be performed 3 to 5 days a week in order to confer health benefits and improve or maintain cardiorespiratory fitness—Winnick and Short (1995) applied these guidelines to the development of a test for children and adolescents with disabilities.[2] The test is called the Target Aerobic Movement Test (TAMT) and is used to measure the *aerobic behavior* of children and adolescents with disabilities.

The TAMT purports to be appropriate for adolescents with a variety of disabilities, including those with spina bifida. No previous research, however, has

been conducted to determine the soundness of established procedures, the appropriateness of intensity and duration requirements, or the consistency of results for children and adolescents with spina bifida. The purpose of this study, therefore, was to determine the feasibility of the TAMT in a group of children and adolescents with spina bifida.

Method

Subjects

Thirty-two subjects with spina bifida myelomeningocele (17 females, 15 males; ages 10 to 18 years old) were recruited from a residential summer camp that was specifically designed for children and adolescents with this condition. The focus of the camp was on arts and crafts, sports and recreational and educational activities. Subjects attended the camp from the following Midwestern states: Illinois, Wisconsin, Indiana, and Michigan. Eleven subjects had a thoracic lesion and 21 subjects had a lumbar lesion. Twenty-eight subjects used a wheelchair routinely, and four subjects used a wheelchair for long distances and sports only. All the subjects were recruited on a volunteer basis. Parental consent was obtained during registration for the camp. Demographic data including birth date, weight, level of cognitive and motor function, and medication usage, were obtained from camp enrollment forms. Height and resting heart rate were measured with subjects lying supine during a 20- to 30-minute rest period. Descriptive characteristics of subjects appear in Table 1.

Target Aerobic Movement Test (TAMT)

Test Description. This test, originally described by Winnick and Short (1995), is designed to assess the ability of children and adolescents to exercise within a recommended target heart rate zone for a sustained period of time. Participants can engage in virtually any physical activity they choose as long as the activity is of sufficient intensity to raise the heart rate into the target heart rate zone for a period of 15 minutes.

Table 1 Descriptive Characteristics ($N = 32$)

Variable	*M*	*SD*	Min	Max
Age (yr)	13.5	2.5	10	18
Ht (cm)	141.8	15.1	106.7	177.8
Wt (kg)	45.8	13.9	21.4	86.4
RHR (bpm)	98.1	12.9	72.0	120.0
WOC (kg)	17.7	3.6	11.8	25.0

Note. RHR = Resting heart rate; WOC = weight of wheelchair.

In the present study, all of the subjects propelled wheelchairs while performing the TAMT. The aim of the test was to get the subjects to sustain moderate physical activity for a period of 15 continuous minutes in a target heart rate zone between 130 and 160 bpm for children and adolescents between the ages of 10 and 18 years. This represented approximately 70 to 85% of the subjects' maximum predicted heart rate with a downward adjustment of 15 bpm for exercise primarily using the arms only (Rimmer, 1994, p. 221).

Heart rates and cumulative distance traveled were recorded at 1-minute intervals during testing. If the heart rate was within the designated target heart rate zone at a recording interval (130 to 160 bpm), subjects were encouraged to keep going and were told that they were doing a great job. If the heart rate was below the target heart rate zone (<130 bpm), subjects were encouraged to go a little faster in order to achieve a higher heart rate. If the value was higher than the top of the target heart rate zone (>160 bpm), subjects were encouraged to slow down. However, subjects were permitted to exceed the upper limit and given a "pass" if they were able to complete 15 continuous minutes at a threshold above 160 bpm. Scores of pass/fail were administered to each subject on the TAMT. Subjects who exceeded the minimal threshold of 130 bpm received a "pass" on the test. A detailed description of the TAMT can be found in Appendix 1.

Test Procedures. A schedule was arranged to perform the TAMT in groups of six subjects or less, a minimum of two times with a 1-day rest between tests. Due to schedule conflicts, two subjects were able to be tested only one time. The remaining 30 subjects were tested two times.

Three different testing sites were used depending on their availability. A non-air conditioned indoor arena with a smooth surface provided a 144 yd (132 m) oval track. Half of three parallel tennis courts with an all-weather surface provided a 133 yd (122 m) track, and the third testing site was a small blacktop parking area with a 97 yd (89 m) track marked off. In order to measure the cumulative distance traveled at 1-minute intervals, eight cones were placed around the track at equal distances from each other.

Subjects wore Polar Accruex II heart rate monitors while performing the TAMT. On the first test, however, three subjects who wore body-contoured back braces did not wear a monitor and instead had their heart rate recorded manually at intervals of 2, 4, 6, 9, 12, and 15 minutes. On the second test, permission was granted to loosen the back braces and place the transmitter, without the strap, on the chest and then fasten the brace over the transmitter to keep it in place. This modification was an effective alternative to regular heart rate monitoring procedures. On Test 2, the heart rate monitor malfunctioned on three other subjects, who then had their heart rate recorded manually at the carotid or radial pulse for 10 seconds by one of the principal investigators at 2, 4, 6, 9, 12, and 15 minutes.

Each subject wheeled to the testing area without assistance. This took approximately 5 minutes and was used as a warm-up. After pushing one lap around the track, a button was pushed on each subject's heart rate monitor to begin testing. The heart rate monitors recorded time and cumulative heart rates for the entire 15-minute testing period. During the test, the lead investigator called off 1-minute intervals, at which time each subject read off their heart rate from the watch. A camp counselor familiar with the test procedures, and one of two lead investigators, recorded the heart rate along with the distance traveled. Distance was measured in cumulative laps plus completed cone increments. There was a separate score sheet for each subject on each testing day.

Results

Several analyses were conducted to evaluate the intensity, duration, and consistency of the TAMT. The first analysis dealt with attainability. Subjects who had their heart rates recorded manually were excluded from this part of the analysis to ensure maximum accuracy. In Test 1, the three subjects who wore back braces, as well as one subject whose monitor yielded inaccurate readings, were eliminated from this analysis. In Test 2, two subjects were unavailable for testing and three subjects had their heart rate monitors malfunction. Twenty-seven out of 28 eligible subjects (96%) on Test 1, and 25 out of 27 eligible subjects (93%) on Test 2, met the criterion for successful completion of the TAMT. One subject in Test 1 and two subjects in Test 2 were unable to maintain a heart rate greater or equal to 130 bpm and, therefore, failed on these trials of the test. The subject who failed Test 1 passed the TAMT in Test 2, and one subject who failed Test 2 passed the TAMT in Test 1. There was only one subject who did not pass the TAMT at least once and this subject completed only one test. Results are not presented by gender because of the high success rates (96% and 93%).

