A Critique of Cardiovascular Fitness Testing With Mentally Retarded Persons

Christine Seidl, Greg Reid, and David L. Montgomery
McGill University

Recently there has been a plethora of research investigating various dimensions of the cardiovascular fitness of mentally retarded persons. It is clearly documented that as a group, mentally retarded persons are particularly low in aerobic fitness. Although there is evidence that such low cardiovascular functioning can be increased, exercise training studies have invariably ignored the important questions of reliability and validity of the dependent measures. Also, there are innumerable testing protocols that make cross-study comparisons tenuous. Several factors are fundamental to the reliability and validity of standardized protocols that have recently been used with retarded persons. These include underlying assumptions of cadence adherence, constant efficiency, learning, and motivation to perform optimally. The development of cardiovascular test protocols for use with retarded persons is necessary to provide for their immediate and future needs in cardiovascular fitness evaluation.

The importance of good cardiovascular fitness for general health and well-being should not be limited to nonhandicapped persons. It is clearly documented that, as a group, mentally retarded persons are particularly low on measures of cardiovascular fitness (Andrew, Reid, Beck, & McDonald, 1979; Beasley, 1982; Coleman, Ayoub, & Friedrich, 1976; Nordgren, 1970, 1971; Rarick, Dobbins, & Broadhead, 1976; Rarick, Widdop, & Broadhead, 1970; Reid, Montgomery, & Seidl, 1985; Schurrer, Weltman, & Brammell, 1985). Such low levels of cardiovascular functioning, as well as of other dimensions of physical fitness, have prompted several studies of an ameliorative nature. However, not all studies reported improvements in cardiovascular fitness (e.g., Andrew et al., 1979; Bundschuh & Cureton, 1982; Coleman & Whitman, 1984; Montgomery, Reid, & Seidl, in press).

More typically, the data suggest that the cardiovascular systems of mentally retarded persons are responsive to exercise regimens (Beasley, 1982; Montgomery et al., in press; Nordgren, 1971; Tomporowski & Ellis, 1984a, 1985). Webster, Tymesn, and Fernhall (1986) have reviewed much of the recent pro-

Request reprints from Greg Reid, McGill University, 475 Pine Avenue West, Montreal, P.Q., Canada, H2W 1S4.
grammatic research and have highlighted a number of concerns and potential solutions. Since demonstration of program efficacy assumes valid and reliable measures of cardiovascular functioning, the purpose of this paper is to focus specifically on the testing issues common to cardiovascular exercise programs. Topics to be addressed are variability in test methodology, cadence adherence, termination point, efficiency, learning, and motivation.

Variation in Test Methodology

Table 1 summarizes the cardiovascular fitness tests used with mentally retarded persons. They have ranged from field tests such as the 300-yard walk-run (Coleman & Whitman, 1984) to maximal treadmill walking and running protocols (Andrew et al., 1979; Bar-Or et al., 1971; Tomporowski & Ellis, 1984a, 1985; Schurrer et al., 1985). Direct comparison among studies is therefore tenuous. There is also evidence of different cardiovascular fitness levels between educable and trainable retarded citizens (Londeree & Johnson, 1974), which further compounds the problem of cross-study comparison.

It is not clear which of the many protocols, if any, are valid. No published studies have established the validities of cardiovascular fitness testing methodologies with mentally retarded persons. With the nonhandicapped population, the direct measurement of oxygen consumption during a maximal graded exercise test is a widely accepted means of accurately assessing cardiovascular fitness. Maksud and Hamilton (1974) used a direct measure during a maximal cycling protocol. Although five reports have used maximal treadmill protocols with mentally retarded persons (Andrew et al., 1979; Bar-Or et al., 1971; Schurrer et al., 1985; Tomporowski & Ellis, 1984a, 1985), only three of these (Andrew et al., 1979; Bar-Or et al., 1971; Schurrer et al., 1985) included direct measurement of oxygen consumption. There appears to be a need for more investigations using direct measures of oxygen consumption during maximal exercise, although this may be impossible for severely and profoundly retarded individuals. Such protocols should be validated in light of the issues raised in the section on learning and motivation.

