DOCUMENT RESUME

ED 103 414 SP 009 021

AUTHOR Bolonchuk, W. W.
TITLE A Critical Analysis of the AAHPER Youth Fitness Test.
SPONS AGENCY Louis W. and Maud Hill Family Foundation, St. Paul, Minn.
PUB DATE 71
NOTE 42p.

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Athletic Programs; Athletics; Elementary Education; *Motor Development; *Physical Education; *Physical Fitness; Reliability; Running; Secondary Education; Secondary School Teachers; Standards; *Test Validity
IDENTIFIERS *AAHPER Youth Fitness Test

ABSTRACT

This document focuses on the question of validity of the AAHPER Youth Fitness Test. Major emphasis of the study is on comparisons between the original test and the UND Revision. Part 1 provides an introduction to the test battery and considers the reliability of the test items and the establishment of validity based on comparisons between the criterion measure and the resulting score. Part 2 discusses the test situation providing the basis for this study. Sections include general test procedure, how the subjects were chosen, the tests administered, and the experimental design and statistical hypothesis of the test. Part 3 presents the results of the two tests including reliability and validity estimates. Part 4 contains conclusions on the reliability of the tests and compares the AAHPER and the UND revision tests. The appendixes attached include materials sent to the subjects, the score card used, and a list of references. (JS)

A Critical Analysis of The AAHPER Youth Fitness Test

by

W. W. Bolonchuk

A Summer Research Project sponsored by a
Hill Family Research Grant

1971

Table of Contents

Introduction

The AAHPER Youth Fitness Test was developed by a special committee of the AAHPER Research Council in 1957. It is a practical test which is easily administered. The purpose of each test is stated and the test procedure is objectively described and illustrated. Norms which have been established on the basis of nationwide sampling are included in the test booklet (1). These factors make the test feasible and acceptable for nationwide use as an integral part of physical education since 1957.

The test battery included seven items which purport to measure the following criteria: pull-up for judging arm and shoulder girdle strength; sit-up for judging efficiency of abdominal and hip flexor muscles; shuttle run for judging speed and change of direction; standing broad jump for judging muscle power of the leg extensors; 50 yard dash for judging speed; softball throw for distance for judging skill and coordination; and 600 yard run-walk for judging cardiovascular efficiency (1).

The reliability of the test items has been examined by several investigators (14, 11, 16, 5, 15, 22, 10, 2, 23, 13). Similarly the validity has been scrutinized by (11, 10, 14, 23, 13). Whereas the reliability is acceptable within the framework of the published reports the validity appears to be less substantial. In selecting the seven-item battery, the planning committee agreed that the items represented specific measures of the components of physical fitness thus accenting face validity for the battery.

The establishment of the validity of a test involves a comparison between the criterion measure and the resulting score (14, 7, 4, 17, 18). For example, the criterion measure for the pull-up test is defined (1) as arm

and shoulder girdle strength. This is an often used criterion for that test. However, if the test is administered to yield a score which is the maximum number of continuous repetitions of the pull-up exercise, then the score is not a valid measure of that stated criterion. Why? The answer is found in the definition of the term strength which was used in stating the criterion measure for the pull-up test.

General agreement for the definition of strength is "The amount of force which a muscle is able to exert to overcome resistance". This definition assumes movement, and thus refers to dynamic muscular strength. The most widely accepted measure of dynamic muscular strength is that test which was used by Delorme (9) to measure the maximum amount of weight lifted in one repetition or 1-RM. If the purported measurement of strength includes a duration of work beyond that which is required to complete on repetition, the measurement now includes muscular endurance or more specifically dynamic muscular endurance, Montoye et al (19). Thus it would appear that the criterion measure cited by Johnson and Nelson (14) and AAHPER (1) should indeed be defined as a measure of dynamic muscular endurance. Johnson and Nelson (14) do exactly this in another section of their text. This example of apparent inconsistancy in identifying the criteria measures for test items is not uncommon. McCloy and Young (18) identify pull-ups and push-ups as tests of muscular strength in a chapter describing strength tests and then describe the same exercises as tests of muscular endurance in a chapter describing endurance tests.

These examples serve to focus on the inconsistencies which exist in the terminology specifically related to the AAHPER Youth Fitness Test and are only an example of inconsistancies related to the prescription of exercise by physical education teachers. The scope of this study however will

focus on the seven items of the fitness test to illustrate an apparent solution.

If a test has proven to be invalid then the validity may be improved in one of two ways. Either the test procedure must be changed so that the score is indicative of the measure that it purports, or the criterion measure must be changed to suit the procedure used to obtain the score. Thus, in the pull-up test if the criterion measure is to be the measurement of dynamic muscular strength, the resistance to the arm flexion movement should be increased to the maximum amount so that the subject could complete only one repetition of the pull-up exercise. This alteration in test procedure would ensure a valid measure of the criterion of dynamic muscular strength. On the other hand should it be advantageous to leave the pull-up test procedure as prescribed in the AAHPER test, then the criterion measure should be altered to be consistent with the procedure.

The matter of altering the criterion measure must carefully be considered in lieu of compounding the error. Generally there are three types of criteria measures that may be established for tests of physical and motor fitness. These are:

1. a definition of the performance task specifically related to the test, for example, the number of continuous pull-ups that a subject can perform.
2. the specific factors which are involved in the performance of the test; for example, dynamic muscular strength, static muscular strength, circulorespiratory endurance, flexibility, dynamic or static muscular endurance and so on.
3. the amount of calculated work that the subject has performed in completing the test.

A statement of the criterion measure in terms of type one or three above is specific to the performance involved in the test, whereas type two is specific to muscle function. Thus, in consideration for the practicality of a test the type two criterion measure is less feasible, even though more desirable. Type one on the other hand is so specific to the activity performed that it offers little value in terms of a subjects' general level of fitness. Type three seems to be the compromise between a criterion measure of one and two since it is descriptive of the performance task, and also allows for some generalization about the specific factors of the type two criterion measure.

When the AAHPER pull-up test is compared with the definitions cited for dynamic muscular strength and dynamic muscular endurance it seems to contain elements of each. The body weight of the subject may be considered as the resistance to arm flexion (an element of dynamic muscular strength); if the resistance is overcome several times, (10 completed pull-ups) then the criterion of dynamic muscular endurance is included because of the duration of the muscular contraction. These two factors of resistance and repetitions plus the distance that the body weight is displaced provide the basis for the type three criterion measure as follows:

Work = Resistance X Distance X Repetitions

Where: Resistance = Body weight in pounds

Distance = Amount of body displacement in inches

Repetitions = A count of completed pull-ups

The score for the pull-up test is now sensitive to the amount of work that the muscles are doing during the test.

Each component of the equation is represented equally so that if the body weight or the distance or the number of repetitions is altered the work

index will reflect that alteration. Thus a criterion measure based on the work performed would appear to combine the most accurate criterion measure for the pull-up without limiting the scope of generalizations to the performance of the item (type 1 criterion measure) yet be free from the stringent definitions of strength and endurance (type 2 criterion measure).