The second analysis evaluated the consistency of the results between Tests 1 and 2. Twenty-four subjects who had functional heart rate monitors on both tests, or who performed both tests, were evaluated in terms of their passing or failing the TAMT according to the previously stated criteria. Out of the 24 subjects, 22 subjects passed both tests (proportion of agreement = .92). A z test for significance of differences between two proportions (Bruning & Kuntz, 1987) revealed no significant difference ($p > .05$) in the proportion of subjects who passed Test 1 or Test 2. All 24 subjects passed one of the two tests.

The average heart rate of subjects over the 15-minute test also was analyzed in an attempt to study consistency of results. Some subjects stayed strictly within the 130 to 160 target heart rate zone, while others exercised above 160 bpm for at least some portion of the test. For subjects exercising between 130 and 160 bpm on both tests ($n = 10$), mean heart rate for Trial 1 was 146.6 bpm (±7.6) and for Trial 2, 147.0 bpm (±7.6). For the 22 subjects exercising within the target heart rate zone of 130 to 160 bpm and/or above 160 bpm, the mean heart rate for Trial 1 was 156.0 bpm (±12.1) and for Trial 2, 157.1 bpm (±13.6). A correlated *t*-test revealed no significant difference in mean heart rate between trials ($p > .05$) for subjects ($n = 22$) exercising within and/or above the target heart rate zone, and a significant correlation was obtained on mean heart rate between trials ($r = .71$, $p < .05$).

Relative to cumulative distance, subjects were able to push further on Trial 1 (1716 yd [1569 m]) compared to Trial 2 (1607 yd [1469 m]), $p < .05$. Although there was a statistically significant difference on cumulative distance traveled between Tests 1 and 2, the intraclass reliability was very high ($R = .94$, $p < .000$). The lower cumulative distance covered during Test 2 was probably the result of the shorter course, which was used for all the subjects in Test 2 and required more turns and, therefore, a greater amount of deceleration around the turns.

Discussion

The present study sought to determine if the Target Aerobic Movement Test (TAMT) could be considered a feasible test for measuring moderate levels of *aerobic be-*

havior in children and adolescents with spina bifida. Feasibility was determined by studying the suitability of the testing procedures for children and adolescents wheeling around a designated area, the attainability of test intensity and duration, and consistency (reliability) of test results. The results of our study suggest that the TAMT is a feasible test for measuring moderate levels of aerobic behavior in children and adolescents with spina bifida.

Regarding the suitability of the testing procedures, testing went smoothly with all of the subjects with the exception of two minor problems. During Test 1, three subjects were not able to wear the heart rate monitor because of interference with body-contoured back braces. During Test 2, however, this problem was alleviated by loosening the back braces and placing the transmitter against the chest and then fastening the brace over the transmitter to keep it in place. The second minor problem involved the heart rate monitor malfunctioning (gave erratic readings or shut down) for one subject on Test 1 and three subjects on Test 2. This may have resulted from the transmitter slipping below the breast level dux+ing testing or from poor contact.

Reliability was primarily determined by examining the passing rate of subjects on successive administrations of the test, as well as analyzing the relationship and significance of differences in mean heart rates for the 15-minute exercise period between Tests 1 and 2. Based on these analyses, it appears that the TAMT is reliable for use in field-based settings. The percentage of subjects passing Test 1 and Test 2 was 96% and 93%, respectively. All 24 subjects who were evaluated on both tests with an intact heart rate monitor passed one of the two tests.

Clearly, one of the disadvantages of the TAMT is that it does not discriminate fitness levels among those who successfully pass it. It is plausible, however, that such distinctions may be enhanced if distance (work) along with heart rate were evaluated. As a matter of interest, pilot data were collected in this regard. These preliminary data suggested that the detection of differences in health-related physical fitness on the TAMT may be enhanced by evaluating distance covered, as an index of workload, in addition to monitoring heart rate. This contention is only suggestive, however, because no attempt was made to control for wheelchair type or to account for differences in course layout or surface texture.

In this study, mean heart rates between trials were virtually identical and the correlation between heart rates on Test 1 and Test 2 was strong. Interestingly, the ability to work at heart rates $\geq$160 bpm, which equates to a training intensity greater than 80% of maximum heart rate for children and adolescents between the ages of 10 to 18 years old performing arms-only exercise was demonstrated in several subjects. Much to our surprise, 10 subjects (31%) on Test 1 and 11 subjects (37%)

Table 2 Passing Rates on the TAMT

Test	Number	Number passing	% passing
1	28	27	96%
2	27	25	93%

on Test 2 ignored our requests to slow down after they were told that their heart rate was above the upper limit of their training zone and finished the test with a mean heart rate ≥ 160 bpm. It was evident that these subjects were performing at a high enough intensity level to reflect a more vigorous level of aerobic behavior. However, more research needs to be conducted to determine if the TAMT, using a target heart rate zone of greater intensity, would be a good indicator of vigorous physical activity.

In conclusion, the TAMT appears to be a reliable and feasible test for measuring moderate levels of aerobic behavior in children and adolescents with spina bifida. Future research should focus on studying the feasibility of the TAMT with other populations with disabilities and, if the test is revised appropriately, use of the test as a reflection of the ability to perform vigorous physical activity and as a more refined discriminator of aerobic behavior and aerobic capacity.

References

American College of Sports Medicine. (1991). *Guidelines for exercise testing and prescription* (4th ed.) Philadelphia: Lea & Febiger.

American College of Sports Medicine. (1995). *ACSM's guidelines for exercise testing and prescription* (5th ed.) Baltimore: Williams & Wilkins.

Bruning, J.L., & Kuntz, B.L. (1987). *Computational handbook of statistics*. Glenview, IL: Scott, Foresman, and Co.

Consensus Development Conference. (1995, December 18-20). *NIH consensus development conference on physical activity and cardiovascular health*. Bethesda, MD.

Janssen, T.W.J., Van Oers, C.A.J.M., Van Der Woude, L.H., & Hollander, A.P. (1994). physical strain in daily life of wheelchair users with spinal cord injuries. *Medicine and Science in Sports and Exercise*, 26, 661-670.