Field tests and submaximal laboratory protocols have commonly been used. However, since these methods have been shown to be population-specific, there is a need to establish their validity. For example, Krahenbuhl, Pangrazi, Burkett, Schneider, and Peterson (1977) demonstrated that a 1-mile run was a more valid field test of cardiovascular fitness than shorter runs for 8-year-old nonhandicapped boys. However, it was not a valid indicant of cardiovascular functioning for the girls. Thus, with mentally retarded persons as subjects, validity of test protocols may prove to be specific to the particular level of retardation, gender, or age. The 300- and 600-yard walk-run tests have been popular (see Table 1), and have been justified for persons with mental retardation because of the subjects’ motivational status. However, these field measures are not considered to be valid means for evaluating cardiovascular fitness in nonhandicapped persons. Thus one must question their validity with mentally retarded persons.

Indeed, in a recent adaptation to the Canadian Fitness Test for trainable
<table>
<thead>
<tr>
<th>Study</th>
<th>Mode of exercise</th>
<th>Test protocol</th>
<th>Subjects</th>
<th>Author comment on test administration or motivation of subjects during testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sengstock (1966)</td>
<td>Running-field</td>
<td>600-yd walk-run (AAHPER Youth Fitness Test)</td>
<td>EMR boys, n = 30; younger non-MR boys, n = 30; older non-MR boys, n = 30</td>
<td>No indication that test scores were affected by motivation of subjects</td>
</tr>
<tr>
<td>Rarick et al. (1970)</td>
<td>Running-field</td>
<td>300-yd walk-run (modified AAHPER Youth Fitness Test)</td>
<td>EMR boys &amp; girls, n = 4, 235; ages 8–18</td>
<td>300-yd walk-run considered to be less motivationally demanding than 600-yd walk-run</td>
</tr>
<tr>
<td>Bar-Or et al. (1971)</td>
<td>Treadmill walking</td>
<td>Maximal walking inclined treadmill, 3 progressive protocols, varying time increments</td>
<td>MR children: males, n = 89; females, n = 72; ages 6–15</td>
<td>15% of subjects could not complete test; mainly attributed to refusal to complete the test</td>
</tr>
<tr>
<td>Thoren (1971)</td>
<td>Cycle ergometry</td>
<td>Protocol not indicated</td>
<td>TMR &amp; EMR boys age 14; TMR children: males, n = 606; females, n = 449; ages 6–19</td>
<td>No comment</td>
</tr>
<tr>
<td>Londeree &amp; Johnson</td>
<td>Running-field test</td>
<td>300-yd walk-run (modified AAHPER Youth Fitness Test)</td>
<td>TMR children: males, n = 606; females, n = 449; ages 6–19</td>
<td>300-yd walk-run not indicated to be affected by motivational problems</td>
</tr>
<tr>
<td>Study</td>
<td>Mode of exercise</td>
<td>Test protocol</td>
<td>Subjects</td>
<td>Author comment on test administration or motivation of subjects during testing</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------</td>
<td>---------------------------------------------------------</td>
<td>-----------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Maksud &amp; Hamilton (1974)</td>
<td>Cycle ergometry</td>
<td>Maximal cycling</td>
<td>EMR boys, n = 62, ages 10–13</td>
<td>Doubtful that poor motivation was cause of early test termination</td>
</tr>
<tr>
<td>Coleman et al. (1976)</td>
<td>Cycle ergometry</td>
<td>Submaximal cycling PWC170 (Modified Sjostrand, 1947)</td>
<td>TMR &amp; EMR adults, males, n = 37</td>
<td>No comment</td>
</tr>
<tr>
<td>Rarick et al. (1976)</td>
<td>Cycle ergometry</td>
<td>Submaximal cycling PWC170 (Sjostrand, 1947)</td>
<td>EMR &amp; non-MR children: boys, n = 207; girls, n = 199; ages 6–13</td>
<td>No comment</td>
</tr>
<tr>
<td>Andrew et al. (1979)</td>
<td>Treadmill running</td>
<td>Maximal running (modified Balke, 1959); submaximal cycling PWC170</td>
<td>MR adults: males, n = 18; females, n = 2; ages 14–34</td>
<td>No comment</td>
</tr>
<tr>
<td>Beasley (1982)</td>
<td>Running-field test</td>
<td>12-min walk-run (Cooper, 1968)</td>
<td>TMR &amp; EMR adults, n = 30, males &amp; females</td>
<td>No comment</td>
</tr>
<tr>
<td>Bundschuh &amp; Cureton (1982)</td>
<td>Cycle ergometry</td>
<td>Submaximal cycling PWC170</td>
<td>EMR adolescents, n = 14, ages 11–19</td>
<td>Possible poor motivation during testing</td>
</tr>
<tr>
<td>Tomporowski &amp; Ellis (1984a)</td>
<td>Treadmill running</td>
<td>Maximal running (Balke &amp; Ware, 1959)</td>
<td>Moderately &amp; profoundly retarded adults, n = 65, males &amp; females</td>
<td>No comment</td>
</tr>
<tr>
<td>Coleman &amp; Whitman (1984)</td>
<td>Running-field test</td>
<td>300-yd run (Special Fitness Test for the Mentally Retarded, AAHPER 1968)</td>
<td>Mildly &amp; moderately retarded adults: males, n = 7; females, n = 10; ages 21–39</td>
<td>No comment</td>
</tr>
<tr>
<td>Study</td>
<td>Mode of exercise</td>
<td>Test protocol</td>
<td>Subjects</td>
<td>Author comment on test administration or motivation of subjects during testing</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------</td>
<td>---------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Tomporowski &amp; Ellis (1985)</td>
<td>Treadmill running</td>
<td>Maximal running (Balke &amp; Ware, 1959)</td>
<td>Severely &amp; profoundly retarded adults, n = 35 males &amp; females</td>
<td>Subjects appeared quite responsive to exercising on the motorized treadmill</td>
</tr>
<tr>
<td>Reid et al. (1985)</td>
<td>Stepping-field test</td>
<td>Submaximal stepping Modified Canada Home Fitness Test (Canadian Standardized Test of Fitness 1981)</td>
<td>TMR &amp; EMR adults: males, n = 105; females, n = 79; ages 20–39</td>
<td>Problems of cadence adherence during stepping exercise; difficulty in completing higher workloads during stepping exercise</td>
</tr>
<tr>
<td>Schurrer et al. (1985)</td>
<td>Treadmill walk-run</td>
<td>Maximal walk-run</td>
<td>TMR adults: males, n = 2; females, n = 3; age 25.2</td>
<td>No comment</td>
</tr>
</tbody>
</table>
retarded persons (Government of Canada, 1985), the running distances have been greatly extended (600 meters for 7- to 9-year-olds, 1200 meters for 10- to 12-year-olds, and 2000 meters for 13-year-olds and over). While longer distances are justifiable from a physiological perspective, all field tests require data-based validation. Thus, studies investigating the relationship between cardiovascular field tests and direct measurement of oxygen uptake with mentally retarded persons appear warranted and overdue.