Ricci (21) illustrated hypothetically that the amount of work performed by a subject doing the pull-up test may be disproportionate to the number of pull-ups which are performed. For example:

Subject	Body Weight(Kgs)	Pull-Ups	Displacement(Meters)	Work(Push-up in Kilogram-meter)
A	50	3	1.5	75(50 X 1.5)
B	100	3	3.0	300(100 X 3)

According to the scoring method employed in the AAHPER test, the performance by student A would be equal to that of student B since both had completed three continuous pull-ups. Yet, student B has performed more work by reason of displacing a greater resistance (100 kg - 50 kg = 50 kgs of body weight) for each repetition. Thus, student A has not demonstrated superior muscular strength nor muscular endurance. If we compare the two performances to the definitions of strength and endurance we find that student B has overcome more resistance and by reason of displacing that resistance (3.0 meters - 1.5 meters = 1.5 meters) over a greater distance he has probably sustained the contractions for a longer period of time. Thus, unless a criterion measure of pull-up performance is established for the pull-up test the validity of the test is in jeopardy. Whereas a criterion measure of work involved in the pull-up test appears to be more specifically related to the intended purpose of the test within the realm of the practical administration of the test.

A battery of tests intended to examine the validity of the AAHPER Youth

Fitness Test has been developed. The battery was identified as the UND Revision of the AAHPER Youth Fitness Test. The battery employs the work (type 3 criterion measure) for the standing broad jump, the pull-ups, sit-ups and softball throw. The fifty yard dash, the 600 yard run-walk and the shuttle run tests will be examined with the revised test employing a type 2 criterion measure.

Norms will be established for the AAHPER test and for the revised test and the subjects' performance will be evaluated on the basis of each set of norms. Scores for the AAHPER test will be correlated with scores for the Revised test to determine the degree of relationship between the two performances. And, since the performance of the subjects may vary from test to test, thus introducing a source of experimental error, a test-retest reliability coefficient will also be determined.

Definitions of Terms

Dynamic Muscular strength - The capacity of a muscle to overcome resistance thus producing movement. A test of this quality to maximum is described by Delorme (9).

Dynamic Muscular endurance - The capacity of a muscle to sustain contraction, and produce movement without benefit of blood flow. Duration of this type of activity is short.

Circulo-respiratory endurance - The capacity of the heart, lungs, arteries, veins to extract oxygen from the atmospheric air and deliver it to the muscles then buffer lactic acid and get rid of carbon dioxide.

Power - is the amount of work which can be accomplished per unit of time.

AAHPER - American Association for Health, Physical Education and Recreation.

METHODOLOGY

Sampling Procedure

A non-probability sample of 189 fifth and sixth grade boys and girls were identified for use as subjects in this project. These students lived in the area adjacent to the University of North Dakota campus which would be convenient for testing purposes. Also, since they were all elementary school students it was presumed that a considerable number would be available at the time of testing.

Since norms were to be established it would be desirable to have data from about 100 subject.

A list of names for all fifth and sixth grade boys and girls was obtained from the principals of Lake Agassiz and West elementary schools. Each subject and/or parent was interviewed by telephone (see interview guide Appendix I). Of the 182 subjects who were contacted by telephone, 81 agreed to take part in the study. Of these 81 subjects 45 (25 girls and 20 boys) completed all items of both tests. Table 1 shows the breakdown of the sampling results.

TABLE I: RESULTS OF THE TELEPHONE INTERVIEW OF THE 5TH AND 6TH GRADE BOYS AND GIRLS REQUESTING THEIR PARTICIPATION IN THE SUMMER RESEARCH PROJECT.

	N	%
Telephone calls made to Prospective subjects	182	100
Unable to participate In the study	60	33
No answer to the Telephone call	41	23
Accepted participation In the study	81	44
Completed both Tests	45	25

If a contact was not made on the first call, then a second call was made. If no contact was made on the second call, the interviewer recorded the result as a "no answer", and that prospective subject was not called again. Forty-one tallies were recorded for this reason. The following reasons were noted for those subjects who were unable to participate:

1. moved.
2. vacations out of town.
3. conflicts with other scheduled activities.
4. did not want to participate.

A letter (Appendix III) describing the purpose of the study and announcing several important dates in connection with the study was mailed to each consenting parent and child. As a follow up to this letter, a demonstration of the test items and procedure was held in the Fieldhouse on the University of North Dakota campus for interested parents and children. This demonstration was held in the evening on June 17.

The Tests:

The AAHPER Youth Fitness Test was administered as it is described and illustrated in the AAHPER publication (1), on Wednesday, June 23. The UND Revision of the AAHPER test was administered on the second test day which was Friday, June 25. The revised test included seven items whose purposes paralleled the seven items of the AAHPER test. These items were:

1. pull-up (or vertical hang)
2. sit-ups
3. standing broad jump
4. dodging run
5. fifty yard dash with a running start
6. softball throw
7. six minute run

The directions for administration of the AAHPER tests were used to administer the pull-ups, sit-ups, standing broad jump and the softball throw of the revised test. The scoring for these tests was changed from

the performance procedure used by the AAHPER to the work index method.

The description of this scoring method and the other items of the revised test are presented in the following section.

UND REVISION OF THE AAHPER YOUTH FITNESS TESTS
Description of the Test Items

1. Pull-Up Test (Boys)

PURPOSE:
To measure the work involved in performing the null-up exercise to a voluntary maximum.

CRITERION MEASURE:
The amount of work accomplished during the exercise phase of the pull-up test.

1. The subject assumes the arms extended position.

2. The subject is instructed to complete as many continuous pull-ups as possible.

3. The test is terminated when the subject cannot continue the pull-ups at the rhythm established in his repetitions.

4. The number of continuous excursions from extended to flexed positions is counted and recorded as the number of repetitions. Only the completed repetitions are counted.

SCORING:
The number of repetitions, the distance that the subject's body was lifted and the subject's body weight will be used to calculate an index of work performance by substituting into the following equation:

$$\text{Work Score} = \text{Body Weight (lbs.)} \times \frac{\text{Distance (ins.)}}{12} \times \text{Reps.}$$

Flexed Arm Hang (Girls)

1. The procedure for this test will include the identical steps involved in the boys pull-up measurement.

2. The test will be performed as it is described in the AAHPER test manual.

3. The duration of the hang is recorded in seconds to the nearest second.

SCORING:
The duration of the hang in seconds, the distance that the subject is lifted from the extended arm to the flexed arm position, and the body weight will be used to calculate a work index by substituting into the following

equation:

$$\text{Work Score} = \text{Body Weight(lbs)} \times \frac{\text{Distance of one pull-up(ins)}}{12} \times \text{Duration of the hang in Seconds}$$

2. Sit-Up Test

PURPOSE:
To measure the work involved in performing the sit-up exercise to a voluntary maximum.