Noreau, L., & Shephard, R.J. (1992). Physical fitness and productive activity of paraplegics. *Sports Medicine, Training, and Rehabilitation*, 3, 165-181.

Pate, R.R., Pratt, M., Blair, S.N., Haskell, W.L., Macera, C.A., Bouchard, C. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association*, 273, 402-407.

Public Health Service. (1991). *Healthy people 2000. National health promotion and disease prevention objectives* (DHHS Publication No. 91-50212). Washington, DC: U.S. Department of Health and Human Services.

Rimmer, J.H. (1994). *Fitness and rehabilitation programs for special populations*. Dubuque, IA: Brown & Benchmark.

Rimmer, J.H., Braddock, D., & Pitetti, K.H. (1996). Research in physical activity and disability: An emerging national priority. *Medicine and Science in Sports and Exercise*, 28, 1366-1372.

Santiago, M.C., Coyle, C.P., & Kinney, W.B. (1993). Aerobic exercise effect on individuals with physical disabilities. *Archives of Physical Medicine and Rehabilitation*, 74, 1192-1198.

Tell, G.S., & Vellar, O.D. (1988). Physical fitness, physical activity, and cardiovascular disease risk factors in adolescents: The Oslo youth study. *Preventive Medicine*, 17, 12-24.

Winnick, J.P., & Short, F.X. (1995). *Project Target test manual*. Unpublished manuscript, State University of New York College at Brockport, NY.

Appendix 1

Target Aerobic Movement Test[3]

Description: This test is designed to assess the aerobic behavior of children and adolescents. Participants attempt to exercise within a target heart rate zone (THRZ) for 15 minutes. Participants can engage in virtually any physical activity they choose as long as the activity is of sufficient intensity to raise the heart rate into the THRZ. In preparation for this test, instructors are encouraged to work with participants to help them identify an appropriate physical activity. For most participants who engage in "whole-body" forms of exercise, the THRZ is defined as 70 to 85% of maximum predicted heart rate (i.e., 140-180 bpm). There are two exceptions to these THRZ values. The first is for participants who have a spinal cord injury that results in low level quadriplegia (spinal lesion between C6-C8 inclusive). For those youngsters, THRZ may be defined in one of two ways. If a youngster has a resting (sitting) heart rate of less than 65 beats per minute, the THRZ is defined as 85 to 100 beats per minute. If a youngster's resting heart rate is 65 or more beats per minute, the THRZ is defined as a range of 20 to 30 beats above the resting value. The second exception applies to those who engage strictly in arm exercise. For those who use arms-only forms of exercise, the THRZ ranges from 130 to 170 beats per minute. The tester checks the participant's heart rate at least once every 60 seconds. If participants are within their THRZ, the tester reinforces the behavior and encourages participants to continue at their present intensity of exercise. If participants are below their THRZ, the tester encourages participants to increase exercise intensity. Should participants fall below their THRZ, they have one minute to regain their minimum value. If they do, the test continues; if not, the test is terminated at that time. If the participant is above the THRZ, the tester should acknowledge the participant's effort but also encourage the participant to decrease exercise intensity. If a participant is above THRZ, the test is continued and passed if the participant exercises for 15 minutes.

Equipment: It is recommended that testers use an electronic heart rate monitor in administering this test. It also is recommended that music with a fast tempo be played during the test to provide motivation and a sense of rhythmic, steady-state exercise.

Scoring and Trials: One test trial is given. This is a pass/fail test item; participants who can stay in or above the THRZ for 15 minutes pass the test. The 15-minute count does not begin until after the participant enters the THRZ. For those unable to pass the test, it is recommended that testers note the length of time that the participant was able to exercise in the THRZ.

Test Modification: If a heart rate monitor is unavailable, the test may be administered using the following procedures: Pulse rate at the wrist (i.e., radial pulse) is counted manually for 10-second intervals at a number of predetermined checkpoints. (The participant's exercise must be briefly interrupted for each pulse check). Pulse rate is checked at the end of a 3-minute warm up period and at the end of each of the following test exercise intervals: 2 minutes, 4 minutes, 6 minutes, 9 minutes, 12 minutes, and 15 minutes. Participants able to exercise for minimum THRZ values for 15 minutes pass the test.

Suggestions for Test Administration

Provide a "cool-down" area and activities of decreasing intensity for participants at the conclusion of the test. Individuals with spinal injuries above T6 are subject to "automatic dysreflexia," a condition which can elevate the heart rate (and blood pressure) as a result of bowel or bladder distention or skin irritation. As a precaution, therefore, it is recommended

that youngsters with spinal cord injuries above T6 empty their bowels and bladders and be checked for tight clothing, straps, or pressure sores which might contribute to skin irritation prior to testing.

Notes

[1]This study was supported, in part, by the Office of Special Education and Rehabilitative Services (OSERS), U.S. Department of Education, as a part of Project Target: Criterion-Referenced Physical Fitness Standards for Adolescents With Disabilities, No. H023C00191, directed by Joseph P. Winnick at the State University of New York, College at Brockport, NY. The contents presented in this document are those of the authors and do not necessarily reflect the position or policy of OSERS, and no official endorsement by OSERS should be inferred.

[2]The 1991 ACSM recommendation for minimal duration was 15 minutes. In 1995, ACSM revised its recommendation to 20 minutes. We chose to employ the 15-minute guideline because the recommendation by ACSM was for training purposes and was not a standard for testing aerobic behavior. The project advisory committee, made up of experts in exercise physiology and adapted physical education, felt that a 15-minute test would adequately represent an individual's aerobic behavior.

[3]This test is a modification of the aerobic movement test developed by Pat Good, Howe School, Dearborn, MI 48124.

Field Tests Employing
Subjects with Cerebral Palsy

1995-96

Brockport, NY
Atlanta, GA

Among the challenges associated with Project Target, selecting test items and establishing standards for youngsters with cerebral palsy are some of the more significant. The variation in the type and extent of the impairment contributes to the difficulty. Gaining access to large numbers of subjects with cerebral palsy to field test items and standards is also problematic. From October 1995 through October 1996 project staff collected research data on 38 subjects with cerebral palsy in conjunction with three disabled sporting events. During October 1995 and October 1996, subjects were tested at the New York State Games for the Physically Challenged held on the campus of SUNY, College at Brockport. In August 1996 the staff traveled to Atlanta, Georgia and tested 13 subjects at the Paralympic Games.

