A MODIFICATION OF THE OHIO STATE UNIVERSITY STEP TEST FOR JUNIOR HIGH SCHOOL AGE GIRLS

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF ARTS IN PHYSICAL EDUCATION

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

COLLEGE OF

HEALTH, PHYSICAL EDUCATION, AND RECREATION

BY

MARIAN FRANCES CROWLEY, B. S.

DENTON, TEXAS
MAY, 1972

Texas Woman's University

Denton, Texas

May 1, 19 72

We hereby recommend that the Thesis prepared under our supervision by Marian F. Crowley entitled A Modification of the Ohio State University Step Test for Junior High School Age Girls

be accepted as fulfilling this part of the requirements for the Degree of Master of Arts.

Committee:

Chairman

Accepted:

Dean of Graduate Studies

ACKNOWLEDGEMENTS

A sincere appreciation goes to the members of my thesis committee: Dr. Marilyn Hinson and Dr. Gene Hayes. A special thanks goes to Dr. Joel Rosentswieg who has been a guiding force since my freshman year at TWU and throughout my professional career. Without his help it is doubtful that I would have ever started graduate work and almost certain that I would not now be completing an advance degree.

I would like to thank my co-workers -- Lennie Cribbs and Jim McKinney -- without their help the data for this study could not have been collected. A very special thanks goes to Ann Farmer. Her encouragement and understanding in times of stress and frustration was the driving force behind the completion of this paper.

And last, but not least, I would like to express my thanks to my parents. Their love and sacrifices made possible the B.S. degree -- their encouragement made possible the M.A.

TABLE OF CONTENTS

LIST OF TABLES

PAGE

TABLES

CHAPTER I

ORIENTATION TO THE STUDY

Introduction

Physical fitness has become a widespread concern to all who are involved in the development of health and welfare of the nation's youth. The meaning of the word fitness has been interpreted in many ways. According to Brouha and Gallagher (1943) there are three primary aspects of physical fitness: (a) "medical or static fitness" having to do with the soundness of the organs of the body; (b) "functional or dynamic fitness," having to do with the functional status, the ability to do strenuous work and physiological efficiency, and (c) the type of fitness that has to do with specific skills, muscle coordination and strength. There has been considerable activity on the part of physical educators to measure the present state of physical fitness of an individual and where necessary to improve it. In many junior high schools, physical fitness testing has been included as an intricate phase of the evaluative program, but relatively little research has been done concerning the measurement of "functional" fitness or cardiovascular efficiency for this age group.

The measurement of cardiovascular efficiency in a class situation has often been a problem because tests are time consuming, frequently require expensive equipment, and necessitate a highly motivated subject to exert the effort required to complete the test. It is believed that these problems can be reduced, if not eliminated, by the use of a sub-

maximal test which would terminate the work before the subject reaches exhaustion (deVries, 1966). A step test, which is a test requiring stepping up and down from a stool as a means of measuring the subject's cardiovascular reaction to work, can be administered to many subjects quickly. The step test may be completed in a very short time allowing the teacher to emphasize the teaching/learning aspect of education. The equipment required for the administration of a step test consists basically of a bench or step of the proper height. The problem of motivation is believed to be minimized by the use of a submaximal test because the test is terminated before the subject is exhausted by the work. Submaximal step tests have been found to correlate at acceptable levels with maximum tests. Truett (1966) found in a study with college men that the length of exercise and the heart rate had a linear correlation up to 180 BPM. Thus, it may be concluded that the maximum work may be predicted at an acceptable level from observation at less stressful levels of activity.

Rational for the Study

Previous studies reported using a step test as a measure of submaximal working capacity (SWC) have been conducted primarily among college students. There is a need for a similar test designed specifically for girls between the ages of twelve and fifteen. It is hoped that this research may result in a test of submaximal working capacity for junior high school age girls that is valid, reliable, objective, and administratively feasible.

Statement of the Problem

The study entailed the modification of the Ohio State University Step Test to determine the submaximal working capacity of one hundred

randomly selected girls enrolled in Bedford Junior High School in Bedford, Texas, during the spring semester of the academic school year of 1970-1971. A test-retest of the modified step test was used to establish the reliability and the Sjostrand Test on the bicycle ergometer was used to establish the validity. Upon the basis of the findings, a conclusion was drawn as to the reliability and validity of the present modification of the Ohio State University Step Test as a measure of the submaximal working capacity of junior high school age girls.