CRITERION MEASURE:
The amount of work accomplished during the exercise phase of the sit-up test.

PROCEDURE:
1. Trunk extension position
   1.1 The subject lies on his back with knees straight and legs together, hands behind the head with fingers interlaced.
   1.2 A partner kneels at the subjects feet and holds both ankles of the subject to keep the heels of his feet in contact with the mat.

2. Trunk flexion position
   2.1 The subject brings his elbows forward and curls up to his knees until his right elbow touches the left knee, the left elbow touches the right knee.
   2.2 The subject returns to the trunk extension position; this is determined when the head touches the mat.

3. The subject is instructed to complete as many sit-ups as possible up to a maximum of 100 for boys and 50 for girls.

SCORING:
The number of continuous excursions from the extended to the flexed trunk positions is counted and recorded. The score for the test is calculated by substituting into the following equation:

$$\text{Work Score} = \text{Upper Body Weight(lbs)} \times \frac{\text{Distance from the Supine Lying to Sit-Up (ins)} \times \text{Reps.}}{12}$$

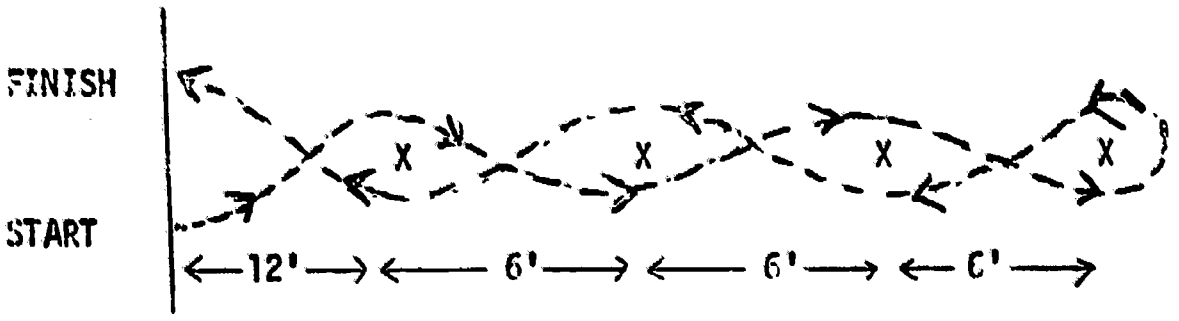
3. Dodging Run Test (14)

PURPOSE:
To measure the agility of the performer in running around four stationary objects.

CRITERION MEASURE:
Time for the run in seconds and tenths.

PROCEDURE:
1. The subject starts behind the starting line on the signal "go" and runs a "figure 7" course around each of the four chairs and returns in the same pattern until he crosses the starting line.

2. Outline of the course:



SCORING:
The score for each performer is the length of time in seconds, to the nearest tenth of a second, to complete the course.

4. Standing Broad Jump Test

PURPOSE:
To measure the work involved in the standing broad jump exercise performed to a voluntary maximum.

CRITERION MEASURE:
The amount of explosive work accomplished in the maximum performance of the standing broad jump exercise.

PROCEDURE:

1. The subject stands behind the take-off line, swings the arms backward and bends the knees, then jumps as far forward as possible.

2. The usual directions for this event apply.

3. Three trials are allowed.

SCORING:
The distance is measured from the take-off line to the point of contact by the heel or other part of the body that is nearest to the take-off line. The distance of all three trials is recorded in feet and inches to the nearest inch. The score for the test is calculated by substituting into the following equation:

$$\text{Work Score} = \text{Body Weight(lbs)} \times \frac{\text{Distance of the jump (ins)}}{12}$$

5. 50 Yard Dash Test

PURPOSE:
To measure the maximum speed over a distance of 50 yards.

CRITERION MEASURE:
The time in seconds and tenths as a measure of speed.

PROCEDURE:
1. Two subjects will run at the same time.

2. The subjects start behind a starting line which is sixty (60) yards from the finish line.

3. On the signal "go" both runners sprint toward the finish line, as they cross the 10 yard line (50 yards from the finish line) another "starter" will signal the timers at the finish line, with an arm signal, to start their stopwatches. Thus, the sprinters will have run 10 yards before the stopwatches are started. This will eliminate the reaction to the "go" single and the gathering of momentum which is measured in other tests of speed.

4. The stopwatch is stopped when the runner breaks the vertical plane of the finish line with any part of his body.

SCORING:
The score is the elapsed time in seconds to the nearest tenth of a second for the fifty yard run.

6. Softball Throw for Distance Test

PURPOSE:
To measure the work involved in throwing a softball for distance, performed to a voluntary maximum.

CRITERION MEASURE:
The amount of work accomplished in the maximum performance of the softball throw exercise.

PROCEDURE:
1. The subject must throw the softball from within a 6 foot restraining area which is drawn parallel to the five yard field markers.

2. Three trials are recorded in feet.

3. The distance of the throw is estimated from the throwing line to the first point of contact with the ground.

SCORING:
The distance of the throw, and the weight of the ball will be used to calculate an index to represent the score for this test. Thus:

Work Score = Wt. of the ball (lbs) x Distance of the Throw (ft)

7. Six Minute Run-Walk Test

PURPOSE:
To measure circulo-respiratory fitness.

CRITERION MEASURE:
Circulo-respiratory Fitness

PROCEDURE:

1. One half of a group will run at a time (if the group has N = 10 subjects five will run the test) the other half of the group will count the number of laps that his "buddy" has completed.

2. The subjects for this test will be located at 1/4 lap intervals to avoid congestion on the track during the run; thus there will be four starting points on the track.

3. To begin the test the subjects and their buddies will proceed to their respective starting lines. On the signal "go" the runners will begin.

4. The counters will remain at each of the four starting lines and count the number of laps that his buddy completes at this point.

SCORING:
The number of laps of the track.

Anthropometric Test Description

In order to calculate the work score for each test it was necessary to determine certain anthropometric measures. These were:

1. sitting height
2. standing height
3. standing body weight
4. lying body weight
5. pull-up
   - 5.1 Extended arm position
   - 5.2 Flexed arm position

1. Sitting Height

The subject sat on a 20" bench with his back to a wall stadiometer. The feet rested on a platform so that the thighs were approximately parallel to the floor. The subject was instructed to sit tall with the back of the head resting against the stadiometer. A 90° angle board was placed on the top of the subject's head and along the

stadiometer. The measurement was read and recorded to the nearest one quarter inch.

2. Standing Height

The subject stood with his back to a linen measuring tape which was hung on a wall. The subject was instructed to place the heels against the wall, to stand tall and to keep his head against the wall. A 90° angle board was placed on the top of the subject's head and along the tape. The measurement was read and recorded to the nearest one quarter inch.

