A COMPARISON OF SELECTED STEP TESTS FOR THE DETERMINATION OF CARDIOVASCULAR EFFICIENCY IN COLLEGE WOMEN

By

Bette Ruth Goldstein

An Abstract of

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Health, Physical
Education and Recreation

1968

ABSTRACT

A COMPARISON OF SELECTED STEP TESTS FOR THE DETERMINATION OF CARDIOVASCULAR EFFICIENCY IN COLLEGE WOMEN

By

Bette Ruth Goldstein

This study involved 103 college women to determine whether an eight-inch step test with a rate of 30 steps per minute for a period of one minute could produce results as good as or better than one of the best tests currently in use for women, the Skubic-Hodgkins Step Test.

Selected anthropometric measurements were taken prior to the performance of both step tests to determine possible effects of weight, height, and leg length variactions upon cardiovascular performance in step tests. Pulse rates were recorded electrically with the use of surface electrodes during the recovery periods of each step test, and were counted in 15-second intervals.

Pearson Product-Moment Correlations were computed for all possible combinations of pulse rates on both step tests and the anthropometric measurements. Analysis of variance was also performed on 30 selected subjects in good, average, and poor cardiovascular fitness to determine the discriminatory powers of the two tests.

On the basis of the results of the analyses, the following conclusions could be drawn:

1. There were significant correlations between the the results of both tests at the .05 level. However, the correlations did not account for a large proportion of the variability in the data.
2. Taller women had significantly better performances on the more severe Skubic-Hodgkins test. The eight-inch test created a relative disadvantage for the taller, longer-legged women in terms of heart rate response.
3. Heavier women consistently performed significantly poorer on the Skubic-Hodgkins Test, whereas there was no significant correlation between weight and any eight-inch pulse interval. This supported the conclusion that the eight-inch test is less sensitive to weight, height, and leg length variations.
4. The Skubic-Hodgkins Test had better definition between groups of fitness levels than the eight-inch test until new variables of raw pulse rates were generated. The new variables involved the pulse counts in the 0:45-1:15 and 1:00-1:15 intervals and were:

A. $\frac{\text{Pulse rate in that interval X Weight}}{\text{Height}}$

B. Pulse rate in that interval X Weight

X Height

These variables created the same discriminatory power for the eight-inch test as the Skubic-Hodgkins Test had achieved.

5. A cardiovascular fitness index for use with the eight-inch test was determined mathematically to be:

Cardiovascular Index =

$$110 - (1.43)\ \frac{\text{0:45-1:15 pulse count X Weight}}{\text{Height}}$$

6. The eight-inch test is well within the capabilities of most women in a wide age range, and would therefore be suitable for use in longitudinal studies.

A COMPARISON OF SELECTED STEP TESTS FOR THE DETERMINATION OF CARDIOVASCULAR EFFICIENCY IN COLLEGE WOMEN

By

Bette Ruth Goldstein

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Health, Physical
Education and Recreation

1968

DEDICATION

To the memory of my Dad

ACKNOWLEDGMENTS

The author wishes to express her sincere appreciation to those persons who have greatly contributed to make this study possible.

To Dr. William W. Heusner, her advisor, for assistance with the concepts and statistics involved, and for his encouragement, without which the task would have been impossible.

To Mr. Robert Wells, whose technical skill and ability made the data collection possible.

To Mr. Walter Smith, whose assistance with the subjects contributed greatly.

To Mr. Brian M. Petrie, whose assistance with statistical calculations, and interest were invaluable.

To the students of Michigan State University who gave so willingly of their time and effort.

TABLE OF CONTENTS

Page

DEDICATION ii

ACKNOWLEDGMENTS iii

LIST OF TABLES vi

Chapter

I. INTRODUCTION 1

Purpose of the Study. 2
Need for the Study 4
Limitations of the Study 6
Definitions of Terms. 7

II. REVIEW OF LITERATURE 8

Brief History of Cardiovascular Testing 8
Pulse-Ratio Test 10
Harvard Step Test. 13
Cardiovascular Tests for Women . . . 14
Brizendine Eight-Inch Step Test . . . 16

III. METHODS OF PROCEDURE 18

Selection of Subjects 18
Preliminary Instructions 18
Step Tests 19
Anthropometric Measurements 19
Preparation of Subjects for Testing. . 20
Testing Routine 21
Recording Procedure 22
Analysis of Data 22

Chapter		Page
IV.	RESULTS	24
	Presentation of Data	24
	Correlation Analysis	26
	Analyses of Variance	30
	Cardiovascular Fitness Index	38
V.	SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	41
	Summary	41
	Conclusions	42
	Recommendations	44
BIBLIOGRAPHY		46

LIST OF TABLES

Table		Page
1.	Means and Standard Deviations of Pulse Counts During All Recovery Intervals .	25
2.	Height and Leg Length Correlations With Pulse Counts Following the Skubic-Hodgkins and Eight-Inch Tests . . .	27
3.	Correlations of Weight With Pulse Counts Following the Skubic-Hodgkins and Eight-Inch Tests.	28
4.	Correlations Between Pulse Counts During the Skubic-Hodgkins 1:00-1:30 Recovery Interval and All Recovery Intervals Following the Eight-Inch Test . . .	29
5.	Height, Weight, and Leg Length Characteristics of the Total Sample, the Selected Samples and Subgroups of Good, Average and Poor Cardiovascular Condition . .	31
6.	Analysis of Variance: Skubic-Hodgkins 1:00-1:30, and All Eight-Inch Recovery Intervals	32
7.	Tukey Test of Comparison of Means Two at a Time Following an F-Test: Skubic-Hodgkins 1:00-1:30 and Eight-Inch 1:00-1:15 and 0:45-1:15 Recovery Intervals .	34
8.	Analysis of Variance: Eight-Inch Test Recovery Intervals, 1:00-1:15 and 0:45-1:15, and Variables Transformed by Use of Anthropometric Data.	36

Table		Page
9.	Tukey Test of Comparison of Means Two at a Time Following an F-Test: Eight-Inch 0:45-1:15 and 1:00-1:15 Recovery Intervals.	38
10.	Measures of Central Tendency and Variation of the Cardiovascular Fitness Index . .	40

CHAPTER I

INTRODUCTION

Beginning about the turn of the century, there developed an increasing interest in cardiovascular fitness, and consequently, an interest in increasing the reliability of its measurement. It was reasoned that since the heart and circulatory system respond to a stress in proportion to the degree of the stress, then the heart's response to a standard, known work load would determine its ability to meet and recover from the demands placed upon it. This response would also serve as a measure of the dynamic fitness of the individual.

In 1943 Lucien Brouha, with the Harvard Fatigue Laboratory, developed the now standard Harvard Step Test. This test, designed for men, utilizes an exercise program of box stepping "so severe that no one can perform in a 'steady state' for more than a few minutes" (2:31). It classifies young men into three groups of least fit, fit, and most fit. The individual's measured heart rate after exercise is used to "evaluate the stress imposed by muscular activity upon the heart and the circulation with a minimum amount of interference with the subjects' freedom of motion and performance ability", due to the close relationship between heart rate, cardiac output, and oxygen consumption (3:197).

From this start, a number of other modified tests, including a very few for women, were devised.

Purpose of the Study

There have been various modifications of the Harvard Step Test for girls and women (5, 13, 19, 27, 28, 29, 30). However, even with the best of these tests, there is still some difficulty. Part of this occurs with the administration of the test, but even more occurs with the subjects' performance of the test.

If one considers the best of the modifications yet developed, a 3-minute test using 24 steps per minute onto an 18-inch bench devised by Skubic and Hodgkins, there are still the problems of:

1. Time consumed in the testing procedure. Three minutes of exercise plus one and one-half minutes of recovery for a total of four and one-half minutes are required for completion of the test;
2. Height of the stepping box. It has been observed that shorter and heavier subjects have a great deal of difficulty executing the 18-inch stepping height; and,
3. Combination of height, rate of stepping, and duration of the test. This combination of factors tends to stress muscular endurance of the legs as well as cardiovascular fitness.