Definitions and/or Explanation of Terms

To contribute to a clear understanding of the problem, the following definitions and/or explanations of terms have been established for use throughout the study:

1. Submaximal Working Capacity - 150 (SWC_{150}): According to Kuruz, Fox and Mathews (1969), "the inning in which the subject's heart rate reaches 150 beats per minute will be recorded as that person's submaximal working capacity." An inning may be defined as thirty seconds of stepping and twenty seconds of rest.
2. Cardiovascular Efficiency: "...may be defined as the ability of the heart and the circulatory system to adjust to the stress of activity (Skubic and Hodgkins, 1963).
3. The Step. The investigator explains the step as a multilevel bench that enables each subject to select a height that will enable her to have an approximate ninety degree angle at both the hip and knee joints. (See Appendix for illustration).

Delimitations of the Study

The study was subject to the following delimitations:

1. The one hundred girls randomly selected from the girls enrolled in Bedford Junior High School during the spring semester of the academic school year of 1970-1971 as representative of the junior high school population.
2. The reliability, validity, and objectivity of the instruments and the techniques of measurement employed in the collection of the data.

Purpose of the Study

The general purpose of the study was to develop a test that could measure the submaximal working capacity of junior high school age girls with a minimum amount of equipment. The specific purpose of the study was to test the null hypothesis that the modified Ohio State University Step Test is not a reliable and valid measure of the submaximal working capacity of junior high school age girls.

Summary

The assessment of dynamic fitness or cardiovascular efficiency should be an integral part of the evaluation of physical fitness in the public schools. Most physical educators fail to measure working capacity because many of the tests are too time consuming, require expensive equipment and necessitate highly motivated subjects. The use of a submaximal step test may reduce or eliminate these problems.

The purpose of the study was to modify the Ohio State University Step Test and to determine if it was a reliable and valid measure of the submaximal working capacity of junior high school age girls. One hundred junior high school age girls from a single school were randomly selected as subjects. All subjects were given a test-retest on the

modified step test and thirty subjects from the original one hundred were tested on the bicycle ergometer to determine a validity coefficient.

Chapter II presents a review of the literature that was found pertinent to the study.

CHAPTER II

RELATED LITERATURE

Introduction

The purpose of the study was to modify the Ohio State University Step Test in order that a valid and reliable measure of the submaximal working capacity of junior high school age girls might be obtained. One hundred girls enrolled in Bedford Junior High School in Bedford, Texas, during the spring semester of the academic school year of 1970-1971 were randomly selected as subjects. The test-retest protocol on the modified step test was used to establish the reliability of the measure and thirty of the initial subjects were tested one time on the bicycle ergometer to establish a coefficient of validity.

Chapter II presents the investigations reported in the literature that are most closely related to the present study. The studies appear under the following headings: The Ohio State University Step Test, Step Tests, and Studies Relating to Testing Procedures.

The Ohio State University Step Test

Kuruz, Fox and Mathews (1969) constructed a submaximal cardiovascular step test at the Ohio State University in Columbus, Ohio. The subjects were seventy-five male volunteers of the Ohio State University ranging in age from nineteen to forty-six.

The equipment for the test consisted of a split-level bench, one level being fifteen inches and the other being twenty inches high,

an adjustable hand bar, a metronome and a stop watch. The test consisted of eighteen innings of fifty seconds duration. Each inning was divided into a thirty second work period and a twenty second rest period. During the rest period, the subject's pulse was counted for ten seconds. The investigator started counting after five seconds of rest had occured and stopped counting on the fifteenth second of the rest period. There were three different work-loads as follows:

Phase I consisted of six innings at 24 steps a minute cadance on the 15 inch bench.

Phase II consisted of six innings at 30 steps a minute cadance on the 15 inch bench.

Phase III consisted of six innings at 30 steps a minute cadance on the 20 inch bench.

The three phases are continuous. The test is terminated when the pulse rate reaches twenty-five for the ten second counting period or 150 BPM. The inning in which the subject's pulse reaches 150 BPM is that subject's score.

The subject stands in front of the bench grasping the chest high restraining bar with both hands. On the command "up" the subject should step up on the bench and straignten his leg completely and continue to do so throughout the work phase of the innings. After the ten second counting period and before the seventh inning, the subject should be told that the cadance will become faster. After the pulse count and before the thirteenth inning, the subject should be told to move over in front of the twenty inch bench. The test is terminated after the eighteenth inning or as soon as the subject's pulse rate reaches 150 BPM.