3. Standing Body Weight

The subject stood in the center of the pressure plate of an upright Toledo Scale. The measurement was read and recorded to the nearest pound.

4. Lying Body Weight

The subject reclined in the supine position on a bench which was placed within eight inches of the Toledo Scale. The hips of the subject were placed within one inch of the end of the bench and the upper body rested on the pressure plate of the scale. The subject raised his arms to head level and put his hands behind his head and interlaced his fingers, to simulate the sit-up starting position. The measurement of lying body weight of the upper body was read and recorded to the nearest pound.

5. Pull-Up

5.1 Extended Arm Position

The subject assumed a hanging position from a horizontal bar. The bar was grasped with an overhand grip. The arms were extended at the elbow; the feet hung free of the floor or any other support. A measurement from the Greater Trochanter of the femur was read by sighting a

right angle board from the Greater Trochanter to a scale on the vertical support for the pull-up bar. This measurement was recorded in inches to the nearest one quarter inch.

5.2 Flexed Arm Position

After the extended arm position was measured, the subject pulled-up to a flexed arm position with the chin placed over the bar and the head held in a horizontal position. A measurement from the Greater Trochanter of the femur was read by sighting a right angle board from the Greater Trochanter to a scale on the vertical support for the pull-up bar. This measurement was recorded in inches to the nearest one quarter inch.

5.3 Difference

The difference between the extended arm and flexed arm positions was recorded. This difference represented the body displacement in a vertical plane.

General Test Procedure

The subjects were scheduled to arrive at the Fieldhouse for testing at twenty minute intervals. To facilitate this schedule, appointment forms (Appendix II) were constructed and mailed to each subject, one week prior to the Wednesday, June 23 test. The appointment form assigned times for each subject for both test days.

Registration for all subjects was held in the hallway on the second floor of the Fieldhouse. The following information was recorded on the subjects' score card at that time:

1. Name
2. Age in years
3. Sex
4. Birthdate

After the subject registered, he proceeded to the anthropometric test stations which were located in the weight training gym. These tests and the pull-ups were administered to each subject individually. The sit-ups, shuttle run, and the standing broad jump were administered in groups of six to eight subjects in a gym area adjacent to the weight training room. The 50 yard dash and the 600 yard run-walk were administered on the indoor track in the fieldhouse and the softball throw was administered on a playing field adjacent to the fieldhouse. All data was recorded on a score card (Appendix IV) which was constructed for this purpose.

Experimental Design

A single group, non-probability sample was employed in the study. The subjects were selected primarily for convenience and therefore are not considered representative of a population.

Inferences from the study were thus limited to a description of the results related to these subjects alone, not to any other individuals or groups. The sample was considered a demonstration group.

Since the reliability and validity of the AAHPER Youth Fitness Test was under study in this project a test-retest procedure was necessary. The Wednesday and Friday tests allowed for sufficient recovery time for the subjects between tests, and yielded sufficient data to fulfill the specific purposes of the study. The data in this study are numerical, continuous and of the interval type.

Since the design was a test-retest procedure, each subject was his own control thus minimizing intergroup error.

The Pearson Product-Moment Correlation was employed to calculate an estimate of the relationship between two tests which employ this type of datum.

The following hypothesis were established to test on the basis of the estimated correlations:

H_0 There was no relationship between the items of the test and the retest.

H_1 There was a relationship between the items of the test and the retest.

RESULTS AND DISCUSSION

Reliability Estimates for the Anthropometric Tests

The correlation coefficients for the Anthropometric tests which are illustrated in Table 1 show an acceptable reliability with the exception of the three pull-up measures. These measures proved to be the most difficult results to obtain. Whereas the standardization of the reference points was a simple matter (see Methodology) controlling the subject and taking the measurements proved to be more subjective than was desirable. It is important to note, however, that even thounh the reliability is low, the mean differences vary less than one half inch for any of the three items.

TABLE 1 : MEANS, STANDARD DEVIATIONS AND RELIABILITIES FOR THE ANTHROPOMETRIC TEST SCORES (N = 45, 25 GIPLS AND 20 BOYS)

No.	Test Item	Mean (1)	S.D.(1)	Mean (2)	S.D. (2)	r
1.	Standing Body Wt. (lbs)	88.7	16.488	89.0	16.344	1.00
2.	Standing Height (ins)	59.0	2.880	59.0	2.873	0.99
3.	Lying Body Weight (lbs)	30.1	1.552	30.0	1.594	0.90
4.	Sitting Height (ins)	37.8	6.270	38.3	6.655	0.97
5.	Pull-Up Flexed Arm (ins)	47.4	1.764	47.7	1.592	0.60
6.	Pull-Up Extended Arm (ins)	27.1	2.597	27.6	2.592	0.75
7.	Pull-Up Difference (ins)	20.3	2.083	20.1	1.992	0.80

The test-retest correlation for lying body weight was 0.90. Represented as a reliability coefficient this value is low, thus seriously questioning the reliability of the test. However, the mean difference between tests is less than one-tenth of a pound and since the raw scores were recorded to

the nearest one-quarter pound, the discrepancy between tests was accepted as reliable.

Reliability Estimates for the AAHPER Youth Fitness Test

It was not the intent of this study to re-exam the test reliability, instead the test-retest procedure was employed as a study control to determine the reliability of the scores produced in this study. Only four of the seven test items were included in the test-retest procedure. The four items, pull-ups (flexed arm hang), sit-ups, standing broad jump and the softball throw were used this way because they could be scored by the work formula (f x d) and validated on that basis. The other three items, 600 yard run-walk, 50 yard dash and the shuttle run required a different validating procedure.

TABLE 2: MEANS, STANDARD DEVIATIONS AND RELIABILITIES OF FOUR AAHPER TEST ITEM SCORES.

Girls (N = 25)	Mean (A)	S.D. (A)	Mean (B)	S.D. (B)	r
Flexed Arm Hand (sec)	14.1	11.337	14.0	9.929	.95
Sit-Up (number)	49.0	3.518	48.6	5.243	.99
Standing Broad Jump (ins)	62.6	6.720	63.1	6.837	.82
Softball Throw (ft)	67.6	22.931	59.3	17.346	.93
Boys (N = 20)					
Pull-Ups (number)	2.9	2.498	2.8	2.971	.92
Sit-Up (number)	84.2	24.348	84.0	24.694	.97
Standing Broad Jump (ins)	66.6	8.059	66.8	9.756	.89
Softball Throw (ft)	112.0	28.780	91.8	21.038	.94

(A) = The AAHPER Youth Fitness Test administered on June 23, 1971.
(B) = The UND Revision of the AAHPER Youth Fitness Test administered on June 25, 1971.

With the exception of the sit-ups for girls (r = .99), the reliability coefficients were not as high as would be expected for reliable measurement. However, in comparing the correlation estimates from Table 2 with those published by Fleishman (11), the present estimates are all higher which would indicate greater reliability for the present test items.