Data collection at all three events focused on items and standards associated with musculoskeletal functioning (strength, endurance, and flexibility). Although items varied slightly from site to site, this report includes data on the following items: Apley Test (modified), Thomas Test (modified), Target Stretch Test, 15-lb. Dumbbell Press, and Dominant Grip Strength. (These were the items recommended for youngsters with cerebral palsy at the time this report was written.) In addition to field testing the items, the purposes of these efforts were: 1) to determine the objectivity of the Target Stretch Test; 2) to determine if evidence of criterion-related validity could be found for either the Apley, Stretch, or Thomas tests; and 3) to determine pass/fail rates for each of the test items.

Objectivity

The Target Stretch Test is a subjective measure of extent flexibility. Testers give a score to a youngster's performance based on the position of a body part in relationship to visual criteria provided in a sketch. Due to the subjective nature of the scoring, it would be important to demonstrate that two different testers generally award the same score for the same performance. During the 1996 Games for the Physically Challenged, two staff members (the project coordinator and a graduate assistant) independently evaluated subjects on the Stretch Test. Testers used visual criteria to award scores between 1 and "5+". (Testing protocol is provided in the Project Target Test Manual.) A total of 175 paired observations were made on a variety of the subtests. Inter-rater reliability correlation coefficients are given below for each of the subtests utilized.

Table 1
Inter-rater Reliability Correlation Coefficients
for the Stretch Test

	r	n
Wrist Extension (R)	.94	4
Wrist Extension (L)	.96	4
Elbow Extension (R)	.83	20
Elbow Extension (L)	.78	20
Shoulder Extension (R)	.80	12
Shoulder Extension (L)	.72	12
Shoulder Abduction (R)	.98	19
Shoulder Abduction (L)	.82	20
External Rotation (R)	.65	12
External Rotation (L)	.82	12
Supination (R)	.95	6
Supination (L)	.98	6
Knee Extension (R)	.67	14
Knee Extension (L)	.75	14

Correlation coefficients presented in Table 1 range from .65 to .98. The lowest Pearson r's are associated with knee extension and shoulder external rotation, the highest with wrist extension and supination. When the subtests are considered together the coefficient for all 175 paired observations is .86. Safrit (1990) recommended that inter-rater reliability coefficients be "relatively high" and she suggested .80 or above as an acceptable criterion. The overall coefficient of .86 appears to represent a level of objectivity that is sufficient for the purposes of this test. Among individual subtests, only knee extension failed to reach .80 on at least one side of the body.

Awarding scores between 1 and 5 on the Stretch Test requires testers to make judgement with a 20% margin of error (each whole number represents 20% of the possible range of motion in a joint). Including 5+ in the scoring scheme should hurt objectivity because testers must make a more precise judgement; namely, does the youngster achieve optimal movement extent?

Technically, this requires the tester to judge with no margin of error. When 5+ was eliminated from the scoring system used with the above data, the Pearson r rose to .89 for the overall test. Since the minimal standard for each of the subtests is 5 while the preferred is 5+, judgements around the minimal standard will be slightly more reliable (i.e., objective) than those pertaining to the preferred.

Criterion-related Validity

Scientific information on the criterion-related validity of standards associated with measures of aerobic functioning and body composition is available and serves as the basis for those standards (FITNESSGRAM, 1994). Direct evidence linking musculoskeletal functioning (MSF) to indices of health, however, are lacking and standards for these measures are frequently based on expert opinion. To determine if any relationship might exist between the measures of musculoskeletal functioning used in this test and indices of health, subjects were asked to complete two questionnaires. The first asked subjects to rate their functional independence on a variety of activities of daily living. Scores ranged from 1 (completely dependent) to 7 (completely independent). On the second questionnaire subjects were asked to estimate the frequency, intensity, and duration of their typical exercise behavior. Scores ranged from 1 to 4 for each variable with 4 representing the "most" or "highest" score. (Both questionnaires are attached to this report.)

The first analysis of this data focused on the relationships among functional movement extent and functional independence on selected ADLs. Functional movement extent was defined as obtaining a score of 5 or 5+ for the preferred limb on selected subtests of the Stretch Test. Scores below 5 were considered nonfunctional. Functional independence was defined as a score of 6 or 7 for a particular ADL. Scores below 6 were considered dependent for that ADL. By so defining functional movement extent and functional independence, a series of 2x2 contingency tables were constructed and phi coefficients were calculated to determine the magnitude of the various relationships. Results of this analysis are provided in Table 2.

Table 2
Relationships Between Items From the Stretch Test
and the Functional Independence Survey

	Eating	Dress Upper Body	Dress Lower Body	Toileting	Transfers (bed,chair, WC)	Propel WC	Stairs
Wrist Extension	.32	.32	-	.32	.41	.50	-
Elbow Extension	.15	.46*	.35*	.15	.25	.38	.31*
Shoulder Extension	-	-	-	-	-	-	-
Shoulder Abduction	.56*	.61*	.48*	.31*	.24	.33	.33*
External Rotation	.42*	.47*	.35*	.13	-.12	.66*	.30
Forearm Supination	.37	.37	.27	.37	-	-	-
Knee Extension	-.04	.03	.11	-.04	-	-	.11

* $p < .05$

The phi coefficients presented in Table 2 generally suggest low to moderate relationships between items from the Stretch Test and Functional Independence Survey. Four coefficients met or exceeded .50: wrist extension and wheelchair propulsion; shoulder abduction and eating; shoulder abduction and dressing upper body; and external rotation and wheelchair propulsion. The magnitude of these relationships is somewhat disappointing since earlier pilot work suggested that coefficients above .70 were possible. There remain strong arguments for the logical validity that certain levels of flexibility are prerequisite (and therefore related) to performing specific ADLs in an independent fashion. This correlation study, however, failed to confirm the logic. It may be that the instruments used were not sensitive enough to detect important differences in function or that function itself was poorly defined. Employing self-report of functional independence with adolescents also may have contributed to measurement error. It may also be true that while a certain level of flexibility is necessary to perform certain ADLs, it is insufficient to explain that performance.