The reliability coefficient for the test as determined by the test-retest method was .94. The validity coefficient of .94 was obtained by comparing the results of the step test for thirty of the subjects to the results of those same subjects on the Balke Treadmill Test.

The findings of the study may be summarized as follows: (1) the test does not overstress the individual; (2) the exercise commences slowly so that even the poorest of conditions can be measured; (3) the third phase (30 steps per minute on the 20 inch bench) requires the subjects to draw on their anaerobic reserves as evidenced by the high volume of oxygen and the high respiratory exchange ratio; and (4) the mechanical efficiency appears to remain constant even at the more strenuous work loads. The test seemed to be sensitive to changes in cardiovascular fitness also as evidenced by the unreliable scores of those subjects who were engaged in a strenuous activity program (not a part of the study itself).

Cotton conducted a study during the fall and winter quarters of 1968-1969 at the Georgia Southern College to develop a modification of the Ohio State University Step Test of cardiovascular fitness which would be satisfactory for mass testing in a class situation. The subjects were 36 male physical education major students between the ages of 20 and 30. The reliability of the modified step test and its correlation with the Balke Treadmill Test was computed.

The modified test was the same as the original Ohio State University Step Test designed by Kuruz, Fox and Mathews with the following exceptions: (a) a 17 inch step (bleachers) was used; (b) no handrail was used; (c) in order to provide an increased work load without the higher step, the stepping cadance is increased to 36 steps per minute in

Phase III; (d) a "buddy system" was used to count the pulse rate. The commands and cadances for the eighteen innings were prerecorded on tape. Each subject had a score sheet listing innings one through eighteen.

Each subject took a test-retest on the Balke Treadmill Test within a week. Prior to each test, the subjects rested for thirty minutes in a seated position. The treadmill test was terminated when the subject's pulse rate reached 180 BPM and the time required for the 180 BPM rate to be reached was the measure of that subject's fitness level. Two mass testing sessions were held within a week of each other so that each subject could be tested on the modified step test. The terminating pulse rate for the step test was 150 BPM. The inning in which the subject's pulse rate reached 150 BPM was used as his fitness score. No subject was given more than one test on a single day and all were tested within a three week period.

Test-retest reliability of the Modified Step Test was .95; reliability of the Balke Treadmill Test was .94. The modified test correlation of .84 with the Balke Test is acceptable, but does not compare favorably with the correlation between the Ohio State University Step Test and the Balke Treadmill Test (.94). The Modified Step Test appears to be a satisfactory measure of cardiovascular fitness. It also has the advantage of not requiring the student to exert maximum effort or possess a high degree of skill.

To evaluate the Modified Step Test in a high school situation, it was administered to 46 boys, grades nine to twelve, in a physical education class at Statesboro High School, Statsboro, Georgia. The same procedure described previously was used. A test-retest was given within a week with the counting partners being changed for the retest. The

reliability of .75 was lower than that obtained for the college men and may be attributed to counting errors present on the score sheets.

Step Tests

Tuttle (1931) was among the first to conduct a study that attempted to measure physical efficiency which would not necessitate expensive equipment, a great deal of time, or an experienced administrator. In 1931 Tuttle developed the pulse-ratio test at the University of Iowa.

The performance selected was "stool climbing" on a thirteen inch bench that was large enough so that it could be mounted and dismounted with ease. The intensity of the exercise could be either increasing the stepping time or by increasing the number of steps taken per minute. The emphasis on a step test is on uniformity: height of the stool, cadance of stepping, duration of the exercise, the body position on the stool, the arm action and the manner in which the pulse is determined.

The pulse-ratio represents the ratio of the resting pulse rate to the pulse rate after exercise. This ratio is found by dividing the total pulse for two minutes after a known amount of exercise by the normal resting heart rate for one minute.

Brouha (1943) conducted a study at Harvard University, Cambridge, Massachusetts, to devise a step test that would measure the general capacity of the body, in particular the cardiovascular system, to adapt itself to hard work and the recovery from the activity. This was the first step test designed specifically to measure the general physical fitness for hard work for normal, healthy men. It has become famous as the Harvard

Step Test (HST). Brouha felt that it was important to have a way of assessing physical fitness in order to classify unknown subjects into excellent, good, average, or poor groups, to follow the progress of a given subject during a training period, to check the adequacy of a training program, to reveal deterioration due to lack of training, staleness, illness, or other variables.