Test Termination and Cadence Adherence

Eighteen studies have used various procedures to predict the cardiovascular fitness of retarded persons (Andrew et al., 1979; Bar-Or et al., 1971; Bundschuh & Cureton, 1982; Coleman et al., 1976; Coleman & Whitman, 1984; Londeree & Johnson, 1974; Maksud & Hamilton, 1974; Nordgren, 1970, 1971; Rarick, Dobbins, & Broadhead, 1976; Rarick, Widdop, & Broadhead, 1970; Reid et al., 1985; Schurrer et al., 1985; Sengstock, 1966; Thoren, 1971; Tomporowski & Ellis, 1984a, 1985; Tomporowski & Jameson, 1985). Only four of them (Bar-Or et al., 1971; Nordgren, 1971; Reid et al., 1985; Tomporowski & Jameson, 1985) commented on the applicability of the selected mode of exercise and protocol used (see Table 1). Nordgren (1971) encountered difficulty with 16 of 63 mildly and moderately retarded subjects in determining physical work capacity on a cycle ergometer. Problems included the inability of subjects to follow the required cadence and difficulty in completing more than one work load, which made extrapolation difficult.

Reid et al. (1985) noted similar problems with a step test and reported a loss of 36 of 220 mildly and moderately retarded subjects, many of whom experienced difficulty in attaining and maintaining required stepping cadences. Recommended cadences for gender and age groups were rarely attained by the subjects, and actual stepping rates were recorded. Also, motivating the subjects proved problematic as 45% stopped stepping prior to reaching target heart rates. Modified procedures of the step test, therefore, required these subjects to continue the test until they could no longer increase their stepping rates despite all possible instructional interventions by the trained evaluators. Actual stepping rate recordings necessitated the authors to interpolate oxygen requirements in order to predict maximal oxygen uptake (VO₂ max) from a regression equation. However, this modified version of the step test was shown to be a reliable (r = .84) indicator of cardiovascular fitness.