The second analysis focused on the relationships between the various musculoskeletal test items and patterns of physical activity. The patterns of physical activity were expressed in terms of frequency, duration and intensity of exercise. In addition a composite score was calculated by adding frequency, duration and intensity scores together. No attempt was made to construct contingency tables for this analysis; test scores were simply correlated with the self-reported patterns of physical activity. The results are presented in Table 3.

Table 3
Pearson Correlation Coefficients Between Test Items
and Patterns of Physical Activity

	Frequency	Duration	Intensity	Composite
Apley (R)	-.11	.01	.01	-.06
Apley (L)	-.16	.10	-.09	-.04
Thomas (R)	.14	-.23	.14	.01
Thomas (L)	.04	-.26	.08	-.09
Wrist (R)	-.57	.69	.36	.65
Wrist (L)	.03	.77	-.25	.47
Elbow (R)	-.07	-.39*	.03	-.28
Elbow (L)	-.09	-.31*	-.27	-.38*
Shoulder Extension (R)	-.02	-.57*	.08	-.43*
Shoulder Extension (L)	-.48*	-.37*	.08	-.52*
Shoulder Abduction (R)	-.08	-.12	.00	-.15
Shoulder Abduction (L)	-.20	-.15	-.07	-.27
External Rotation (R)	-.39*	.09	.31	.02
External Rotation (L)	-.28	.24	.00	.05
Supination (R)	-.18	.04	.30	.14
Supination (L)	.02	.29	-.11	.20
Knee (R)	.00	-.18	.07	-.07
Knee (L)	-.14	-.27	-.02	-.25
Seated Push-up	.11	.68*	-.20	.46
Dominant Grip	.17	.55*	.11	.41
Dumbbell Press	.12	.35	-.37	-.01

* $p < .05$

Most of the Pearson r values in Table 3 could be described as depicting low to moderate relationships. Of some interest is the fact that many of the coefficients associated with the Stretch test items are negative. In fact, three of the composite coefficients are moderate and significant indicating that higher levels of physical activity tend to be associated with lower flexibility test scores! The news is a little better for measures of strength and endurance. Most of the coefficients are positive and both seated push-ups and grip strength have moderate and significant relationships with duration. Still, the overall result is that evidence of criterion-related validity for the musculoskeletal functioning test items for youngsters with cerebral palsy remains elusive.

Pass/Fail Rates

Pass/fail rates for the musculoskeletal functioning items are summarized in Table 4. All subtests of the Stretch test are combined in this analysis. USCPAA classes are shown along with accompanying standards. Standards for the Apley, Thomas, and Stretch tests reflect levels of performance described in the test protocol. A score of 3 is a maximum score for the Apley and Thomas tests. A score of 5 for the Stretch Test represents 80% of optimal movement extent in the joint being assessed. The standards for the dumbbell and grip strength items represent the 20th and 60th percentile values for samples of nondisabled boys and girls at various ages (10-17). The scores of subjects with CP who were older than 17 were compared to the standards for 17-year-olds.

Table 4
Pass/Fail Rates on Recommended Standards for MSF Items

Item	Apley			Thomas		Stretch		Seated		Dumbbell		Grip	
Class	C1	C2 U	C2U-C8	C5	C6-C8	C1-C2	C3-C8	C2U C3	C4 C6	C3- C5	C7- C8	C4 C5	C7 C8
Standard	1	2	3	2	3	5	5	5	20	20th	60th	20th	60t
Males	1/0	0/0	18/6	5/4	9/2	10/6	92/7	7/0	6/1	3/3	0/6	4/5	1/8
Females	2/0	0/0	3/3	0/0	3/0	17/5	30/1	3/0	3/0	2/1	0/3	4/1	4/1
Total	3/0	0/0	21/9	5/4	12/2	27/11	122/8	10/0	9/1	5/4	0/9	8/6	5/9
% Pass	100	-	70	56	86	71	94	100	90	56	0	57	36

Separate standards are provided for the Apley and Thomas tests as a function of classification. For the Apley all three class 1 subjects were able to achieve the criterion (touch hand to lips). No class 2L subjects were tested (the standard requires subjects to touch their hand to the top of the head). Subjects from classes 2U through 8 were required to touch their preferred hand to the opposite scapula (superior medial angle). Seventy percent of 30 subjects were able to do so. For the Thomas test classes 6 through 8 were required to pass it in the traditional way;

that is, subjects are expected to maintain a straight leg in contact with the table surface when the opposite leg is pulled to the chest until the back is flat. Twelve out of 14 subjects (86%) were able to achieve this criterion. The standard is eased for those is class 5 due to the nature of the disability. For class 5 the standard allows the leg to come off the table surface, but no more than approximately 15 degrees at the hip joint. This seems to be a rigorous standard for class 5; five out of nine subjects (56%)achieved the standard. In general, however, the standards for the Apley and Thomas tests appear to be reasonable.

The standard associated with the Stretch Test proved to be "easy." One hundred- thirty subtests were administered to subjects in classes 3 through 8. Subjects made a passing score on 122 of the tests for a 94% pass rate. In fact, even classes 1 and 2 were able to be successful 71% of the time. Currently an "individualized" standard is recommended for classes 1 and 2, but these data suggest that these youngsters generally can make the 80% of movement extent criterion. Given the high rate of success for subjects with CP perhaps a more challenging definition of "functional" movement extent should be considered.

The 5-second criterion for the seated push-up was easily attainable by subjects from the four relevant classes; all 10 achieved the standard. Nine out of 10 subjects also were able to attain the 20-second standard. The only failure was a 10-year-old boy. Since the 5-second standard is tied directly to a recommendation for the reduction of risk in developing pressure sores it is appropriate for that standard to exist across the 10-17 age range. It may be, however, that the preferred standard, which has been more arbitrarily determined, should vary more as a function of age to more closely parallel other strength and endurance items. Potentially such a standard would be more challenging for older youngsters.

Standards for the dumbbell press and grip strength were more difficult to attain. Fifty-six percent of the relevant subjects were able to achieve the minimal standard on the dumbbell press, while 57% were successful for the grip strength. In each case these standards represented the 20th percentile value for nondisabled youngsters. The 60th percentile standard proved to be very challenging especially for the dumbbell press. None of the nine subjects who took the dumbbell press reached this criterion of success. Five out of 14 (36%) met this standard on the grip strength. Actually this is a reasonable level of success since by definition only about 40% of nondisabled youngsters would be expected to score at the 60th percentile. Overall the standards for these two items generally seem to be appropriate, but there is a caveat when interpreting these statistics. The sample of subjects includes adult Paralympians who were compared to the 17-year-old standards. Since strength generally peaks in the late 20's the pass rates here may be overestimates assuming that people with cerebral palsy follow the same developmental profile as people without disabilities.