The subjects for the test were the freshman entering college during the Fall semester of the academic school year of 1943. They were tested upon entering, and retested twelve weeks later after a training period of four hours of work a week. The test consisted of stepping up and down on a twenty inch platform thirty times per minute for five minutes. The pulse was counted at the carotid artery from 1 to $1\frac{1}{2}$, 2 to $2\frac{1}{2}$, and 3 to $3\frac{1}{2}$ minutes after the work stopped. The score was obtained by dividing the duration of the exercise by the sum of pulses in recovery according to the following formula:

$$\text{Index} = \frac{\text{duration of exercise in seconds} \times 100}{2 \times \text{sum of pulse counts in recovery}}$$

The heart rates during the recovery period following exercise were used in the index because a previous study (Brouha and Heath, 1943) indicated that the heart rate before exercise is usually not related to an individuals physical fitness.

The Harvard Step Test gave group scores comparable to those obtained with laboratory techniques such as the treadmill and the bicycle ergometer. The results of the study were as follows: the test differentiated between the levels of physical fitness in unknown subjects; the scores indicated the effects of training; the scores showed deterioration in fitness scores for the students who had been excused from the training

program because of injury or illness; and finally, the scores of the unfit improved more than the scores of the good group.

Approximately nine months after the completion of the study to devise the Harvard Step Test (HST), Brouha and Gallagher (1943) collaborated to modify the HST slightly so that it could be appropriate for girls. The subjects were the girls at the Phillips Academy, Andover, Massachusetts, during the 1943 academic school year. The aspect of fitness that was tested was the ability to perform strenuous work in a physiologically efficient manner.

The test consisted of stepping up and down on a sixteen inch bench for four minutes at the rate of thirty steps per minute. If the girl fell behind the cadance for fifteen seconds because of fatigue, the test was terminated. At the end of the test, the subject's pulse was taken by a partner at the carotid artery by the palpitation method. The pulse was counted for the first thirty seconds of each of the three minutes immediately following the exercise. The test is based on the principle that the more fit the individual, the more rapidly will the heart rate decelerate after hard work. The subject's score was calculated in the same manner as in the HST.

The study found that the HST could be considered a simple method for testing the dynamic physical fitness of adolescent girls; the method is rapid, reasonably accurate, and requires neither expensive equipment nor trained personnel for its administration.

Skubic and Hodgkins (1963) conducted a study to find a test of cardiovascular fitness for girls and women which would be less strenuous and easier to score than the HST. The study was limited to females between the ages of eleven and twenty-four and was based on the following assumptions: (a) that differences in states of training or condition

are reflected by differences in cardiovascular efficiency; (b) that those persons who are most fit can best adapt to a given amount of exercise; (c) that adaptations to exercise can be determined by the recovery of heart rate after exercise; and (d) that girls who have trained for at least three hours per day, six days per week for seven months are considered "trained" and in excellent state of cardiovascular fitness.

Skubic and Hodgkins modified the HST to a shorter period of stepping (three minutes) and used only one recovery period measurement of the heart rate. A preliminary assessment of the modified test was given to 96 females between the ages of twelve and twenty-five who volunteered to take both the five minute and the three minute step test. The rate of stepping for both test was 24 steps per minute on an 18 inch bench. For the modified test the pulse was taken for thirty seconds after one minute of rest after completion of the exercise. The only variable was the length of the tests. Some subjects were given the 5 minute test first, while others were given the 3 minute test first. The correlation between the two tests was .790 and was accepted as sufficiently high to warrant further study of the three minute test.

For the actual study the heart rate was monitored throughout exercise by means of telemetering of the EKG. In those tests where only the recovery pulse was needed, the pulse was taken by the palpitation method at the carotid artery.

The subjects for the study were divided into four groups:

1. Trained girls between the ages of 11 and 17 who were involved in a competitive swimming program in which they swam for 3-5 hours a day. They had been in training at least 7 months.
2. Untrained girls between the ages of 12 and 16 and were not trained.
3. Active women between the ages of 17 and 23 who participated in

daily vigorus activity.

4. Sedentary women between the ages of 17 and 23 who did not participate in any kind of vigorous activity on a regular basis.