These last two factors tend to limit performance and bias the results in favor of the taller and lighter subjects.

Another factor to be considered, and one which should be mentioned in support of this test, is that the authors claim that it will distinguish between the 3 cardiovascular fitness classifications of highly trained, active, and sedentary at the .01 level of significance. However, when this test was administered to 30 students categorized into 3 groups of good, average, and poor cardiovascular fitness, it distinguished between good and poor, and average and poor at the .05 level of significance, but did not distinguish between good and average at this level.

In 1967, S. F. Brizendine, reporting on "A Comparative Study of the Harvard Step Test with Tests of Varied Times and Heights" (1), found a high positive rank-order correlation of .89 between the results of the Harvard Step Test and the results of an 8-inch step test at 30 steps per minute for 1-minute duration. As with the Harvard Step Test, Brizendine found that his test did not discriminate significantly between good and average cardiovascular fitness levels. It did discriminate between poor and good, and poor and average degrees of fitness at a higher level of significance than did the Harvard Step Test.

It was decided that this problem should be studied with a view towards the determination of a cardiovascular efficiency test that would be more feasible for use with typical college-age women. Such a test should create neither the extreme fatigue during testing, nor the discomfort present for a period of several days after testing, as is prevalent with those tests currently in use. It was decided that the effects of height, weight, and leg length upon performance should be studied as part of the investigation.

Need for the Study

There is a recognized need for the measurement of cardiovascular efficiency of both men and women. Unfortunately, very little of the work which has been done in this field has been conducted on women. As a result, the tests used for women and girls have been modified versions of the Harvard Step Test and are extremely difficult tests for females.

The Skubic-Hodgkins combination of an 18-inch step with a rate of 24 steps per minute for 3 minutes is so severe that 13.3 per cent of a sample of 2,360 college women throughout the country could not complete the testing routine (13). Observation of those students who could complete the routine showed that the stepping rate is difficult to maintain and that when the test is administered to large groups simultaneously, the pace of the entire group can be disrupted by a few who lose cadence. Due to the demands of this and other tests, those who experience difficulty modify their own performance by their inability to establish an erect position on the stepping box, irregular stepping, and missed steps. This negates the standardization of the work of the test and may be the result of insufficient leg strength rather than lack of cardiovascular efficiency as is presumed (14, 29). In some instances, there is the added difficulty of counting extremely rapid pulse rates caused by the severe test conditions.

It is believed that the amount of time required to complete the test, and the severity of performance with its resulting inaccuracies are responsible for practitioners in the field of physical education not making as much or as advantageous use of cardiovascular efficiency testing as is possible (13, 27). In order to realize full utilization then, a satisfactory test should be relatively short and not too strenuous and should involve a common activity in daily use. This type of activity would eliminate the possibility of improved performance through practice and training (9, 10). A test such as the 8-inch test, used by Brizendine, at a rate of 30 steps per minute for one minute with one and one-half minutes of recovery would meet these requirements (1). It would also eliminate the need for specially prepared benches, as 8 inches is the height of a common stair step.

In addition, if a test can be found which will give the same information as the more strenuous Harvard Step Test, such a test could be used at all age levels. This would prove invaluable in longitudinal studies on both males and females, an area in which there is a dearth of information.

Limitations of the Study

One of the prime difficulties of studies involving recordings of heart rate is control of the subjects. Activity, smoking, eating, and sleeping habits will all affect an individual's heart rate. Testing during the spring season with heightened activity out-of-doors must also be considered as an affecting agent. In an effort to control these factors, the subjects for this investigation were tested and then retested, during the same hour of the day, within a period of from 3 to 7 days. They were also instructed to continue their normal daily patterns of living with no alterations other than cessation of smoking, and avoidance of large food consumption 1 hour prior to testing. Subjects were cautioned not to participate in strenuous activity for a 24-hour period prior to testing. Unfortunately, these directives leave quite an area for possible error.

An additional difficulty occurred with the placement of the electrodes on the female subjects. Due to positioning of the electrodes on either side of the sternum, and to the limited number of electrodes used, respiratory and muscle action patterns were also recorded for some subjects. This tended to "clutter" a few of the recordings and created a difficulty in reading those particular pulse counts.

Another limitation, brought to light as a result of the study, was the failure to take measurements of the subjects' leg strength. Observation and evaluation have shown that it is not the length of the subject's legs which permit her to complete the Skubic-Hodgkins Step Test. Quite possibly, leg strength, as a component of muscular endurance, is the determining factor.

Definitions of Terms

In the interest of a common basis of understanding, it is best to establish a standard terminology.

The test to be used as a criterion measure is that devised by Vera Skubic and Jean Hodgkins. It is a 3-minute step test, using a stepping height of 18 inches and a rate of 24 steps per minute. The subject performs the prescribed exercise and then rests in a sitting position for 1-minute. At the end of this period, a 30-second pulse count is taken. This test will be referred to hereafter as the Skubic-Hodgkins Test.

The experimental test is one developed for use with college-age men by Stanley F. Brizendine. It utilizes an 8-inch stepping height in combination with a rate of 30 steps per minute for a duration of 1-minute. Pulse rates are recorded for the last 30 seconds of a one and one-half minute recovery period.

Leg length of the subjects is defined as the distance from "the superior surface of the greater trochanter of the femur to the floor" (21:563).

CHAPTER II

REVIEW OF LITERATURE

The use of standardized exercise as a means of eliciting a cardiovascular response which could evaluate an individual's level of fitness is not a recent development. Such tests have a history of almost a century of research and application. Experimenters have persevered in their efforts to establish such a cardiovascular test due to its relative ease of administration to large groups. However, the research which has come down through the years pinpoints a great many inaccuracies and conflicting studies.

Brief History of Cardiovascular Testing

Originally, experimentation with cardiovascular measurement arose as a result of the interest vested by physiologists in muscle strength. The ergograph, invented by Mosso in 1884, "gave rise to the development of cardiovascular tests" (20:223). This development was further strengthened by an interest in cardiac function.

W. P. Bowen was the first to publish results of a cardiovascular function study on the relationship of pulse rate to exercise and physical fitness. His work was followed by Crampton, reporting on heart response to activity in 1905, McCurdy in 1910, and Barringer reporting in 1916 on the relationship between blood pressure and physical deficiency (20).

Then, in the 1920's, researchers at Guy's Hospital in London, England, began experimentation on a test which would evaluate the physical fitness level of men (4, 9, 19). Originally, they worked on a staircase, with subjects running up and down the staircase a given number of times within specified time limits. It was found that the exercise was uniform for each individual, but was not standardized between individuals. Each subject had his own particular style in ascending and descending, which affected the total amount of work performed by each individual. The researchers then attempted to determine a standard exercise to provide a specific task for all subjects. They decided to use stepping on a 13-inch bench at a rate of 24 and 30 steps per minute for 3 minutes. This task was equal in total height stepped to the staircase exercises (9). They used a stepping exercise because practice causes improvement; and rather than allow unfair advantage to any subject, they wanted an exercise which was a daily activity for all. It was their opinion that "the rate of the pulse after muscular work to which the subject is accustomed is a good test of efficiency, for the reaction of the pulse to muscular exercise involves adequate adaptations of the nervous, respiratory, circulatory, and muscular systems" (9:373).

The Guy's Hospital researchers then proceeded with a pulse-ratio type of test. After an exercise of stepping on a 13-inch bench 28 times a minute for 3 minutes, the subjects were allowed 5 seconds to sit quietly, and their pulse was then counted for 2 minutes. The ratio was determined as the number of beats in that 2 minute period divided by the resting pulse rate, or:

$$\text{Pulse ratio} = \frac{\text{Heart beats in 2 minutes}}{\text{Resting pulse 1 minute}}$$

After working with the pulse-ratio, the researchers decided that it was not as useful as the rate of decline of the pulse after exercise, which "gives a more accurate and more detailed knowledge of the behavior of the pulse and a better indication of a man's fitness than does the pulse ratio" (4:272). However, after all their experimentation, their consensus was that no single measurement they had used was infallible.