Validity Estimates for the AAHPER Youth Fitness Test

The 600 yard run-walk correlated inversely (r = -.60 boys and r = -.63 girls) with the six minute run. This would indicate that the two tests measure distinctly different criteria. This result seems to be in agreement with the explanation of Balke (3), that as the duration of an activity decreased the limiting factor to performance would be the anaerobic tolerance. The 600 yard run-walk test purports to measure cardiorespiratory endurance as does the six minute run-walk test. However, the duration of the two varies from a mean of one minute and thirty-five seconds for boys and one minute and fifty-nine seconds for girls in the 600 yard run-walk test to six minutes for the six minute run-walk test. This difference in duration of these tests may in fact result in different criteria measurements. Whereas the six minute duration may measure aerobic capacity the shorter duration of the 600 yard run-walk test may measure anaerobic capacity.

It is generally accepted that the best test of circulorespiratory endurance is the maximal oxygen consumption test. Both Balke (3) and Cooper (8) have developed field tests of maximal oxygen consumption which correlate highly with laboratory tests, thus establishing their validity. Cooper's twelve minute run-walk test was compared by Bolonchuk (6) with a six minute run-walk test to determine the relationship between these field tests. The resulting correlation was r = .90. Thus the six minute run-

walk test was employed in this study as a valid field test of circulorespiratory endurance. A test of circulorespiratory endurance must allow sufficient duration of a moderate intensity to allow muscular action to enhance circulation of the blood from the heart to the muscles, through the lungs and back again. Muscle contraction which exceeds a moderate intensity, tends to restrict blood flow. When the blood flow is restricted the build-up of metabolic waste particularly lactic acid increases rapidly and the duration of the exercise must be reduced. Thus, when a test involves short duration, high intensity muscular contraction, the capacity for energy production without sufficient oxygen, or anaerobic energy is being tested. Cardiorespiratory involvement in this type of exercise is limited, since the flow of blood is restricted. Thus this type of test involves less cardiorespiratory endurance and more muscular endurance.

The AAHPER 50 yard dash was compared with the 50 yard dash with a running start. The essential difference between the two tests was simply that the 50 yard dash with a running start was timed after the subjects had achieved sprinting speed. This procedure helped to eliminate the time involved in a subject's reaction to the starting command, moving from the stationary starting position and gaining momentum to sprinting speed. Since the criterion measure for the AAHPER 50 yard dash item indicates speed which is measured in time, it seemed logical to compare these two tests in an attempt to determine what effect the stationary as opposed to the running start would have on the performance.

The mean time for the 50 yard dash with the running start was less than the mean time for the AAHPER 50 yard dash. The correlations for these items were low indicating that the relationship was not a strong one. The positive correlations illustrate the similarity between the tests to measure

speed, however, their low values indicate that the stationary versus the running start did make a difference.

TABLE 3 : MEANS, STANDARD DEVIATIONS AND VALIDITY ESTIMATES FOR THE AAHPER YOUTH FITNESS TEST.

Boys Test	Mean (A)	S.D. (A)	Mean (R)	S.D. (R)	r
Pull-Up	2.8	2.971	401	418	.97
Sit-Up	84.0	24.094	7851	2208	.77
Standing B. J.	66.9	9.756	476	82	.64
Softball Throw	91.8	21.099	42	10	1.00
50 Yard Dash	8.0	0.567	6.9	0.6	.47
600 Yd. Run-Walk	135 sec.	25.516	9.4	1.0	-.60
Shuttle Run	10.6	0.904	7.4	0.5	.49
Girls Test					
Flexed Arm Hang	14.0	9.929	2054	1494	.90
Sit-Up	48.6	5.243	4664	1082	.33
Standing B.J.	63.1	6.837	468	104	.51
Softball Throw	50.3	17.348	27	8	1.00
50 Yard Dash	8.4	0.663	7.2	0.8	.75
600 Yd. Run-Walk	159 sec.	27.160	8.4	1.1	-.63
Shuttle Run	11.0	0.658	7.8	0.5	.55

(A) = The AAHPER Youth Fitness Test administered on June 23, 1971

(R) = The UND Revision of the AAHPER Youth Fitness Test administered on June 25, 1971.

A definition of agility includes the ability to move with maximum speed and change direction. Both the shuttle run and the figure eight run contain elements of agility. However, this element of agility is more definite in the shuttle run than in the figure eight test. A subject is required to come to a momentary stop to pick up the object in the shuttle run thus reducing speed, and then increasing speed to sprint to the next object, whereas the figure eight test allows for a more constant maintenance of speed throughout the test. The validity of the Dodging Run test has been reported by Gates and Sheffield (12) as .820. The reliability was estimated as .934 for boys and .802 for girls respectively. These values are acceptable standards for validity and reliability thus the figure eight test was used as a criterion test. The comparison between the two tests indicated a low positive relationship, r = .55 for girls and r = .49 for boys, illustrating that the tests are similar but not highly related.

The pull-up, sit-up, standing broad jump and softball throw were validated on the basis of work scores. The force and distance components of each test were identified and an attempt was made to measure each. The June 25 test scores for the four test items were converted to units of work expressed in foot pounds or in the case of the flexed arm hang, a work index. The raw scores were correlated with the work score to determine the degree of relationship.

A high positive correlation resulted between work scores and raw scores for the pull-up and flexed arm hang tests (see Table 3). Thus the raw score for these items appears to be closely related to the amount of work accomplished in this exercise. It is clear then that the subject with the greatest body weight and body displacement is accomplishing the greatest amount of work and also doing the greatest number of pull-ups.

One must accept this result with some reservation, however, since a number of zero scores were obtained for both boys and girls. These zero scores tend to bias the results in terms of the stronger subjects, who in this sample, were the bigger boys and girls.

The sit-up test comparisons resulted in low positive correlations (see Table 3). The sit-up test comparison for the girls was particularly low, r - .33. These estimates indicate that the sit-up raw score is not indicative of the work accomplished during the test. Thus, validity of the test in this case is questionable. A similar result is evident for the comparisons between the raw scores and the work scores for the standing broad jump. The correlations were positive, however, they are low, indicating questionable validity.

Work performance and the raw scores of the softball throw correlated as r = 1.00 for both boys and girls. Since the resistance (the weight of the softball) was constant, the calculation of the work is identical in magnitude to the score units (distance of the throw). Thus, although the numbers for the work and raw scores are different, the relative values are the same.

Norm Comparisons

Norms were constructed for the scores from the UND Revision of the AAHPER tests. Table 4 and 6 illustrate the norms for the raw scores taken from this test and Table 5 and 7 shows the work score equivalents for the raw score norms.