Summary and Recommendations

1) The objectivity of the Stretch Test is sufficient for the purposes of an assessment instrument. The subtest with the lowest objectivity is knee extension. A few adjustments have been made to the protocol for this item which hopefully will improve its objectivity.

2) No strong evidence linking MSF items to either ADLs or physical activity patterns were found. Consequently standards for MSF items largely will be based on expert opinion which, if not as desirable, at least is consistent with procedures used with other criterion-referenced tests.

3) Generally speaking, the pass/fail rates of the current standards for youngsters with cerebral palsy appear reasonable. Hopefully field testing at the New York State Games for the Physically Challenged in May 1997 will help to confirm this impression. Should adjustments to standards be considered on the basis of pass/fail, however, the standards for the Stretch Test and the seated push-up (preferred standard) might be too low and those associated with the dumbbell and grip strength might be too high (if, in fact, the presence of adult subjects skewed the data).

Disk: F.X. Short
File: fieldrep.cp

Empire State Games
for the
Physically Challenged 1997

Project Target
Research Report

Project staff continued to field test Target test items with youngsters with physical disabilities on two occasions during 1997. In May, staff traveled to Uniondale, N.Y. on Long Island to test youngsters at the state-wide Games for the Physically Challenged. Testing took place at Nassau County Community College. Staff also tested youngsters at the regional Games conducted at SUNY Brockport during October. Twelve subjects were tested at each site. Of the 24 youngsters tested, 19 had cerebral palsy (CP) and five had a spinal cord injury (SCI). The purpose of the testing was to continue to gather data on the feasibility and attainability of tests and standards recommended for youngsters with physical disabilities (especially CP). One planned purpose of the Brockport testing (test-retest reliability of certain items) was abandoned due to low number of subjects. Videotaping was conducted at the Brockport Games in preparation for a training tape to be developed on the test. The following test items were administered at one or both sites: Target Stretch Test, Apley Test, Thomas Test, 40-meter Push/Walk, Ramp Test, Seated Push-up, Dominant Grip, and Dominant Dumbbell Press. Test protocols and associated standards can be found in the August 1997 version of the test manual.

Results

Pass/fail results for seven of the test items are provided in Table 1. For test items with both minimal and preferred standards, the minimal standards were used to evaluate performance. Due to low subject numbers, no definitive conclusions can be reached pertaining to the attainability of the standards recommended for youngsters with CP. These numbers, however, are at least encouraging in suggesting that the standards may be realistic for this group.

Table 1
Pass/Fail Results for Test Items Taken by
Subjects with Cerebral Palsy

	Pass #	**Fail #**
Apley	4	3
Thomas	9	4
40-meter	6	0
Ramp	2	0
Seated Push-up	4	0
Grip Strength	5	1
Dumbell Press	1	3

Pass/fail information collected on the Target Stretch Test (TST) was also encouraging. Utilizing the revised scoring system (0-2), 15 subjects with CP (class 3 and higher) were given a variety of TST subtests. For the 45 total subtests administered, a score of 2 (indicative of optimal or typical range of motion) was awarded 32 times (71%). A score of 1 was given 12 times (27%). A score of 1 implies that the individual has at least functional range of motion. A score of 0 was given just one time (2%); a 0 indicates that the range of motion is not functional in a general sense. The 98% pass rate (for attaining at least minimal standards) is close to the 94% rate obtained when using an earlier version of the test.

Videotaping much of the testing at the Brockport Games proved to be valuable as protocols for two test items were modified as a result of viewing the tape. Staff identified problems with the modified Thomas test and the knee extension subtest of the TST. In the Thomas test the greater trochanters of the femur are lined-up with a tape mark placed 11" from the end of a table and scoring is predicated on this 11" distance. A review of the videotape, however, indicated that when one knee is drawn toward the chest and the pelvis rotates back, the trochanters move off the 11" line (i.e., further away from the edge of the table) for some youngsters. Consequently, the Thomas protocol now includes a suggestion that testers place a tape mark on the greater trochanters and verify that those marks are located above the 11" line prior to scoring the test.

The problem with the knee extension subtest was that even though it appeared that some subjects could achieve complete knee extension, they could not get the lower leg to the required horizontal position (eg., 3 o'clock) to get the maximum score (i.e., 2). (Previous work on the TST indicated that the knee extension test had the lowest objectivity of any of the subtests in the battery.) This may be due to a tapering of the upper leg that causes the femur to be less than horizontal in a sitting position. If the femur is not horizontal, the extended lower leg cannot get to horizontal either. At any rate, a decision was made to conduct this test in a side-lying position rather than in a sitting position.

The videotape was also used to collect some criterion-related validity data on the TST. Twenty TST subtests were administered to three subjects and videotaped from a tester's perspective. Three graduate students were given a five-minute training session on scoring the TST, were then shown the videos of the 20 TST subtests, and scored them. The project coordinator served as a fourth tester. Criterion scores were established by taking goniometry readings of the 20 joint actions from the videotape. TST scores ranging from 0-2 were given for each of the 20 subtests based on the goniometry values. The scores obtained for the four testers were then compared to the criterion scores obtained through goniometry (a total of 80 comparisons). The testers correctly scored the tests 85% of the time. Individual accuracy scores ranged from 75% to 95% among the four testers. When the goniometry values were rounded to the "nearest half-hour" (the protocol required of testers on the TST), the testers' accuracy improved to 90%. Safrit (1990) indicated that validity coefficients determined in this manner should exceed 80%.

Summary

Although testing conducted at the ESGPC at both Uniondale and Brockport failed to generate a large subject base, these efforts were helpful in fine tuning the final test battery. Adjustments were made to two test protocols, evidence that the standards appear to be reasonable emerged, and data collected

on the TST lends some additional confidence that this is an adequate measure of range of motion. Furthermore, the feasibility of administration of these test items to youngsters with physical disabilities was further enhanced as a result of the 1997 ESGPC testing.