Pulse-Ratio Test

Experimentation with a pulse-ratio test continued on into the 1930's. W. W. Tuttle (35) felt that the response of the heart rate to exercise was an index of cardiovascular efficiency. He believed that the "cardiovascular response reflects the physiological condition of the individual" (31:5).

The pulse-ratio "represents the ratio of the resting pulse rate to the rate after exercise" following a known amount of exercise (31:5). It is determined by counting the pulse for 2 minutes post-exercise and then dividing this by the resting pulse for 1 minute. The object of the test was to achieve an empirical pulse ratio of 2.5, an arbitrary goal which was within the capabilities of the majority of subjects originally tested. This technique necessitated 3 separate test periods to determine the pulse rate ratio. The first test was calculated to produce a pulse ratio below 2.5, the second a pulse ratio above 2.5, and the third the theoretical 2.5 pulse ratio. This was accomplished by using a 13-inch stool on all tests and varying the amount of work by adjusting the number and rate of steps within a one, two, or three-minute exercise period. This fitness index is therefore a function of the amount of work required to produce a 2.5 pulse ratio.

Two of the limitations inherent in this test are obvious. First, conditions for the first and second tests are estimations of a person's performance. Therefore, if the experimenter over-estimates the work for the first trial, the subject's fatigue level would be high. This would adversely affect his performance in the second and third trials. In comparison, if a person's work load was greatly under-estimated, then it is possible that he would be in a far less fatigued state by the third trial. Thus, the performances for 2 subjects, even though both attained pulse ratios of 2.5, could not really be regarded as equivalent.

Second, to maintain a standard work load for each test, regardless of intensity, each subject must assume an erect position. In addition, Tuttle (34) stated that application of this test must be limited to those individuals whose pulse rate was within the normal range.

Other experimenters dealing with the application of the Pulse-Ratio Test have reported conflicting results. Flanagan states the pulse ratio is a reliable evaluative technique for endurance (6), while others have found that not only is it unsound as a criterion for endurance, but it is also not as reliable as it should be for use with individual subjects (7, 11). The consensus is that the pulse-ratio test may be applicable for discrimination between groups of subjects, but not within groups.

Tuttle also attempted to apply the Pulse-Ratio Test for use with women. Working with Henryetta Frey in 1930, he conducted a study on college women to determine physical efficiency by means of the Pulse-Ratio Test. In their conclusions, they stated that "the use of this test indicated that the pulse ratio after mild exercise is a good index to physical efficiency" (33:20). However, further investigation by other researchers disclosed a measure of reliability too low to justify the expenditure of time required for testing.

It was results such as these that kept researchers working to find a more reliable and efficient cardiovascular test. Among these researchers was Lucien Brouha and his colleagues at the Harvard Fatigue Laboratory.

Harvard Step Test

Brouha felt that in order for a given task to be a valid criterion of cardiovascular fitness, it must provide an exercise so severe that no one could perform in a 'steady state' for more than a few minutes. The 2 factors involved in this test then, are the length of time one could continue the task, and the deceleration of the pulse during recovery. It was stated that the "easier the work, the less clear cut are the differences between the fit and the unfit" (2:31). Therefore, the task chosen for the test must be one of great severity. For this purpose, Brouha chose a step test with a duration of 5 minutes at a height of 20 inches and a rate of 30 steps per minute. This was the test which has come to be known as the Harvard Step Test.

The subjects perform this test and then rest as their pulse is counted at 30-second intervals for 1 to 1½, 2 to 2½, and 3 to 3½ minutes of recovery. The fitness index then is computed by multiplying the length of the exercise for each individual by 100 and dividing by twice the sum of the pulse counts:

$$\text{Index} = \frac{\text{Duration (sec.) X 100}}{\text{2 X Sum of pulse counts}}$$

An individual's index score places him in one of 5 fitness categories: Poor, Low Average, High Average, Good, and Excellent. However, when the Harvard test is applied to groups of good, average, and poor cardiovascular fitness, it differentiates significantly between good and poor, and average and poor, but does not differentiate significantly between good and average levels (1).

The major feature of the Harvard test is its severity. Unfortunately, it is this factor which causes the major difficulties. The test is just too severe for some men and a larger number of women. There are inaccuracies with counting and computing the 3 separate pulse counts, and the test is quite time-consuming and fatiguing. It is for the above reasons that this test had to be modified before it could be applied with any success for use with women and girls.

Cardiovascular Tests for Women

As nearly all the cardiovascular fitness research to this time had been done by and for men, the tests produced had to be modified or adapted for use with women.

In 1943, Harriet Clarke, among others, attempted to find a cardiovascular fitness test for women that would be as good as the Harvard Step Test. The test she chose was, again, a step test so as to reduce the element of skill and practice required. The exercise consisted of "stepping up and down on an 18-inch bench every 2 seconds" (5:358), for as long a time as possible up to a maximum of 4 minutes. Clarke retained the Harvard pulse count and fitness categories, but modified both height and duration of stepping. This test, too, is extremely severe and time-consuming, so that its practical application is limited. In addition, further research by Scott, Mordy and Wilson on this Clarke-Brouha Functional Physical Fitness Test indicated that it "has little or no relationship to the individual's capacity to do work" (26:128).

Further modifications of the Harvard Step Test were devised by A. W. Sloan. He compared the fitness indexes of two groups of young women performing the Harvard Step Test at varying heights of 16, 17, 18, and 20-inches, with up to a maximum exercise period of 5 minutes, to corresponding groups of men performing the standard Harvard test. He found that a t-test showed no significant difference between the fitness indexes of men and women. The closest agreement on fitness indexes occurred between the Harvard test and the modified 17-inch test. "Not only did the mean fitness indexes correspond, but also the distribution of individuals in the several categories of physical fitness" (29:985).

Sloan admits that the 20-inch height of the Harvard test "has a particular disadvantage for women in that the limiting factor is often local fatigue of leg muscles; in these cases, the test is not a true measure of the capacity of the body as a whole for sustained muscular effort" (29:986). He further states that this limiting factor can be prevented by lowering the stepping height to 18-inches, "since large muscle groups are used, the limiting factor is usually cardiorespiratory embarrassment rather than local muscular fatigue" (29:986).

Following current trends, Sloan also experimented with the possibility of replacing cardiovascular tests with the resting pulse rate as a fitness index, but found it "a much less reliable measure of physical fitness than is the Harvard Step Test" (30:169).

The most recent of the modifications for girls and women is a test devised by Skubic and Hodgkins, which was used as the criterion test for this study. They desired a test which was short, efficient, not too difficult for women to perform quick to evaluate, and yet would still be reliable and valid. To meet these needs, they developed an 18-inch step test with a rate of 24 steps per minute for 3 minutes. After the test, the testees rest for 1 minute and then the pulse is counted for 30 seconds. Unfortunately, even with these modifications, 13.3 per cent of a sample of 2,360 college women, to whom the test was administered, could not complete it (13:455). Although this test retains severity of performance, and cumbersome administration, its value lies in its ability, according to its authors, to discriminate between trained and active, trained and sedentary, active and sedentary, and trained and untrained subjects at the .01 level of significance (27). Thus, the discriminatory power of the test is far greater than any other cardiovascular test devised for men or women to date.

Brizendine Eight-Inch Step Test

One further modification of the Harvard Step Test is worthy of note for this study. Experimentation and comparison

of the Harvard test with other step tests using varied times and heights produced a very high positive rank order correlation between the Harvard test and a 1-minute, 8-inch, 30-steps-per-minute test. It is also significant that in one sample of men, the Brizendine Eight-Inch Test had the same powers of discrimination as did the Harvard Step Test (1).

On the basis of this high correlation, it was decided that the eight-inch test should be used as the experimental test in this study.

CHAPTER III

METHODS OF PROCEDURE

Selection of Subjects

The subjects involved in this study were 103 young college women enrolled in HPR 105: Foundations of Physical Education, and HPR 109: Individual Skills Courses, at Michigan State University. Each subject volunteered for participation in the study. Prior to agreement to serve as a subject, the students knew only that they were to take part in a research program which their instructor was conducting.