The findings of the study indicated that the three minute step test was sufficiently strenuous to be classified as hard work for girls and women. The test was found to discriminate to a high degree among subjects who are moderately active, subjects in an excellent state of physical condition, and those who are sedentary. Age was not found to be a factor in the step test among females of junior high, high school, and college age. The conclusion that the modified step test was a valid and reliable instrument for determining the cardiovascular efficiency of girls and women appeared to be warrented.

After the initial study, Skubic and Hodgkins (1963) sought to establish national norms for the Three Minute Step Test. An attempt was made to secure an adequate sample from all parts of the nation and to include all types on institutions. The physical education departments who responded and indicated their willingness to participate in the study were sent complete instructions including a description of the test, directions for administering the test, suggested methods for obtaining a random sample of subjects, and directions for recording results. A total of 2,360 usable test scores were received from 66 institutions.

In order to present an accurate picture of the total sample of the population, a formula was devised to take into account both the recovery pulse rate and the number of seconds of exercise. It was found that the rapid form of the HST could be used with slight modification.

$$\text{cardiovascular eff. score} = \frac{\text{no. of sec. completed} \times 100}{\text{recovery pulse} \times 5.6}$$

All pulse rates were converted to efficiency scores by means of this formula and all further calculations made in this study utilized these scores rather than the recovery pulse rate.

When rated on a national scale for cardiovascular efficiency, the majority of the 2,360 college women in the study were rated fair. Thirteen percent of the subjects, however, could not complete the 3 minute step test and the majority of these subjects were rated poor and very poor. Subjects in the Eastern District (AAHPER Districts) had better cardiovascular efficiency scores than the subjects in the other five districts. The subjects in the Southern District made poorer scores than subjects in any other district. The height of the subject did not effect scores on the step test but there was a clear indication that heavy women tend to score less well than lighter women. And, finally, when the subject's scores were correlated with their major fields of study, the physical education major students scored the highest and the education major students scored the lowest from a cardiovascular efficiency standpoint.

In 1964 Skubic and Hodgkins (1964) conducted a study to gather scores from junior and senior high school age subjects in order to provide standard cardiovascular efficiency scores for girls of secondary school age and to compare these results with those of college women. They also wished to determine whether or not air temperature or diurnal variations affected the scores and to compare the scores obtained from the six AAHPER districts. Recovery pulse rates following a 3-minute step test were obtained from 686 junior high school and 1,332 high school age girls who were students in 55 different secondary schools throughout the six AAHPER districts.

Correlations were determined between the cardiovascular efficiency scores and the following variables: age, air temperature, time of testing, and the area of the country in which the subjects resided. A comparison of the 686 junior high school students and the 1,332 high school students indicated that the younger girls made significantly better scores than the older girls on the cardiovascular efficiency scores. When compared to the scores of the college women, the junior high school girls scored significantly higher, but there was no difference between the high school and the college scores. Subjects tested in cool temperature (between 54°F and 65°F) did not score differently from those subjects tested in the warm temperature (80°F to 90°F). Tests given at four different times during the day revealed that the time of testing does not affect test scores. And, finally, the high school students in the Central district made the best scores. In the junior high school level, the students from the Central and Southern districts made the best cardiovascular efficiency scores.

Day (1966) conducted a study that evaluated the reliability of the Rhyming Step Test for the prediction of aerobic capacity. The Rhyming Step Test is a modification of the HST and was designed to predict an individuals maximum oxygen intake. Fifty-two male undergraduate students at the University of Oregon in Eugene, Oregon, volunteered as subjects. Each subject was tested twice within a one week period. The pulse rate was monitored during exercise by an instrument that transmitted an audible beep with each heart beat. The heart beat rate was counted the last fifteen seconds of each minute. The exercise period consisted of stepping on a 40 cm. bench for five minutes or until the heart rate stabilized. A steady value was usually reached after two minutes and

this was the subject's score.

Statistical treatment of the data yielded the following information. The variables of age and university class correlated well only with each other (r=.630). The correlation between the subject's weight and the amount of work completed correlated very highly (r=.997). Body weight and aerobic capacity correlated well at .681 for the first test and .671 for the retest. Heart rate scores correlated positively with each other and negatively with aerobic capacity. The test-retest reliability was represented by the correlation between the predicted aerobic capacity scores on the first test and second test administration (r=.795). A high significant improvement between the first and second test occurred, possibly because of the subject's familarity of the test.