The profiles for two subjects were constructed on Tables 4 to 7 to illustrate the comparison between the raw score and work score status of

Table 4

UNIVERSITY OF NORTH DAKOTA
DEPT. OF HPER
Summer Research Project
AAHPER Youth Fitness Test Norms for Grade 5 and 6 Girls (N = 25)

JULY 1971

Standard Scale	Sit Ups No.	Flexed Arm Hang Sec.	Figure 8 Sec. & Tenths	Standing Broad Jump Ft. & Ins.	50 Yard Dash Running Start Sec. & Tenths	Softball Throw Ft.	Six Minute Test Distance in Laps of the F H Track	Standard Scale
100	0	0	0.0	0.0	0.0	0.0	0.0	100
95	0	0	0.0	0.0	0.0	0.0	0.0	95
90	0	0	0.0	0.0	0.0	0.0	0.0	90
85	0	0	0.0	0.0	0.0	0.0	0.0	85
80	0	0	0.0	0.0	0.0	0.0	0.0	80
75	0	34	0.0	0.0	0.0	93	0.0	75
70	0	30	6.9	6'2"	5.9	87	10.1	70
65	0	25	7.2	5'11"	6.2	80	9.7	65
60	0	22	7.4	5'9"	6.5	73	9.3	60
55	51	18	7.6	5'6"	6.8	66	8.8	55
50	49	14	7.8	5'3"	7.2	59	8.4	50
45	47	10	8.0	5'0"	7.5	52	8.0	45
40	44	6	8.3	4'10"	7.8	45	7.5	40
35	42	2	8.5	4'7"	8.2	38	7.1	35
30	40	0	8.7	4'4"	8.5	31	6.7	30
25	38	0	8.9	4'1"	8.8	24	6.2	25
20	36	0	0.0	3'11"	0.0	0.0	5.8	20
15	34	0	0.0	0.0	0.0	0.0	0.0	15
10	32	0	0.0	0.0	0.0	0.0	0.0	10
5	30	0	0.0	0.0	0.0	0.0	0.0	5
0	28	0	0.0	0.0	0.0	0.0	0.0	0
MEAN	49	14	7.8	5.03	7.2	59	8.4	MEAN
SD	5	9	0.544	0.068	0.827	17	1.090	SD
N	25	25	25	25	25	25	25	N
H.S.	50.000	33.000	8.800	6.020	8.800	90.000	10.000	H.S.
L.S.	25.000	0.0	7.000	4.010	6.100	30.000	6.000	L.S.
RANGE	25.000	33.000	1.800	2.010	2.700	60.000	4.000	RANGE
SSC	0.419	0.794	0.043	0.005	0.066	1.047	.087	SSC
NO. TRLS	1	1	2	3	1	3	1	NO. TRLS

Table 5

UNIVERSITY OF NORTH DAKOTA
DEPT. OF HPER
Summer Research Project

JULY 1971

Modification of the AAHPER Youth Fitness Test Norms for Grade 5 and 6 Girls (N = 25)

Standard Scale	Flexed Arm Hang Work Score	Sit Up Work Score	Shuttle Run Time - Sec & Tenths	Standing Broad Jump Work Score	50 Yard Dash Time - Sec & Tenths	Softball Throw Work Score	Six Minute Test Distance in Laps of the F H Track	Standard Scale
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100
95	0.0	0.0	0.0	0.0	0.0	0.0	0.0	95
90	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90
85	0.0	0.0	0.0	760	0.0	0.0	0.0	85
80	3871	7261	0.0	718	0.0	0.0	0.0	80
75	5043	5828	0.0	677	0.0	43	0.0	75
70	4445	6395	6.9	635	5.9	39	10.0	70
65	3247	5962	7.2	593	6.2	36	9.7	65
60	2250	5529	7.4	551	6.5	33	9.3	60
55	2652	5097	7.6	510	6.8	30	8.8	55
50	2054	4664	7.8	468	7.2	27	8.4	50
45	1456	4231	8.0	426	7.5	24	8.0	45
40	858	3798	8.2	384	7.8	21	7.5	40
35	261	3365	8.5	342	8.2	18	7.1	35
30	0.0	2933	8.7	301	8.5	14	6.7	30
25	0.0	0.0	8.9	0.0	8.8	11	6.2	25
20	0.0	0.0	0.0	0.0	0.0	0.0	5.8	20
15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
MEAN	2054	4664	7.8	467	7.2	27	8.4	MEAN
SD	1494	1081	0.544	104	0.827	7.829	1.090	SD
N	25	25	25	25	25	25	25	N
H.S.	5115.00	6933.00	8.800	740.000	8.800	41.000	10.000	H.S.
L.S.	0.0	3326.00	7.000	339.000	6.100	14.000	6.000	L.S.
RANGE	5115.00	3607.00	1.800	401.000	2.700	27.000	4.000	RANGE
SSC	119.55	86.56	0.043	8.355	0.066	0.626	.087	SSC

Table 6

UNIVERSITY OF NORTH DAKOTA
DEPT. OF HPER
Summer Research Project
AAHPER Youth Fitness Test Norms for Grade 5 and 6 Boys (N = 20)

JULY 1971

Standard Scale	Sit Ups No.	Pull-Ups No.	Figure 8 Sec. & Tenths	Standing Broad Jump Ft. & Ins.	50 Yard Dash Running Start Sec. & Tenths	Softball Throw Ft.	Six Minute Test Distance in Laps of the F H Track	Standard Scale
100	0	0	0.0	0.0	0.0	0.0	0.0	100
95	0	0	0.0	0.0	0.0	0.0	0.0	95
90	0	0	0.0	0.0	0.0	0.0	0.0	90
85	0	0	0.0	0.0	0.0	0.0	0.0	85
80	0	10	0.0	0.0	0.0	142	0.0	80
75	0	9	0.0	0.0	0.0	(133)	0.0	75
70	0	(8)	6.6	6'12"	0.0	125	11.0	70
65	0	6	(6.8)	6'8"	6.2	117	10.6	65
60	(103)	5	7.0	(6'4")	6.4	108	10.2	60
55	94	4	7.2	6'0"	6.7	100	9.8	55
50	84	3	7.4	5'7"	6.9	91	9.4	50
45	74	2	7.6	5'3"	(7.1)	83	9.0	45
40	65	0	7.8	4'11"	7.3	74	8.6	40
35	55	0	8.0	4'7"	7.6	66	(8.2)	35
30	45	0	8.3	4'3"	7.8	58	7.8	30
25	36	0	8.5	3'11"	8.0	49	7.4	25
20	0	0	8.7	3'7"	8.3	0.0	7.0	20
15	0	0	8.9	3'3"	8.5	0.0	0.0	15
10	0	0	0.0	2'11"	0.0	0.0	0.0	10
5	0	0	0.0	2'7"	0.0	0.0	0.0	5
0	0	0	0.0	0.0	0.0	0.0	0.0	0
MEAN	84	3	7.4	5.07	6.9	91	9.4	MEAN
SD	24	2.994	0.531	0.101	0.578	21.011	1.028	SD
N	20	20	20	20	20	20	20	N
H.S.	100.000	10.000	8.800	6.090	8.300	138.000	10.750	H.S.
L.S.	45.000	0.0	6.600	2.090	6.200	51.000	7.250	L.S.
RANGE	55.000	10.000	2.200	4.000	2.100	87.000	3.500	RANGE
SSC	1.928	0.238	0.042	0.008	0.046	1.082	.082	SSC
NO. TRLS	1	1	2	3	1	3	1	NO. TRLS