Preliminary Instructions

Once a student had volunteered, she was informed as to the purpose and procedures of the study. This served not only as a source of information for the subjects, but also as a means of motivating them to perform well in each of the tests.

An attempt was made to control possible biasing of the results due to circadean variability and external factors such as training, activity, eating, sleeping, and smoking habits. Each subject was scheduled for 3 test periods at the same hour of the day, at least 3 days but not more than 1 week apart. They were instructed to follow normal sleeping, eating, and smoking patterns, with the exceptions of cessation of smoking and abstinence from large food consumption 1 hour prior to testing. The subjects also were instructed to refrain from any severely demanding physical activity for a 24-hour period prior to testing.

Step Tests

Two step tests were used in this investigation. The criterion test was the Skubic-Hodgkins Test which consists of stepping for three minutes on an 18-inch step at a rate of 24 steps per minute (13). The experimental test, developed by Brizendine (1), consists of stepping for one minute on an 8-inch step at a rate of 30 steps per minute.

Anthropometric Measurements

To determine the effects of body structure upon performance of the tests, it was necessary to take various anthropometric measurements. The weights of the subjects were measured on a springless scale calibrated in kilograms to the nearest hundredth of a kilogram. Height was measured against a fixed centimeter scale. The length of each subject's left leg, from the greater trochanter to the floor, also eas measured in centimeters by the use of an anthropometer. To determine test-retest reliability for leg length, 22 subjects were tested a second time at a three week interval. The reliability coefficient was found to be .846, which was considered sufficiently reliable to warrant the use of leg length as a tool of measurement. All measurements were taken with the subject barefoot and wearing only a light gymnasium uniform consisting of thin cotton shorts and a blouse.

Preparation of Subjects for Testing

Since this study required accurate determinations of heart rate, immediately following the exercise of stepping up and down on boxes eight and eighteen inches high, a Gilson Recorder was used to record both exercise and recovery pulse rates. Three Grass surface electrodes were placed on the subject prior to exercise. Of these three electrodes, one was placed on the small of the back, one inch to the left of the spinal column, as a ground. The remaining two electrodes were placed on the subject's chest. One was placed directly below the left nipple in the fold of the left breast, just underneath the line of the brassiere, and one inch toward the sternum. The third electrode was placed in the fold of the right breast, below the nipple and one inch toward the sternum.

The sites of application of the electrodes were first prepared with a coating of an adhesive substance. Then, by scraping with a sharp blade, an area was cleared for the placement of each electrode. Electrode jelly was used. The electrodes were secured to the body with crossed strips of adhesive tape, as were the wires leading from them to the recorder.

Testing Routine

The subjects reported for testing a total of three times and followed a set routine each day.

DAY 1: Upon arrival at the laboratory, the subject sat and rested for a period of five minutes. After this, her resting pulse was taken and recorded for further use as an indication of cardiovascular fitness. The electrodes then were applied to the subject. Height and weight measurements were taken next. Finally, the subject was led to the testing area, where the temperature was kept between 22 and 25 degrees Centigrade. The subject was allowed to rest for a few moments while the electrodes were attached to the recorder and the equipment for the specific test to be performed was arranged. The order of tests for each subject test had been predetermined by random assignment. Following this procedure, the subject again was instructed in the proper technique for executing the test she was to perform that day. The subject then performed the required step test, and immediately upon completion of the test, sat in a chair placed three feet behind the testing box.

DAY 2: The subjects followed the same procedure as on the first day of testing. However, on this day, each subject performed the second of her two tests.

DAY 3: On this day, the subjects did not perform either of the step tests. Leg length measurements were taken and recorded. The greater trochanter of the femur was located by palpation.

To enable the subjects to maintain the cadence of the step tests, an electric metronome was used. The metronome was calibrated, and periodically re-calibrated, for 96 and 120 beats per minute which yielded a rhythmic cadence of 24 and 30 steps per minute respectively. In addition, a twelve-inch diameter timer was placed in clear view of both the subject and the test administrator, so that each could see the expiration of time for each test. The subjects were verbally cued at the half-way point in the stepping exercise and at fifteen seconds prior to the end of the test.

Recording Procedure

Heart beats were recorded electrically during the last 30 seconds of activity and during the first one and one-half minutes of recovery. Fifteen-second intervals were marked on the recordings by means of a built-in timing device on the Gilson Recorder. Once a test was completed, the heart beats in each 15-second interval were counted manually and the statistical results were recorded on data sheets for analyses.

Analysis of Data

Pearson product-moment correlation coefficients were computed between all possible combinations of recovery pulse counts and anthropometric measurements. Thirty students, ten each in relatively good, average and poor cardiovascular fitness, were selected for various tests of significance of mean differences between subgroups. The minimum acceptable significance level was set at .05. On the basis of these results, the discriminatory power of the 8-inch test warranted generation of new variables to eliminate any possible effects of anthropometric variability. Finally, a preliminary index of cardiovascular fitness was formulated.

CHAPTER IV

RESULTS

Presentation of Data

The study involved separate testing sessions for the administration of the Skubic-Hodgkins Step Test and the Brizendine Eight-Inch Step Test. Each subject completed both step tests. Recovery heart beats were counted in 15-second intervals to facilitate comparisons between tests. One and one-half minutes of recovery was deemed sufficient since: (a) in the Skubic-Hodgkins test, which was used as a criterion, only the pulse count during the 30-second period between one and one-half minutes after exercise is recorded; and (b) the eight-inch step test is of a lesser degree of severity and thus should produce a significant drop in pulse rate earlier than the Skubic-Hodgkins test. It follows that the highest correlations between the results of the two tests would not be expected in corresponding time intervals.

The raw data collected on each subject consisted of anthropometric measurements of height, weight, and leg length, and recovery pulse counts during 15-second intervals from 0:00 through 1:30 for both the Skubic-Hodgkins and eight-inch tests. The 15-second pulse counts were then combined over larger time intervals for further analyses. These intervals and the pulse count means and standard deviations are shown in Table 1.

TABLE 1.--Means and standard deviations of pulse counts during all recovery intervals.

Recovery Intervals	Skubic-Hodgkins		Eight-Inch	
(Minutes & Seconds)	$\bar{X}$	S.D.	$\bar{X}$	S.D.
0:00-0:15	43.22	2.87	32.40	3.40
0:15-0:30	39.35	3.02	27.37	3.66
0:30-0:45	36.65	3.30	24.60	3.19
0:45-1:00	34.78	3.49	22.17	3.27
1:00-1:15	33.01	3.27	20.83	3.15
1:15-1:30	31.60	3.38	20.37	3.27
0:00-0:30	82.57	5.52	59.77	6.69
0:00-0:45	119.22	8.45	84.37	9.59
0:00-1:00	154.00	11.66	106.54	12.42
0:00-1:15	187.01	14.73	127.38	15.11
0:00-1:30	218.61	17.86	147.75	18.10
0:15-0:45	76.00	6.08	51.97	6.72
0:15-1:00	110.78	9.37	74.15	9.69
0:15-1:15	143.79	12.48	94.98	12.50
0:15-1:30	175.39	15.65	115.35	15.55
0:30-1:00	71.43	6.62	46.78	6.29
0:30-1:15	104.44	9.75	67.61	9.17
0:30-1:30	136.04	12.96	87.98	12.25
0:45-1:15	67.79	6.66	43.01	6.24
0:45-1:30	99.39	9.90	63.38	9.34
1:00-1:30	64.61	6.53	41.20	6.28

The anthropometric characteristics of the subjects and their 15-second pulse counts were punched onto IBM cards for analyses by a Control Data 3600 computer, using a variety of programs.

Correlation Analysis

Utilizing all data, including anthropometric measures, raw (15-second) pulse counts and transformed (combined) pulse counts, Pearson product-moment correlation coefficients were computed for all possible combinations of variables.

It was interesting to note that the correlations showed a significant advantage, at the .05 level, for tall women with the Skubic-Hodgkins test. However, tall, long-legged women fared less well and had higher pulse rate responses during the recovery period following the eight-inch test. In addition, there was no significant correlation between heart rate response and height in the eight-inch test until 30 seconds of recovery time had elapsed. After 30 seconds, the taller women had higher pulse rates (Table 2).