Brizendine (1967) conducted a study to determine whether a test could be devised having a shorter stepping time and a lower stepping height than that employed by the original HST that would yield results equal to or better than those yielded by the HST to facilitate the speed and ease of administration. An initial study using fifteen physical education major students from Andrews University indicated that after 225 tests (fifteen for each subject) that a correlation of .89 existed between the original HST and a test administered on an eight inch bench for one minute. The modified test was designated the B-1 test.

The final study employed twenty-nine subjects randomly selected from the same population as the subjects for the pilot study. A wider span existed in the physical fitness of the final group ranging from poor to very good. The subjects were given the B-1 test and twenty-four hours later the original HST. The statistical treatment of the data included a rank order correlation between the two tests, the application of the

ANOVA and the Scheffe' test to compare the means of the between and within groups on both the HST and the B-1 test.

The statistical treatment of the data yielded the following information. The B-1 test compared with the HST by rank order correlation with a high positive correlation of .89. The B-1 test produced the same results as the HST, that is, it classified the individuals into groups of good, average and poor in relationship to cardiovascular fitness. The results of the study revealed that the B-1 test and the HST both discriminated at the .05 level of confidence between the average and poor groups. Neither differentiated statistically between the good and average groups. The B-1 test distinguished between the good, average and poor groups with a higher level of significance than the HST.

Myers (1969) conducted a study concerned with the reliability of the HST. The subjects were thirty-one freshman men selected at random from the 116 students in the classes of one instructor in the basic instructional program in physical education at the State university of New York at Buffalo. The tests were given individually with the test procedures played on a tape recorder and all pulse counts taken by the same test administrator at the carotid artery. A reliability coefficient of .84 was computed by the Pearson Product-Moment method on a test-retest after a one-week interval using the HST.

In order to determine the reliability of the HST for a mass administration, Myers gave the test to 119 eighth grade boys from two different schools. The pulse rates were counted by the subjects in teams of two. They were given a test-retest on a one-week interval. A single test administrator gave all the tests utilizing the tape recorded testing procedure and teacher assistance in class organization. The reliability

coefficient was .65 which was thought to be encouraging considering the limited but difinitive instruction the students were given as to the method of counting the pulse at the carotid artery.

Harvey and Scott (1970) conducted a study to investigate the reliability and validity of the Kent State University (KSU) Step Test and the relationship between digitally self-determined and stethoscopically determined pulse rates. The KSU step test consists of stepping up and down on an eighteen inch bench for one minute at the rate of thirty steps per minute. The cadance is set by an electric metronome. The step test had been used for a number of years as a part of the Foundation for Movement course. Scoring was accomplished by each student digitally palpitating the carotid artery and recording both the resting pulse and the post-10 second exercise pulse in a ratio which was used as a cardio-vascular score:

$$\text{score} = \frac{\text{post-exercise pulse}}{\text{resting pulse}}$$

The subjects for the study were thirty-two freshman women at the Kent State University in Kent, Ohio, between the ages of seventeen and twenty and fourteen women athletes between the ages of eighteen and twenty. They were tested at the beginning of the fall semester using the KSU 1-minute step test and the 3-minute Skubic-Hodgkins test. The tests were administered on consecutive days. The subjects were retested on the KSU step test at the end of the quarter, after the completion of the Foundations for Movement course and at the end of the athletic seasons (field hockey and swimming).

During the test, resting and post-10 second exercise pulses were taken simultaneously with the stethoscope by the authors and at the

carotid artery by the subjects. A 30-second stethoscopic reading was recorded after one minute post-exercise because the Skubic-Hodgkins test had been validated at this point. Pulses were taken with the stethoscope each half-minute for 15 seconds for five minutes starting with the reading at one minute and forty-five seconds. The post-exercise procedures were identical for both tests so that multiple comparisons could be made. To determine the relationship between the selp-palpitation and the stethoscopic methods of counting the pulse rates, thirty-two college women were selected. They were unfamiliar with the KSU Test. The test was administered to them on two consecutive days while the pulse rates were checked simultaneously by self-palpitation and stethoscope. The rank order correlation of reliability was computed between the first and second test and correlation coefficients were determined between the two methods of counting pulse rates.

The following findings of the study were reported. The KSU One-Minute Step Test was found to be both valid and reliable as a measure of cardiovascular efficiency under the conditions of the study. The KSU Test distinguished between women athletes and non-athletes at a significant level. And, finally, that the students counting their own pulses at the carotid artery are sufficiently accurate to score the KSU Test.