PROJECT TARGET
RESEARCH REPORT

BROCKPORT CENTRAL SCHOOL DISTRICT

1995-96

Purpose and Overview

Over the course of this project, project staff has sought to identify health-related physical fitness tests which are appropriate for a particular disability. This search has sometimes uncovered potential test items which might be considered "nontraditional" or possibly "alternative" measures of fitness. Although an item may appear to have merit for a specific class of youngsters with disabilities, these nontraditional items typically lack a reference point by which to judge performance. So a test may be described in the literature or it may come highly recommended by a professional in the field, but oftentimes it is difficult to know how to assess a particular test score because little research has been done with that item. Of noteworthy import to project staff is to know the percentile values for nondisabled youngsters because in some instances staff may choose to set criterion-referenced standards based on certain percentile values (e.g., 20th and/or 60th) in much the same way as FITNESSGRAM. The purpose of this study, therefore, was to test nondisabled youngsters on certain nontraditional tests of fitness to determine "typical" performance. These data may be used to help set criterion-referenced standards for youngsters with disabilities.

Subjects

The subjects for this study came from three schools within the Brockport Central School District: the Hill School (an elementary school), Brockport Middle School, and Brockport High School. A total of 904 boys and girls between the ages of 9 and 17 were tested on at least one of the test items. Subjects were tested as part of their regular physical education class. High school students were tested in the fall of 1995, middle school students were tested during January 1996, and Hill School students were tested during May 1996.

Methods and Procedures

All subjects were tested by graduate students associated with the adapted physical education program at SUNY Brockport. Testing was conducted in the physical presence of Project Target investigators. Each student was very familiar with Project Target test procedures. The so-called nontraditional test items included the 35-lb bench press, the 15-lb dumbbell press, the 10-lb dumbbell press, the extended arm hang, and the isometric push-up. In addition to these tests, three more traditional measures of fitness were included in the test battery to determine the degree of relationship between the traditional and nontraditional tests. The three traditional measures included the flexed arm hang, push-ups, and grip strength. A final element of the Brockport testing was a test-retest reliability study on the bench press. Sixty-four male subjects performed a second trial on the bench press during their next physical education class. Test protocols for all of the items in the battery are described in the October 1995 version of the Project Target Test Manual.

Results

Descriptive results are summarized in Tables 1-8. Subject performance on each test item is broken down by gender and age (10-17 only) and includes number of subjects, mean, standard deviation, and 20th, 50th, and 60th percentile values.

Table 1. Bench Press (30 rep max for girls; 50 for boys)

	N	M	SD	20th	50th	60th
Girls						
10	-	-	-	-	-	-
11	73	11.2	7.4	5	10	12
12	36	14.0	8.4	6	13	15
13	2	21.5	12.0	-	-	-
14	42	22.6	8.9	14	26	30
15	52	20.5	8.7	12	20	26
16	35	22.3	7.6	17	22	25
17	18	23.8	10.3	12	24	30
Boys						
10	-	-	-	-	-	-
11	83	20.1	11.3	12	18	20
12	42	22.6	12.6	12	20	23
13	5	29.0	13.1	17	25	29
14	42	40.8	10.7	31	46	50
15	65	46.7	8.2	50	50	50
16	40	45.9	7.3	40	50	50
17	21	49.5	2.2	50	50	50

Table 2. 15-lb. Dumbbell Press (50 rep max)

	N	M	SD	20th	50th	60th
Girls						
10	-	-	-	-	-	-
11	33	5.6	3.5	2	5	6
12	19	7.6	5.0	2	7	8
13	4	18.3	10.0	-	-	-
14	101	13.3	6.0	8	14	15
15	58	15.8	6.6	10	15	16
16	36	16.5	4.7	12	16	17
17	18	16.8	6.3	11	15	16
Boys						
10	-	-	-	-	-	-
11	78	12.6	6.6	7	12	13
12	42	17.6	8.8	9	17	18
13	44	21.4	8.9	14	21	21
14	56	28.4	10.9	20	26	28
15	68	32.5	9.9	24	31	34
16	41	30.4	9.3	20	30	37
17	21	39.9	10.0	27	43	45

Table 3. 10-lb. Dumbbell Press (50 rep max)

	N	M	SD	20th	50th	60th
Girls						
10	-	-	-	-	-	-
11	40	16.4	6.5	12	17	20
12	17	21.5	7.5	13	23	25
Boys						

10	-	-	-	-	-	-
11	74	23.2	8.8	15	22	25
12	34	24.4	9.1	17	23	25

Table 4. Extended Arm Hang (120 sec max)

	N	M	SD	20th	50th	60th
Girls						
10	50	54.3	38.7	17	43	60
11	100	45.9	31.4	20	37	49
12	34	42.3	26.3	19	36	51
13	5	51.2	23.9	21	54	61
14	100	63.7	35.7	30	56	68
15	57	62.1	34.7	30	60	62
16	36	52.4	24.6	29	50	60
17	18	47.9	37.9	16	30	46
Boys						
10	53	72.0	34.2	40	69	88
11	122	60.0	33.2	30	56	62
12	44	70.8	39.1	33	65	79
13	44	84.6	35.4	45	93	115
14	56	74.4	31.8	50	78	84
15	67	76.5	31.7	46	74	86
16	40	91.0	25.6	63	91	104
17	21	90.4	31.2	61	90	120

Table 5. Isometric Push-up (120 sec max)

	N	M	SD	20th	50th	60th
Girls						
10	53	63.5	41.9	23	48	75
11	27	68.7	42.5	26	51	94
12	-	-	-	-	-	-
13	3	92.7	26.1	-	-	-
14	58	93.9	29.9	71	99	120
15	3	77.0	39.6	-	-	-
16	-	-	-	-	-	-
17	-	-	-	-	-	-
Boys						
10	52	81.2	37.4	41	81	119
11	41	78.1	37.5	41	75	98
12	4	88.3	36.7	-	-	-
13	40	104.2	22.5	80	120	120
14	15	90.9	30.8	57	98	120
15	3	67.3	55.1	-	-	-
16	-	-	-	-	-	-
17	-	-	-	-	-	-

Table 6. Flexed Arm Hang

	N	M	SD	20th	50th	60th
Girls						
14	14	17.5	12.5	7	13	20
15	29	11.7	9.1	1	11	15
16	8	13.6	11.7	2	12	17
17	-	-	-	-	-	-
Boys						
14	18	15.4	9.2	7	18	19