There was a consistent trend, although non-significant, for long-legged women to have lower heart rates and faster recovery times following the Skubic-Hodgkins test. These women were at a significant disadvantage in terms of heart rate response during the entire recovery period after the eight-inch test (Table 2).

In nearly every recovery interval, the heavier women had higher pulse rates following the Skubic-Hodgkins test; whereas, there were only a few significant correlations between pulse counts and weight, during the recovery intervals after the eight-inch test. This is illustrated in Table 3.

TABLE 2.--Height and leg length correlations with pulse counts following the Skubic-Hodgkins and eight-inch tests.*

Recovery Intervals (Seconds)		Height	Leg Length
S.-H	0:00-0:15	NS	NS
	0:15-0:30	-.19	NS
	0:30-0:45	NS	NS
	0:45-1:00	-.16	NS
	1:00-1:15	-.20	NS
	1:15-1:30	-.18	NS
	0:00-0:30	-.18	NS
	0:00-0:45	-.18	NS
	0:00-1:00	-.18	NS
	0:00-1:15	-.19	NS
	0:00-1:30	-.19	NS
	0:15-0:45	-.18	NS
	0:15-1:00	-.18	NS
	0:15-1:15	-.18	NS
	0:15-1:30	-.19	NS
	0:30-1:00	NS	NS
	0:30-1:15	-.18	NS
	0:30-1:30	-.18	NS
	0:45-1:15	-.18	NS
	0:45-1:30	-.19	NS
	1:00-1:30	-.19	NS
8"	0:00-0:15	NS	.25
	0:15-0:30	NS	.24
	0:30-0:45	NS	.23
	0:45-1:00	.17	.24
	1:00-1:15	.22	.28
	1:15-1:30	.20	.29
	0:00-0:30	NS	.26
	0:00-0:45	NS	.26
	0:00-1:00	NS	.26
	0:00-1:15	NS	.27
	0:00-1:30	NS	.28
	0:15-0:45	NS	.24
	0:15-1:00	NS	.25
	0:15-1:15	NS	.26
	0:15-1:30	NS	.27
	0:30-1:00	NS	.24
	0:30-1:15	.17	.26
	0:30-1:30	.18	.27
	0:45-1:15	.20	.27
	0:45-1:30	.20	.28
	1:00-1:30	.21	.29

*Correlation between height and leg length; r=.82.
$r_{.05} = .16$

TABLE 3.--Correlations of weight with pulse counts following the Skubic-Hodgkins and eight-inch tests.*

Recovery Intervals (Seconds)	Skubic-Hodgkins	Eight-Inch
0:00-0:15	.20	NS
0:15-0:30	.19	NS
0:30-0:45	.24	NS
0:45-1:00	.21	NS
1:00-1:15	.18	NS
1:15-1:30	NS	NS
0:00-0:30	.21	NS
0:00-0:45	.23	NS
0:00-1:00	.23	NS
0:00-1:15	.22	NS
0:00-1:30	.20	NS
0:15-0:45	.22	NS
0:15-1:00	.22	NS
0:15-1:15	.21	NS
0:15-1:30	.19	NS
0:30-1:00	.23	NS
0:30-1:15	.22	.17
0:30-1:30	.19	.18
0:45-1:15	.20	.20
0:45-1:30	.17	.20
1:00-1:30	NS	.21

*$r_{.05} = .16$

Overall, the eight-inch test appears to be less sensitive to weight, height and leg length variations.

Since the Skubic-Hodgkins test yields a cardiovascular index based upon the last 30-seconds of the one-and-one-half-minute recovery period, it was necessary to correlate the pulse counts during this specific interval with the pulse counts during all of the eight-inch recovery intervals (Table 4). Although all of these correlations were significant at the .05 level, the results were deemed inconclusive by virtue of the fact that the variability in the eight-inch

TABLE 4.—Correlations between pulse counts during the Skubic-Hodgkins 1:00-1:30 recovery interval and all recovery intervals following the eight-inch test.*

Recovery Intervals (seconds) Eight-Inch Step Test	Correlation Coefficients
0:00-0:15	0.4412
0:15-0:30	0.5178
0:30-0:45	0.5195
0:45-1:00	0.4849
1:00-1:15	0.4097
1:15-1:30	0.3814
0:00-0:30	0.5073
0:00-0:45	0.5270
0:00-1:00	0.5349
0:00-1:15	0.5250
0:00-1:30	0.5074
0:15-0:45	0.5294
0:15-1:00	0.5306
0:15-1:15	0.5147
0:15-1:30	0.4943
0:30-1:00	0.5164
0:30-1:15	0.4948
0:30-1:30	0.4723
0:45-1:15	0.4614
0:45-1:30	0.4421
1:00-1:30	0.4044

*All correlations significant at .05 level.

test results could account for only a small amount of the variability in the Skubic-Hodgkins data. Even the highest correlation of .53, which occurred in the recovery interval of 0:00-1:00 after the eight-inch test, could account for only 29 per cent of the variability.

It was then reasoned that a high correlation might not be the answer to the problem at hand. The object of this study was to find a simple cardiovascular efficiency test that is at least as effective, if not more so, than the

Skubic-Hodgkins test currently in use. If two tests measure the same thing equally well, then high correlations between test results should exist. Conversely, if one test is capable of discriminating between subjects to a greater degree than the other, or is of a higher power, then one would not necessarily expect to find high correlations between test results. Working under this assumption, it was decided to further test the power of each of the step tests used in this study by means of analyses of variance between selected groups of subjects.

Analyses of Variance

To facilitate the analyses of variance, three subgroups of subjects in good, average, and poor cardiovascular condition were selected from among the total sample of 103 students. Thirty students were selected--ten for each subgroup. The subjects were assigned to subgroups on the basis of subjective evaluation by the investigator, stated levels of activity, resting pulse rates, and height-weight relationships. Unfortunately, this technique is far from precise and leaves a great margin for error.

One possible source of error was eliminated in the selection of the subgroups due to the fact that the total sample was quite homogeneous with regards to anthropometric measurements. There were no significant differences between the total sample of 103 subjects, the selected sample of thirty subjects, and any of the subgroups of ten subjects classified as good, average, and poor, as far as height and leg length were concerned. However, there was a significant difference between groups on the basis of weight at the .01 level (Table 5). The subjects in the poor subgroup weighed more than the subjects in the other groups.

TABLE 5.--Height, weight, and leg length characteristics of the total sample, the selected samples and subgroups of good, average and poor cardiovascular condition.

Group	Height		Weight		Leg Length	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Total 103	165.21	5.61	59.07	6.88	82.50	4.20
Sub-Group Total 30	166.40	5.92	61.38	6.28	82.88	3.23
Good Gr.I	165.54	6.16	57.60	3.79	82.74	3.99
Average Gr.II	167.59	5.71	57.89	3.90	83.00	2.92
Poor Gr.III	166.07	5.59	68.59	6.57	82.83	3.07

The analyses of variance to determine the discriminatory powers of the tests between subgroups were performed on the results obtained during the Skubic-Hodgkins 1:00-1:30 recovery interval and during all of the 15-second and combined recovery intervals of the eight-inch test.