Finally, Tomporowski and Jameson (1985) provided further support for the notion that retarded persons experience difficulty in maintaining predetermined cadences. Behavioral data on 19 severely and profoundly retarded adults revealed that few subjects were able to maintain a regular cadence while exercising with cycling and rowing ergometers. Most subjects performed quick bursts of pedaling or rowing. Furthermore, performance deteriorated following increases in work demand.

1There are also established time criteria for 200-meter portions of the running. The failure to meet these criteria results in termination of the test.
Bar-Or et al. (1971) indicated four reasons why 15% of the mentally retarded children performing maximal graded treadmill walking did not complete the test to exhaustion. Causes of noncompletion included refusal to use the mouthpiece, poor coordination, refusal to walk, and refusal to complete the walk.

It is not clear why so few studies report problems in test administration or completion. Were subjects in the other studies higher functioning or pretested/selected in some manner to eliminate difficulties? Or did Nordgren (1971), Reid et al. (1985), and Tomporowski and Jameson (1985) not provide sufficient orientation or precise enough teaching of the testing protocol to prevent problems? Bar-Or et al. (1971) suggested that there was a striking relationship between lack of completion and place of residence (i.e., institution) and that level of IQ may be a contributing factor. Researchers should carefully chronicle the characteristics and selection criteria of the subjects as well as the techniques employed to facilitate adjustment to the testing methodology. For example, physical prompting appears to be a useful method to promote test adjustment (Tomporowski & Ellis, 1984b). If subjects still present problems in test administration, then such problems should be articulated so that the data are interpreted by others in that light.

With the exception of Bar-Or et al. (1971), studies using the treadmill have not identified any apparent problems with the protocol, possibly because the cadence is mechanically determined. Tomporowski and Jameson (1985) described treadmill running as the most efficient exercise for the mentally retarded since the treadmill moves at a set speed and allows precise control of pace and ensures continuous exercise. However, according to Shephard (1969), treadmill walking and running are acquired skills with appreciable components of learning and habituation. The issue of learning is dealt with in the next section.

Although the studies listed in Table 1 generally demonstrated inferior cardiovascular fitness levels for retarded persons compared to nonhandicapped individuals, it remains unclear whether the observed scores are entirely due to poor cardiovascular fitness per se or the inability of subjects to perform optimally on the test. Therefore, a closer look at the underlying assumptions and procedural requirements of these tests is warranted. These include motivation and learning, and physiological efficiency.

Motivation and Learning

A subject is expected to be equally motivated on a cardiovascular test across trials and there should be no systematic fluctuation of scores. However, several studies have reported that motivating the subjects can be a problem (Bundschuh & Cureton, 1982; Montgomery & Reid, 1985; Nordgren, 1971; Reid et al., 1985; Tomporowski & Jameson, 1985). Thus, is the predicted oxygen uptake score a function of the cardiovascular system or the motivational status?

When subjects are familiar with the test apparatus, repeated tests of cardiovascular fitness using a given protocol should result in similar scores. How-

---

2Efficiency was not assessed by typical physiological procedures but rather used as a descriptive term by the authors.
ever, there is some evidence that this may not occur with retarded subjects. Montgomery and Reid (1985) noted extreme variability in some of their subjects on the Léger Test (Léger & Lambert, 1982) when a baseline of performance was determined. Furthermore, Andrew et al. (1979) reported VO\textsubscript{2} max values (ml • kg\textsuperscript{-1} • min\textsuperscript{-1}) of two control subjects who decreased their cardiovascular functioning over 12 weeks from 43.5 to 25.7, and 50.5 to 40.0. Conversely, Scherrer et al. (1985) indicated that one of their five subjects showed an increase in VO\textsubscript{2} max from 20.6 to 42.3 as a function of the exercise program. It is difficult to attribute such changes to physiological mechanisms, and one must question if the tests were providing valid indicants of cardiovascular fitness for these individuals. Such extreme fluctuations in scores may more accurately represent changes in motivation rather than physiological functioning.