Table 7

UNIVERSITY OF NORTH DAKOTA
DEPT. OF HPER
Summer Research Project
[illegible] of the AAHPER Youth Fitness Test Norms for Grade 5 and 6 Boys (N = 20)

JULY 1971

Standard Scale	Pull Up Work Score	Sit Up Work Score	Shuttle Run Time - Sec & Tenths	Standing Broad Jump Work Score	50 Yard Dash Time - Sec & Tenths	Softball Throw Work Score	Six Minute Test Distance in Laps of the F H Track	Standard Scale
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100
95	0.0	0.0	0.0	0.0	0.0	0.0	0.0	95
90	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90
85	0.0	0.0	0.0	0.0	0.0	0.0	0.0	85
80	0.0	0.0	0.0	0.0	0.0	65	0.0	80
75	1237	12298	0.0	641	0.0	61	0.0	75
70	1070	11384	6.6	608	6.2	57	11.0	70
65	902	10501	6.8	575	6.4	53	10.6	65
60	735	9617	7.0	542	6.6	50	10.2	60
55	569	8734	7.2	509	7.1	46	9.8	55
50	401	7850	7.4	476	7.4	42	9.4	50
45	233	6967	7.6	444	7.6	38	9.0	45
40	66	6084	7.8	411	7.8	34	8.6	40
35	0.0	5200	8.0	378	8.0	30	8.2	35
30	0.0	4317	8.3	345	8.3	27	7.8	30
25	0.0	3433	8.5	312	8.5	23	7.4	25
20	0.0	0.0	8.7	279	0.0	0.0	7.0	20
15	0.0	0.0	8.9	0.0	0.0	0.0	0.0	15
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
MEAN	4006	7850	7.4	476	6.9	42	9.4	MEAN
SD	418	22	0.531	82	0.578	9.6	1.028	SD
N	20	20	20	20	20	20	20	N
H.S.	1200	11440	8.800	608.000	8.300	63.000	10.750	H.S.
L.S.	0	4130	6.600	297.000	6.200	23.000	7.250	L.S.
RANGE	1200	7310	2.200	311.000	2.100	40.000	3.500	RANGE
SSC	33.457	176	0.042	6.573	0.046	0.769	.082	SSC

their performances. The profiles illustrated in Tables 4 and 5 represent those scores achieved by one subject, a girl, in the study sample. The mean standard score for this subject over the seven items was 47.6 based on the raw score norms and 53.4 based on the work score norms. This discrepency is amplified if we consider only those four items which were actually calculated in work units (pull-ups, sit-ups, standing broad jump and the softball throw). The mean standard score form the raw score norms is now 48.7 as compared with a mean standard score of 58.7 from the work score norms.

Tables 6 and 7 illustrate the performance of another subject, a boy, who participated in the study. The same general trend is illustrated by comparing the raw scores and work scores on the basis of mean standard scores. The mean standard score over seven items for the raw scores norms was 56.1 and 63.4 for the work score norms. Once again this discrepency is apparent over the four items which were calculated on the basis of work. In this case the mean standard score was 66.2 from the raw score norms and 72.5 from the work score norms.

The overall fitness performance for each of these subjects leads to a different interpretation depending on which norms are used to construct the profile. Thus if one is to evaluate fitness performance, the relative status of their performances will be affected on the basis of the norms which are used.

The work score-raw score comparisons indicate by reason of the low correlations a need for additional study of these methods employed for testing fitness and also for prescribing exercise. Since physical educators prescribe exercise as a routine procedure in their classes, it may be somewhat startling to realize that the amount of work involved in performing

an exercise may be so variable.

It is interesting to speculate on the attitudes which we develop in our students as a result of experiences in physical education directly related to testing fitness and prescribing exercise. For example, if one instructs a class of thirty junior high school girls to perform twenty sit-ups, that prescription of exercise may elicit a very light to an exhaustive load on the individuals within the class. Yet all must do twenty sit-ups or suffer the experience of failure when in fact they have exercised to their maximum capacity.

The student is obviously aware of this failure yet at the same time is keenly aware of the fact that she has worked harder to perform this exercise than has her peers. Similarly the high school football tackle who may weigh 200 pounds is pushed to maximum to complete five pull-ups. His classmate who weighs 130 pounds is able to perform eight pull-ups. The measurement of fitness by the AAHPER test results in an evaluation of fitness which indicates that the football player is less fit than the 130 pound student, yet the football player has demonstrated his strength and power and is regarded by his peers as stronger than the smaller boy.

Such contradictions between fitness testing, exercise prescription and peer ranking must lead to skepticism on the part of students such as these. A re-examination of the factors involved in the performance of an exercise and corresponding sensitivity on the part of physical educators to the measurement and prescription of exercise based on muscular work seems essential to the accurate practice of physical education.

Conclusions

Within the limitations of this study the following conclusions appear justified.

1. The reliability of the anthropometric tests were acceptable, except for the pull-up items.

2. The reliability for the four items of the AAHPER test was acceptable for this study.

3. Comparisons between the AAHPER Youth Fitness Test and the UIID Revision of the AAHPER Youth Fitness Test indicated a substantial difference in the scores achieved for both boys and girls, except in the pull-up or flexed arm hang and the softball throw for distance.

APPENDICES

APPENDIX I

SUMMER RESEARCH PROJECT 1971
TELEPHONE INTERVIEW GUIDE

1. Introduce Yourself
   Name - a student at UND
   majoring in physical education

2. Purpose for the call: I'm calling for Professor Bolonchuk who is conducting a summer research project. The project is a study involving 5th and 6th grade boys and girls. We would like to include your son/daughter (name) in the project and we need your permission.

3. Test: AAHPER Youth Fitness Test. The test was developed by The American Association for Health, Physical Education and Recreation.

4. The purpose of this study is to determine the accuracy of the test to measure physical fitness.

5. We will send a detailed description of the project by mail.

6. The tests will be administered on June 23 and June 25 from 8:30 A.M. to 12:00 noon.

APPENDIX II

Summer Research Project 1971
Dept. HPER
The University of North Dakota

An appointment for the AAHPER Youth Fitness Test has been made for at A.M. on June 23 and June 25. Please be prompt. Each child should wear tennis shoes, shorts, or swimsuit and a T-shirt.