The Skubic-Hodgkins test was shown to be highly discriminative between subgroups, with a significance level of less than .0005 (Table 6). Although none of the pulse counts in the eight-inch recovery intervals were as

TABLE 6.--Analysis of variance: Skubic-Hodgkins 1:00-1:30, and all eight-inch recovery intervals.*

Recovery Interval	$\bar{X}$	S.D.	F	P
Skubic-Hodgkins 1:00-1:30				
Group I	61.0	3.40	12.02	<.0005
Group II	62.8	5.47		
Group III	70.0	3.86		
Eight-inch 0:00-0:15				
Group I	32.2	3.36	1.28	.295
Group II	34.3	1.89		
Group III	33.9	3.79		
Eight-inch 0:15-0:30				
Group I	26.4	4.04	1.27	.297
Group II	27.2	2.83		
Group III	26.9	3.79		
Eight-inch 0:30-0:45				
Group I	23.2	3.64	2.98	.067
Group II	24.0	2.50		
Group III	26.5	3.21		
Eight-inch 0:45-1:00				
Group I	19.7	3.06	6.70	.004
Group II	21.7	2.41		
Group III	24.3	2.95		
Eight-inch 1:00-1:15				
Group I	18.3	3.02	7.04	.003
Group II	20.3	1.95		
Group III	23.1	3.45		
Eight-inch 1:15-1:30				
Group I	18.2	3.05	5.60	.009
Group II	20.4	2.03		
Group III	22.7	3.71		
Eight-inch 0:00-0:30				
Group I	58.6	7.26	1.15	.323
Group II	61.5	3.84		
Group III	62.9	7.08		
Eight-inch 0:00-0:45				
Group I	81.8	10.78	1.76	.192
Group II	85.5	5.91		
Group III	89.3	9.44		
Eight-inch 0:00-1:00				
Group I	101.5	13.14	2.99	.067
Group II	107.2	7.45		
Group III	113.6	11.78		
Eight-inch 0:00-1:15				
Group I	119.8	15.75	3.91	.032
Group II	127.5	8.96		
Group III	136.7	14.85		
Eight-inch 0:00-1:30				
Group I	135.0	18.90	4.33	.023
Group II	147.9	10.87		
Group III	159.4	15.34		
Eight-inch 0:15-0:45				
Group I	49.6	7.6[illegible]	[illegible]	[illegible]
Group II	51.2	[illegible]		
Group III	55.4	[illegible]		
Eight-inch 0:15-1:00				
Group I	69.3	10.19	3.47	.046
Group II	72.9	7.05		
Group III	79.7	9.37		
Eight-inch 0:15-1:15				
Group I	87.6	12.89	4.42	.022
Group II	93.3	8.73		
Group III	102.8	12.65		
Eight-inch 0:15-1:30				
Group I	105.8	15.69	4.76	.017
Group II	113.6	10.69		
Group III	125.5	16.17		
Eight-inch 0:30-1:00				
Group I	42.9	6.39	4.96	.015
Group II	45.7	4.55		
Group III	50.8	5.96		
Eight-inch 0:30-1:15				
Group I	61.2	9.13	5.88	.009
Group II	66.0	6.32		
Group III	73.9	9.30		
Eight-inch 0:30-1:30				
Group I	79.4	12.07	5.95	.007
Group II	86.4	8.24		
Group III	96.6	12.78		
Eight-inch 0:45-1:15				
Group I	38.0	5.87	7.18	.003
Group II	42.0	4.27		
Group III	47.4	6.35		
Eight-inch 0:45-1:30				
Group I	56.2	8.87	6.84	.004
Group II	62.4	6.11		
Group III	70.1	9.83		
Eight-inch 1:00-1:30				
Group I	36.5	6.04	6.49	.005
Group II	40.7	3.83		
Group III	45.8	7.01		

*I = Group in good cardiovascular condition.
II = Group in average cardiovascular condition.
III = Group in poor cardiovascular condition.

leg length were concerned. However, there was a significant difference between groups on the basis of weight at the .01 level (Table 5). The subjects in the poor subgroup weighed more than the subjects in the other groups.

TABLE 5.--Height, weight, and leg length characteristics of the total sample, the selected samples and subgroups of good, average and poor cardiovascular condition.

Group	Height		Weight		Leg Length	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Total 103	165.21	5.61	59.07	6.88	82.50	4.20
Sub-Group Total 30	166.40	5.92	61.38	6.28	82.88	3.23
Good Gr.I	165.54	6.16	57.60	3.79	82.74	3.99
Average Gr.II	167.59	5.71	57.89	3.90	83.00	2.92
Poor Gr.III	166.07	5.59	68.59	6.57	82.83	3.07

The analyses of variance to determine the discriminatory powers of the tests between subgroups were performed on the results obtained during the Skubic-Hodgkins 1:00-1:30 recovery interval and during all of the 15-second and combined recovery intervals of the eight-inch test.

The Skubic-Hodgkins test was shown to be highly discriminative between subgroups, with a significance level of less than .0005 (Table 6). Although none of the pulse counts in the eight-inch recovery intervals were as discriminative as the Skubic-Hodgkins test, several were highly significant. The best 15-second recovery interval discrimination occurred at 1:00-1:15, and the best combined-interval discrimination between subgroups occurred at 0:45-1:15 (Table 6).

It was then necessary to obtain comparisons between individual subgroup means to determine how well the two tests could distinguish between each of the pairs of subgroups. The Tukey (T-Method) Test of the Comparison of Means two at a time following an F-test was applied to the data (8). The results showed that the Skubic-Hodgkins test discriminated between the means of the good and poor subgroups, and the average and poor subgroups at the .01 level, but did not discriminate between the good and average subgroups. The eight-inch test differentiated between the means of the good and poor subgroups at the .01 level, but did not discriminate significantly between the good and average, or the average and poor subgroups during either of the two best eight-inch recovery intervals: 1:00-1:15, and 0:45-1:15 (Table 7).

It is of interest to note that in this investigation, the Skubic-Hodgkins test did not have the discriminative power ascribed to it by its authors (27). A possible explanation for this situation is provided by the homogeneity of the sample of subjects used in the study. The young college women involved were all in good health, maintained roughly the same level of activity which could in no way be

TABLE 7.--Tukey test of comparison of means two at a time following an F-test: Skubic-Hodgkins 1:00-1:30 and eight-inch 1:00-1:15 and 0:45-1:15 recovery intervals.*

Comparison	Significance
Skubic-Hodgkins 1:00-1:30	
I , II	NS
I , III	.01
II , III	.01
Eight-Inch 1:00-1:15	
I , II	NS
I , III	.01
II , III	NS
Eight-Inch 0:45-1:15	
I , II	NS
I , III	.01
II , III	NS

*I = Group in good cardiovascular condition.
II = Group in average cardiovascular condition.
III = Group in poor cardiovascular condition.

considered either sedentary or unusually active, and were in a highly restricted age group. The sample was drawn from what one could easily term a 'typical' population of female college students. The ideal simple test or evaluative tool would be capable of discriminating between levels of cardiovascular efficiency within such a group as this, rather than only between extreme groups, such as highly trained, active, and sedentary individuals which are not typical of college populations.

As a further step in the search for such an evaluative tool, it was decided to re-evaluate the data, in light of pulse counts and anthropometric measurements, with the idea of eliminating the effects of height and weight upon the cardiovascular responses to a test performance. For this purpose, new variables were generated, making use of the pulse counts during the two most discriminative recovery intervals following the eight-inch test (1:00-1:15 and 0:45-1:15) and anthropometric data. Analyses of variance were employed to determine the effectiveness of the new transformed variables in differentiating between subgroups.

The transformed variables were:

1. $\frac{\text{Pulse count in the interval X height}}{\text{Weight}} = \frac{\text{P.C. x Ht.}}{\text{Wt.}}$

2. $\frac{\text{Pulse count in the interval X weight}}{\text{Height}} = \frac{\text{P.C. X Wt.}}{\text{Ht.}}$

3. Pulse count in the interval X weight X height = P.C. X Wt. X Ht.

4. $\frac{\text{Pulse count in the interval}}{\text{Height X Weight}} = \frac{\text{P.C.}}{\text{Ht. X Wt.}}$

After this treatment of the data, the greatest significance was found to occur when the pulse counts during either of the recovery intervals under consideration were multiplied by weight and divided by height or were multiplied by both weight and height (Table 8).

As a further step in the search for such an evaluative tool, it was decided to re-evaluate the data, in light of pulse counts and anthropometric measurements, with the idea of eliminating the effects of height and weight upon the cardiovascular responses to a test performance. For this purpose, new variables were generated, making use of the pulse counts during the two most discriminative recovery intervals following the eight-inch test (1:00-1:15 and 0:45-1:15) and anthropometric data. Analyses of variance were employed to determine the effectiveness of the new transformed variables in differentiating between subgroups.