An increasing trend across trials might be interpreted as learning or accommodation. Indeed, learning was a critical dimension in the early studies evaluating the relative efficacies of treadmill tests, step tests, and cycle ergometer tests (Andersen & Hart, 1963; Rode & Shephard, 1973; Shephard, 1969; Zarraugh, Todd, & Ralston, 1974). Learning is undesirable\textsuperscript{3} because it confounds the internal validity of a study. Investigations that examine the validity of cardiovascular tests with mentally retarded persons should assess the extent of learning over several administrations of the test. Physiological efficiency, our last concern here, is related to learning.

Physiological Efficiency

Indirect tests that predict oxygen consumption from work performance or heart rate assume a constant efficiency. Poor efficiency, or a greater energy expenditure for a given work demand, can be attributed to such factors as poor coordination or lack of experience with a task or piece of equipment. If mentally retarded persons are not efficient exercisers compared with standards for nonhandicapped persons, then group comparisons using identical prediction equations are fraught with difficulties.

Recent evidence demonstrated that mentally retarded persons were less efficient than nonhandicapped persons when performing a stepping exercise. Prompted by the difficulties in cadence adherence in Reid et al. (1985), Seidl (1986) compared the physiological efficiency of stair stepping of educable and trainable mentally retarded women to nonhandicapped women using a protocol similar to the step test in the Canadian Standardized Test of Fitness (Government of Canada, 1981). The completion rate improved after 4 days of stepping practice. There was a 20% increase in the number of retarded women who were capable of completing the third and fastest stepping rate of the step test. Mentally retarded women as a group were least efficient at the fastest stepping rate. Furthermore, these women stepped less efficiently across all stepping rates (15.7%) compared to nonhandicapped women (17.1%). However, the stair-stepping effi-

\textsuperscript{3}As a reviewer pointed out, learning in itself is not undesirable. However, test results used to evaluate exercise programs cannot be clearly interpreted unless there is a plateau in learning.
ciency of mentally retarded women was only 0.3% less than the assumed 16% stepping efficiency for the general population.

Implications of these findings relate directly to the questioning of valid measures of fitness for retarded persons. No studies listed in Table 1 using indirect measurement of VO₂ max verified assumptions of exercise efficiency. In Seidl (1986), poor stepping efficiency at faster cadences lends support to previous reports of early test termination with increasing work rates. Improved test completion over repeated trials may be evidence of increased motivation or learning.

**Summary and Recommendations**

Recently a number of studies have investigated the cardiovascular fitness levels of persons with mental retardation. As noted in Table 1, a variety of test protocols have been used. It seems researchers have assumed that cardiovascular tests developed for nonhandicapped populations are appropriate for mentally retarded persons; however, no published study was identified that has assessed the validity of the cardiovascular testing methodology, and only two studies reported reliability coefficients for the testing schemes employed (Bundschuh & Cureton, 1982; Reid et al., 1985).

Most cardiovascular fitness tests are built around the basic premise that there is a linear relationship between heart rate, oxygen uptake, workload, and energy expenditure. While this relationship is not based on intellectual level, cardiovascular fitness tests with the mentally retarded population have problems with variations in testing methodology, difficulties of test termination, cadence adherence, variable performance by the participant because of learning or motivational changes, and variable efficiency. These influences can undermine the validity of the cardiovascular measures and thus cast doubt on the studies, which depend upon accurate and precise measures to determine the effectiveness of an exercise program.

It is recommended that further study address the issues raised and the extent to which they impinge upon the validity of cardiovascular measures of the individual with mental retardation. Indeed, systematic validity studies are greatly needed. Also, determining a reliability estimate for the selected measure should become integral to all research studies. Furthermore, greater detail of subject selection, orientation to the test protocol, and administrative difficulties should be presented in publications. Finally, there is sufficient evidence of variable performance by mentally retarded persons to question the typical pre/post research design. It would seem desirable that multitesting occur prior to and after an exercise program or that researchers adopt a single-subject time-series approach (Watkinson & Wasson, 1984). This latter method need not be restricted to one subject per study; rather, each of the subjects can be tested on the dependent measure on a number of occasions.

Although programming cardiovascular exercise for mentally retarded persons is on the upswing, there are still some serious measurement issues that must be addressed with regard to cardiovascular assessment. Exercise program effectiveness can only be determined with valid and reliable measures for the population in question.
References


Acknowledgments

We wish to thank Gail Webster, Garth Tymeson, and Bo Fernhall for their important input. Also, the financial support of Fitness Canada is gratefully acknowledged.