15	33	22.4	14.4	8	22	27
16	9	23.8	10.8	15	22	28
17	3	37.3	24.2	-	-	-

Table 7. Push-ups

	N	M	SD	20th	50th	60th
Girls						
13	4	15.5	11.7	-	-	-
14	86	12.6	8.3	5	12	15
15	26	11.1	7.7	3	11	12
16	28	15.2	9.6	7	13	15
17	18	15.2	10.2	6	14	16
Boys						
13	39	21.5	9.1	15	23	25
14	38	23.3	10.7	13	25	28
15	35	25.8	16.9	11	25	28
16	31	25.1	9.1	15	26	29
17	16	27.2	10.0	20	26	28

Table 8. Dominant Grip Strength

	N	M	SD	20th	50th	60th
Girls						
10	53	20.4	4.2	17	19	20
11	27	21.0	4.3	17	20	21
12	-	-	-	-	-	-
13	4	32.0	3.4	-	-	-
14	101	30.2	5.0	26	31	32
15	55	32.7	5.1	29	33	33
16	36	33.4	4.5	29	34	35
17	18	32.6	5.7	28	31	32
Boys						
10	52	21.8	3.8	18	22	22
11	85	25.5	4.9	22	25	26
12	25	27.9	5.2	24	27	29
13	43	34.6	7.9	29	33	36
14	56	39.4	7.3	33	38	41
15	66	45.5	8.0	38	46	48
16	40	47.5	7.6	41	48	49
17	19	56.2	6.4	49	57	57

To determine the degree of relationship among these test items Pearson correlation coefficients were calculated. The correlation matrix is presented in Table 9.

Table 9. Intercorrelations

	FAH	Push	Bench	Dumb15	EAH	Iso	Dumb10	DGrip
FAH	1.0							
Push	-	1.0						
Bench	.40	.55	1.0					
Dumb15	.33	.49	.81	1.0				
EAH	.54	.51	.35	.29	1.0			
Iso	-	.55	-	.14	.44	1.0		
Dumb10	-	-	.41	.58	.27	-	1.0	
DGrip	.35	.39	.77	.76	.24	.17	.26	1.0

The results of the "mini-study" on bench press reliability are presented in Table 10.

Table 10. Test-Retest Reliability for 35-lb. Bench Press

	Trial 1	Trial 2
M	21.3	24.1
SD	12.1	10.9

n = 64
a = .92

Discussion

The Brockport testing sessions were successful in the sense that they provided insight into typical performance for atypical fitness tests. Of the nontraditional items, the bench press and the 15-lb dumbbell press appeared to present the "best" developmental progressions. For both of these items performance generally improved linearly with age and the means for both items were generally associated with relatively small standard deviations. Both tests appeared to be easily administered although the bench press requires more equipment while the dumbbell press requires a certain degree of subjectivity on the part of the examiner. (eg., when youngsters begin to fatigue, they do not always press the dumbbell "straight-up").

Consideration has been given to using the 10-lb. dumbbell press as a lead-up to the 15-lb. item, especially for ages 10-12. Correlational data, however, suggest that this may not be a good

idea. Despite the apparent similarity in these two tests, the Pearson r between these two items was only moderate, r = .58
(n = 107). If these two tests are not measuring the same construct, which one is doing the best job of measuring the desired construct? If bench press is used as the criterion measure of muscular strength and endurance, the answer to this question appears to be the 15-lb. dumbbell press. The 15-lb. dumbbell press had a Pearson r of .81 (n = 490) with the bench press while the Pearson r between the 10-lb. dumbbell press and the bench press was just .41 (n = 172). Based on these data, the best approach might be to use only the 15-lb. press as a formal part of the test. Data suggest that even the youngest subjects generally are able to do at least one repetition with the 15-lb. weight. For those who cannot do at least one repetition with the 15-lb. weight, a suggestion could be to substitute the 10-lb. weight, but provide no standard for performance. Testers could collect "individualized" data in this case.

The extended arm hang and isometric push-up are nontraditional items that are direct "descendants" from the traditional items flexed arm hang and push-ups. The data for the extended arm hang and isometric push-up are not as "clean" as the data associated with either the bench or dumbbell presses. Mean values for the extended arm hang and isometric push-up generally improve with age, but there are frequent exceptions to the rule. For instance, for the extended arm hang, 10-year-old boys and girls made better scores than their 11- and 12-year-old counterparts. Furthermore, the standard deviations for these items are fairly large in relation to the means; the standard deviations tend to run between 1/3 and 2/3s the size of the mean. The large standard deviations speak to the variability in subject performance and may be suggestive of lower test reliability. On the other hand, the "parent" items (flexed arm hang and push-ups) also tend to be associated with larger standard deviations so perhaps greater variability is to be expected. Clinically, it appeared that there was a fairly large subjective component to the isometric push-up test. Directions for this test should be reviewed to determine whether more specific criteria by which to judge test performance might be helpful to a tester. Conversely the extended arm hang was quite objective (although some bar swinging on the part of some subjects was observed), but was associated with a common practical problem. To a large extent subjects complained about their hands hurting at the conclusion of the test. This was especially true for those who made high scores. Any standard selected for this item, therefore, should reflect concerns about the tendency for this item to hurt youngsters' hands.

Fairly high correlation coefficients were found among bench press, 15-lb dumbbell press, and dominant grip strength. Grip strength correlated at .77 and .76 for bench press and dumbbell press, respectively, while the coefficient between the two presses was .81, the highest correlation among any of the test items. These coefficients suggest that testers might be able to select any one of these items to assess the upper body strength and endurance of a youngster. The correlation coefficients between the "parent" items and their "descendants" were moderate. Specifically, the Pearson r between flexed arm hang and extended arm hang was .54 and the r between push-ups and isometric push-up was .55. These coefficients seem to suggest that the "descendant" items cannot be considered replacements for the "parent" items, but they might be considered reasonable substitutes (or lead-ups) when the "parent" item is too difficult or otherwise inappropriate.

The test-retest data for the bench press was encouraging . The alpha coefficient determined from the performance of 64 males was .92 which is indicative of good reliability.

In conclusion, it would appear that the nontraditional items may have some good utility for Project Target. In establishing standards which may be based on the performance of the students from Brockport Central, however, it will be necessary to "smooth" the data in some cases (especially with extended arm hang and isometric push-up) to reflect desirable developmental progressions.

Disk: F.X. Short
File: bcsd.rep

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☐ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☑ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").