The transformed variables were:

1. $\frac{\text{Pulse count in the interval X height}}{\text{Weight}} = \frac{\text{P.C. x Ht.}}{\text{Wt.}}$

2. $\frac{\text{Pulse count in the interval X weight}}{\text{Height}} = \frac{\text{P.C. X Wt.}}{\text{Ht.}}$

3. Pulse count in the interval X weight X height = P.C. X Wt. X Ht.

4. $\frac{\text{Pulse count in the interval}}{\text{Height X Weight}} = \frac{\text{P.C.}}{\text{Ht. X Wt.}}$

After this treatment of the data, the greatest significance was found to occur when the pulse counts during either of the recovery intervals under consideration were multiplied by weight and divided by height or were multiplied by both weight and height (Table 8).

TABLE 8.--Analysis of variance: Eight-inch test recovery intervals, 1:00-1:15 and 0:45-1:15, and variables transformed by use of anthropometric data.

Source	$\bar{X}$	S.D.	F	p
(1:00-1:15 X Ht.)/Wt.				
Group I	52.74	9.31	1.29	.293
Group II	58.74	4.01		
Group III	56.43	10.52		
(1:00-1:15 X Wt.)/Ht.				
Group I	6.37	1.10	15.84	<.0005
Group II	7.04	1.03		
Group III	9.54	1.69		
1:00-1:15 X Ht. X Wt.				
Group I	175224.44	11550.90	13.13	<.0005
Group II	197703.74	30900.50		
Group III	263002.10	49627.62		
(1:00-1:15)/(Ht. X Wt.)				
Group I	.002	0.00	.78	.470
Group II	.002	0.00		
Group III	.002	0.00		
(0:45-1:15 X Ht.)/Wt.				
Group I	109.68	19.18	1.22	.311
Group II	121.49	8.64		
Group III	115.81	20.40		
(0:45-1:15 X Wt.)/Ht.				
Group I	13.22	2.04	17.06	<.0005
Group II	14.57	2.24		
Group III	19.57	3.25		
0:45-1:15 X Ht. X Wt.				
Group I	364041.13	73443.25	13.29	<.0005
Group II	409362.67	68048.12		
Group III	539521.04	93291.75		
(0:45-1:15)/(Ht. X Wt.)				
Group I	.004	0.002	.77	.475
Group II	.004	0.002		
Group III	.004	0.002		

Again, it was necessary to perform the Tukey Test for comparisons between means (8) to determine how well each of these two transformed variables could distinguish between pairs of subgroups (Table 9).

It is interesting to compare the discriminative power of the Skubic-Hodgkins test with that of the eight-inch test. When the raw data were analyzed, it was found that the Skubic-Hodgkins test gave better discrimination. However, when transformed variables were used, the eight-inch test had a discriminative power equal to that of the Skubic-Hodgkins test.

Because the discriminative powers of the two tests were found to be equivalent when transformed variables were used with the eight-inch test, it was felt that the eight-inch test would be a much better tool than the Skubic-Hodgkins test for general application. Several factors contributed to this decision. These were:

1. The eight-inch test is less sensitive to variations in weight and height.
2. The eight-inch test is much less demanding in terms of leg strength and endurance.
3. The eight-inch test is easier to administer and is less time-consuming.
4. The eight-inch test is well within the capabilities of most women in a wide age range and thus would be suitable for use in longitudinal studies.

TABLE 9.--Tukey test of comparison of means two at a time following an F-test: Eight-inch 0:45-1:15 and 1:00-1:15 recovery intervals.*

Comparison 0:45-1:15	Sign.	Comparison 1:00-1:15	Sign.
P.R. X Wt./Ht.		P.R. X Wt./Ht.	
I , II	NS	I , II	NS
I , III	.01	I , III	.01
II , III	.01	II , III	.01
P.R. X Wt. X Ht.		P.R. X Wt. X Ht.	
I , II	NS	I , II	NS
I , III	.01	I , III	.01
II , III	.01	II , III	.01

*I = Good cardiovascular condition group.
II = Average cardiovascular condition group.
III = Poor cardiovascular condition group.

In order to facilitate practical application of this test, it was felt that an index should be formulated so that practitioners, knowing the mean and standard deviation of such an index, could place their subjects on a cardiovascular fitness continuum.

Cardiovascular Fitness Index

An index was developed so as to convert the existing range of transformed scores into a continuum with the ideal score approaching 100. The pulse count during the 0:45-1:15 recovery interval multiplied by weight and divided by height $\left(\frac{\text{0:45-1:15 pulse count X Weight}}{\text{Height}}\right)$ was chosen as the transformed variable to be used in the construction of the index since this generated variable had the highest

F-statistic in the analyses of variance and a discriminative power equal to that of the Skubic-Hodgkins Test. However, this equation yields a limited range of relatively low scores. It was felt that for the test to have more practical application, the index should permit a wider range of scores and should present them in a manner readily understandable to research workers, clinicians, teachers and subjects. For this reason, a scale with its upper limit approaching 100 was deemed appropriate.

To obtain this scale, calculations were performed with the data obtained from the study subjects as well as with hypothetical extreme values of pulse counts, heights, and weights. It was found that the application of two constants to the transformed scores would cause the ideal cardiovascular fitness index score to approach the upper limit. The complete equation, including the constants, is:

$$\text{C. V. Index Score} = 110 - (1.43)\,\frac{\text{Pulse Count X Weight}}{\text{Height}}$$

When the raw data for all subjects were converted by the above equation, the scores achieved ranged from a low of 72.7 to a high of 93.7. Measures of central tendency and variation for the total sample, as well as for the three fitness subgroups, are listed in Table 10.

It must be kept in mind that these results are limited to a very homogeneous sample and that further research is necessary before a table of normative values can be established.

TABLE 10.--Measures of central tendency and variation of the Cardiovascular Fitness Index.

Subject Group	Mean	S.D.
Total population 103 subjects	88.02	4.18
Group I--Good 10 subjects	91.10	2.91
Group II--Average 10 subjects	89.17	3.20
Group III--Poor 10 subjects	82.00	4.65

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

There have been a number of recent studies related to the testing of cardiovascular fitness. However, most of the studies have been conducted on men. Lucien Brouha and his associates (2) developed one of the most popular tests of cardiovascular efficiency in 1943 at the Harvard Fatigue Laboratory. A less strenuous modification of this test was developed for women by Skubic and Hodgkins (29). However, even this test was found to be cumbersome to administer, and extremely severe for some subjects (13). In 1967, Brizendine (1) reported the results of a study on men in which the Harvard Step Test was compared with other step tests. An eight-inch test, of one-minute duration with a rate of 30 steps per minute, appeared to be useful for general application.

The main objective of the present study on women was to compare the more severe Skubic-Hodgkins Test with the easier, more feasibly administered eight-inch test. In addition, the influence of selected anthropometric variables, upon cardiovascular performance in step tests, was studied.

To obtain the information desired for comparative purposes, 103 young college women were tested on both procedures. Their heights, weights, and leg length measurements also were taken and recorded. In terms of these measurements, the group was remarkably homogeneous. Pulse counts following exercise were recorded electrically with the use of surface electrodes.

Pearson product-moment correlation coefficients were computed between all possible combinations of recovery pulse counts and anthropometric measurements. Thirty students, ten each in relatively good, average, and poor cardiovascular condition, were selected for various analysis of variance tests of significance of mean differences between subgroups. On the basis of these results, the discriminatory power of the eight-inch test warranted generation of new variables to eliminate any possible effects of anthropometric variability. Finally, a preliminary index of cardiovascular fitness was formulated.

Conclusions

From the results of the analyses of the data, the following conclusions were drawn:

1. Taller women had a significant advantage, at the .05 level, on the Skubic-Hodgkins Test. The eight-inch test created a disadvantage, in terms of heart rate response, for the taller, longer-legged women.