Yours sincerely,

Bill Bolonchuk

Project Director

APPENDIX III

Summer Research Project 1971
Department of Health, Physical Education & Recreation
University of North Dakota

Dear Parents and Participants in the Summer Research Project:

Thank you for participating in the summer research project which was briefly described to you in a recent telephone conversation. Your co-operation is appreciated and there is little doubt that the successful completion of this project began with your affirmative response.

The purpose of this project is to critically examine the American Association for Health, Physical Education and Recreation Youth Fitness Test. The test was published in 1957 and has been used extensively in physical education programs to measure the physical fitness of school children ranging in age from 7 to 17 years. Although a great deal of research has been done on this test since 1959, little evidence exists to show the influence of body size on the scores which are achieved by the students taking the test. For example: Does a heavy boy or girl need to jump as far as a boy or girl whose body weight is less? What is the difference between the raw score for a test and the work accomplished during that test? These questions are the basis for this summer research project.

To more completely explain the procedure for the study a demonstration is planned for the early evening of June 17. This meeting is designed to acquaint you and your child with the purpose of the study, the procedure that we will use for testing the children and a description of the items involved in the tests. The meeting will be held in the Fieldhouse Arena beginning at 7:30 P.M. (Please use the East entrance) and will last for about an hour.

The program for the meeting will include:

- A brief explanation of the purpose of the study
- A film entitled: "The Fitness Challenge", this film describes each of the items of the AAHPER Youth Fitness Test
- Discussion of the procedure and related matters

If you are unable to attend, your absence will not disqualify your child's participation in the study and we will expect him or her on the scheduled test days. The test schedule is not complete yet, however, when the schedule is finalized we will mail an appointment form to you. All of the testing will be done on Wednesday, June 23 and Friday, June 25, from 8:30 A.M. to 12:00 noon.

Yours sincerely,

Bill Bolonchuk

Bill Bolonchuk
Associate Professor
Dept. HPER
UND

APPENDIX IV

Summer Pesearch Project 1971
Score Card

Name: _____ Sex: M F (circle)
Last First

Test Date: _____

Age: _____ Years Birthdate: _____(D) _____(M) _____(Y)

Body Weight: _____lbs. Standing Height: _____ins.

Sitting Height: _____ins. Upper Body Weight: _____lbs.

Pull-up Measures: C of I to floor extended arms _____ins.
C of I to floor flexed arms _____ins.
Difference _____ins.

Weight of the softball: _____ozs.

Station Number	Test	Trials	Score Units	Trial Score	Best Score
1.	Pull-Up or Flexed Arm Hang	1	Total No. of Continuous Reps or Time		
2.	Sit-Ups	1	Total No. of Continuous Reps		
3.	Shuttle Run or Agility Run	2	Time in Sec. & Tenths		
4.	Standing Broad Jump	3	Distance in feet and inches to the nearest inch		
5.	50 Yard Dash	1	Time in Sec. & Tenths		
6.	Softball Throw for Distance	3	Distance in feet to the nearest foot		
7.	600 Yard Run-Walk or 6 Minute Run-Walk	1	Minutes and seconds to the nearest second		

NAME _____

Reference

1. AAHPER Youth Fitness Test Manual NEA Publications-Sales, Washington, D.C., 1965 edition.
2. Askew, N.R., Reliability of the 600 Yard Run-Walk Test at the Secondary School Level, Research Quarterly, 37 : 451-54, 1966.
3. Balke, B., A Simple Field Test for the Assessment of Physical Fitness, CARI Report 63-18, Oklahoma City : Civil Aeromedical Research Institute, Federal Aviation Agency, September, 1963.
4. Barrow, H.M., and McGee, R., A Practical Approach to Measurement in Physical Education, Lea and Febiger, 1964.
5. Baumgartner, T.A., Stability of Physical Performance Test Scores, Research Quarterly, 40 : 257-61, 1969.
6. Bolonchuk, W.W., The Accuracy of the Six Minute Run Walk Test to Measure Cardiorespiratory Fitness, unpublished report to the UND committee on research, June, 1970.
7. Clarke, H.H., Application of Measurement to Health and Physical Education, 4th edition, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1967.
8. Cooper, Kenneth H., A Means of Assessing Maximal Oxygen Intake, Correlation between Field and Treadmill Testing, J.A.M.A. Vol. 203:201, January 15, 1968.
9. Delorme, T.L. and Watkins, A.L., Progressive Resistance Exercise: Technique and Application, New York : Apple-Century-Crofts, Inc., 1951.
10. Doolittle, T.L. and Bigbee, R., The Twelve-Minute Run-Walk : A Test of Cardiorespiratory Fitness of Adolescent Boys, Research Quarterly, 39 : 491-95, 1968.

11. Fleishman, E.A., Examiner's Manual for the Basic Fitness Tests, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1964.

12. Gates, D.P., And Sheffield, R.P., "Tests of Change of Direction as Measurement of Different Kinds of Motor Ability in Boys of the 7th, 8th and 9th Grades," Research Quarterly, 11 : 136-37, October, 1940.

13. Gulsvig, J., An Investigation of the Reliability and Validity of the Grand Forks Fitness Test as administered to forty-two Central High School Boys, unpublished Masters Thesis, University of North Dakota, 1969.

14. Johnson, B.L., and Nelson, J.K., Practical Measurements for Evaluation in Physical Education, Burges Publishing Company, Minneapolis, 1969.

15. Klesius, S.E., Reliability of the AAHPER Youth Fitness Test Items and Relative Efficiency of the Performance Measure, Research Quarterly, 39 : 809-11, 1968.

16. Marmis, C., Montoye, H.L., Cunningham, D.A. and Kozar A.J., Reliability of the Multi-trial Items of the AAHPER Youth Fitness Test, Research Quarterly, 40 : 240-45, 1969.

17. Mathews, D.K., Measurement In Physical Education, W.B. Saunders Co., 3rd edition, Philadelphia, 1968.

18. McCloy, C.H., and Young, N.D., Tests and Measurements in Health and Physical Education, New York : Appleton-Century-Crofts, Inc., 1954.

19. Montoye, H.L., An Introduction to Measurement in Physical Education, Phi Epsilon Kappa Fraternity, Indianapolis, Ind., Vol. 4, 1970.

20. Reiff, G., Kish, L. and Harter, J., Selecting a Probability Sample of School Children in the Coterminous United States, Research Quarterly, 39 : 409-14, May, 1968.

21. Ricci, B., AAHPER's Open Forum for a Moratorium on "Physical Fitness" Testing, Journal of HPER, pp. 28-30, March, 1970.

22. Rathermel, B.L., Pollock, M.L., and Cureton, T.K., AAHPER Physical Fitness Test Score Changes Resulting from an Eight-Week Sports and Physical Fitness Program, Research Quarterly, 39 : 1127-29, 1968.

23. Soderberg, P., The Reliability and Validity of the Modified AAHPER Fitness Test as applied to Freshman College males at the University of North Dakota, Master of Science Thesis, UND, 1970.