2. Heavier women consistently had higher pulse counts on the more strenuous Skubic-Hodgkins Test. There were no significant correlations between weight and any pulse count following the eight-inch test. The eight-inch test is less sensitive to weight, height, and leg length variations.
3. There was a significant correlation between the results of the two tests at the .05 level. However, the correlations did not indicate the existence of a large amount of concomitant variability in the data.
4. Analyses of variance showed that the Skubic-Hodgkins Test had a greater power of discrimination than did the raw data from the eight-inch test. The Skubic-Hodgkins Test was capable of discriminating between the good and poor, and the average and poor, but not the good and average cardiovascular fitness subgroups at the .01 level. The eight-inch test could only discriminate between the good and poor cardiovascular fitness subgroups at that level of significance. However, it was found that when the eight-inch test pulse counts for the 0:45-1:15 and 1:00-1:15 recovery intervals were either multiplied by weight and divided by height, or multiplied by weight and height, the same differentiating ability as the Skubic-Hodgkins Test was achieved.

5. The cardiovascular fitness index, calculated on the basis of this study, holds promise as a possible tool for the evaluation of subjects on a fitness continuum. Such a continuum could also be used for the development of a table of normative values for use with the eight-inch test. The index equation derived from the data was determined mathematically to be;

Cardiovascular Fitness Index =

$$110 - (1.43)\ \frac{\text{0:45-1:15 count X Weight}}{\text{Height}}$$

6. The eight-inch test is well with the capabilities of most women in a wide age range and would, therefore, be suitable for use in longitudinal studies.

Recommendations

The results of this study have indicated directions for further research. Recommendations are:

1. Tests of reliability and validity should be applied to the eight-inch test. The validity of the cardiovascular index, especially, should be evaluated.
2. The eight-inch test requires further study to develop normative values. For this reason, it

is suggested that a larger, more heterogeneous sample be used. The sample should be separated into objectively derived fitness subgroups prior to testing.

BIBLIOGRAPHY

1. Brizendine, Stanley Farrell. A Comparative Study of the Harvard Step Test with Tests of Varied Times and Heights. Unpublished Master of Arts Thesis, Michigan State University, 1967.

2. Brouha, Lucien. "The Step Test: A Simple Method of Measuring Physical Fitness for Muscular Work in Young Men." Research Quarterly. 14:31-35, March,

3. Brouha, L., Radford, E. P. The Cardiovascular System in Muscular Activity. Edited by W. R. Johnson. Science and Medicine of Exercise and Sports. New York: Harper and Row, 1960.

4. Campbell, J. M. H. "Weight, Vital Capacity, Pulse Rate Before and After Exercise and Physical Fitness in Health." Guy's Hospital Reports. 75:263-281, 1925.

5. Clarke, Harriet L. "A Functional Physical Fitness for College Women." Journal of Health Physical Education and Recreation. 14:7, September, 1943.

6. Flanagan, K. "Pulse Ratio Test as a Measure of Athletic Endurance in Sprint Running." Research Quarterly. 6:64, October, 1935, Supplement.

7. Franklin, H. and Farmer, D. "The Reliability of the Pulse Ratio Test." Research Quarterly. 9:81, May 1938.

8. Guenther, William C. Analysis of Variance. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1964.

9. Hambly, W. D., et al. "Tests for Physical Efficiency: Part II." Guy's Hospital Reports. 72:367-385, 1922.

10. Hambly, W. D.; Pembrey, M. S.; and Warner, E. C. "The Physical Fitness of Men Assessed by Various Methods." Guy's Hospital Reports. 75:388-394, 1925.

11. Henry, F. M. "Influence of Athletic Training on the Resting Cardiovascular System." Research Quarterly.

12. Henry, F. M., and Kleeberger, F. L. "The Validity of the Pulse-Ratio Test of Cardiac Efficiency." Research Quarterly. 9:32, March, 1938.

13. Hodgkins, Jean, and Skubic, Vera. "Cardiovascular Efficiency Test Scores for College Women in the United States." Research Quarterly. 34:454-461, December, 1963.

14. Hodgson, P.; Lopez, A.; Pilliard, M.; and Newman, A. J. "A Study of Some Relationships Between Performance Tests and Certain Physiological Measures Associated with Maximal and Submaximal Work." Research Quarterly. 17:208-224, October, 1946.

15. Hrdlicka, Ales. Practical Anthropometry. Stewart, T. D. (Ed.). Philadelphia: Wistar Institute of Anatomy and Biology, 1947.

16. Keen, E. N. and Sloan, A. W. "Observations on the Harvard Step Test." Journal of Applied Physiology. 13:241-243, 1958.

17. Massey, B. H.; Husman, B. F.; and Kehoe, C. L. "The Effect of Posture on the Brachial Sphygmogram as an Indicator of Cardiovascular Condition." Research Quarterly. 24:194, May, 1953.

18. McCurdy, J. H. and Larson, L. A. "Study of Organic Efficiency of United States Veterans at the Domiciliary Facility, Bay Pines, Florida." Research Quarterly. 8:3, May, 1937.

19. Metheny, E.; Brouha, L.; Hohnson, R. E.; and Forbes, W. H. "Physiologic Response of Women and Men to Moderate and Strenuous Exercise." American Journal of Physiology. 137:318, September, 1942.

20. Meyers, C. R. and Blesh, T. E. Measurement in Physical Education. New York: The Ronald Press Company, 1962.

21. Montagu, M. F., and Ashley. An Introduction to Physical Anthropology. 3rd ed. Springfield, Illinois: Charles C. Thomas Publications, 1960.

22. Montoye, H. J. "Interrelation of Maximum Pulse Rate During Moderate Exercise, Recovery Pulse Rate, and Post-Exercise Blood Lactate." Research Quarterly. 24:453, December, 1953.

23. Murphy, M. A. "A Study of the Primary Components of Cardiovascular Tests." Research Quarterly. 11:57-71, March, 1940.

24. Phillips, M.; Ridder, E.; and Yaekel, H. "Further Data on the Pulse-Ratio Test." Research Quarterly. 14:425-429, December, 1943.

25. Salit, E. P. and Tuttle W. W. "The Validity of Heart Rate and Blood Pressure Determinations as Measures of Physical Fitness." Research Quarterly. 15:253, October, 1944.

26. Scott, G. M., Mordy, and Wilson. "Validation of Mass Type Physical Fitness Tests with Tests of Work Capacity." Research Quarterly. 16:2, May, 1945.

27. Skubic, V., and Hodgkins, J. "Cardiovascular Efficiency Test for Girls and Women." Research Quarterly. 34:191-198, 1963.

28. Skubic, V., and Hodgkins, J. "Cardiovascular Efficiency Test Scores for Junior and Senior High School Girls in the United States." Research Quarterly. 35:184-192, 1964.

29. Sloan, A. W. "A Modified Harvard Step Test for Women." Journal of Applied Physiology. 14:985-986, 1959.

30. Sloan, A. W. "Effect of Training on Physical Fitness of Women Students." Journal of Applied Physiology. 16:167-169, 1961.

31. Tuttle, W. W. "The Use of the Pulse-Ratio Test for Rating Physical Efficiency." Research Quarterly. 2:5-17, 1931.

32. Tuttle, W. W., and Dickinson, R. E. "A Simplification of the Pulse-Ratio Technique for Rating Physical Efficiency and Present Condition." Research Quarterly. 9:73-80, 1938.

33. Tuttle, W. W., and Frey, Henryetta. "A Study of the Physical Efficiency of College Women as Shown by the Pulse-Ratio Test." Research Quarterly. 1:17-25, 1930.

34. Tuttle, W. W., and Skein, J. S. "The Efficiency Rating of High School Boys as Shown by the Pulse-Ratio Test." Research Quarterly. 1:19-33, 1931.

35. Tuttle, W. W., and Wells, G. "The Response of the Normal Heart to Exercises of Graded Intensity." Arbeitsphysiologie. IV:519-526, 1931.

36. Walters, C. E. "A Study of the Effect of Prescribed Strenuous Exercises on the Physical Efficiency of Women." Research Quarterly. 24:102, 1953.

37. Wyndham, C. H. et al. "Maximum Oxygen Intake and Maximum Heart Rate During Strenuous Work." Journal of Applied Physiology. 14:927-936, 1959.

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 03061 5227