Studies Relating to Testing Procedures

Adams, Bengtsson, Berven and Weglius (1957) conducted a study to determine the physical working capacity of 196 Swedish city and country children of both sexes between the ages of ten and twelve. One hundred and two of the subjects were from Stockholm and the remaining 94 were from two country schools in the same nation. The socioeconomic status

of most of the city children was middle and upper class while the country children's families had lived on the farm for two or more generations. From the group of city children, 47 were retested in September to determine the effect of the summer vacation on the working capacity.

The children were divided into three groups: highly trained, moderately trained, and poorly trained. Using the Sjostrand Method of determining working capacity, the children were asked to perform two consecutive work load trials on an electric bicycle ergometer. The rate of pedaling was maintained between 60-70 revolutions per minute. Each work trial lasted six minutes. The work loads were scheduled so that the first work load produced a heart rate of 140 BPM and the second work load a heart rate of 170 BPM. The working capacity was calaculated by plotting on graph paper the heart rate at the end of each trial against the work load and drawing a straight line through the two points and intersecting the line of 170 BPM. The estimated amount of work that would produce a heart rate of 170 BPM was recorded as the individual's physical working capacity.

All the collected material obtained on each subject was coded and the information transferred to cards that were analyzed electronically on a digital computer. Logarithms were used to make the regressions linear and the variability homogeneous. The data were analyzed to study the relationship of height, weight, surface area, working capacity, and heart volume to age and sex. It was found that the working capacity of these young subjects increased with age, height, weight, heart volume, and the degree of physical fitness.

A study of the data indicated that the Swedish boys showed greater working capacity than the girls of the same age, body size, and

heart volume. There were no significant differences in the slopes of the regression lines for the country and city boys or country and city girls.

Alderman (1965) undertook a study to study the fatigue curves for whole body work as a criterion for work capacity. Forty male college students from the University of Alberta in Edmonton, Alberta, Canada, were tested and retested forty-eight hours later on the bicycle ergometer. The heart rate was monitored continuously during the three experimental periods: 1) Pre-exercise period for five minutes; 2) the exercise period until 180 BPM were reached; and 3) the post-exercise period until the pulse slowed to 100 BPM. The friction on the bicycle ergometer started as .25 Kpm per minute and was increased .50 Kpm each minute during exercise. Data were yielded as to the length of the exercise period, basic recovery time, and recovery times for the intermediate heart rate levels of 140 and 160 BPM. The following conclusions were drawn from an analyzation of the data. The time scores increased in reliability as the heart rate criteria increased progressively from 100 to 160 BPM. The reliability at 160 BPM was found to be greater than that at 180 BPM. And, as the work load increased at a steady progression, the heart rate increased in a linear fashion up to 180 BPM.

McArdle, Zwiren, and Magel (1968) undertook a study to determine the validity of the post-exercise heart rate as a means of estimating the rate during work. Ten male volunteers from Queens College, New York, New York, were the subjects for the study. The average age of the subjects was twenty. To evaluate the effect of work intensity on the rate of recovery, each subject was progressively exercised to a work level which produced a heart rate of either 100, 120, 140, 160, or 180 BPM. The subjects pedaled at a rate of 50 rpm with the work load at 150 Kpm per minute until the predetermined pulse rate was reached. The desired

exercise was maintained for 20 to 30 seconds to determine the effect of the duration of work at a set heart rate on the rate of recovery. In order to secure an accurate reading of the heart rate during a specific recovery interval, cardiac cycles and fractions thereof were counted to the nearest .1 and extrapolated to beats per minute from the collected electrocardiograms.

The error in estimating the exercise heart rate from reading taken during the 10 seconds immediately following strenuous work (180 BPM) averaged 2.7%. This increased to 7.6% during the first 10 seconds of recovery from exercise rates of approximately 140 BPM. If one allowed 4 seconds to find the pulse and then counted during a 10-second interval, the average underestimation of the exercise heart rate increased to 5.7% in strenuous exercise and 13.5% in recovery from moderate exercise. The rate of recovery is not similar for all individuals. This variability may introduce serious errors when post-exercise heart rate is used to infer the heart rate during exercise.

Summary

The present study is based on the research of Kuruz, Fox and Mathews conducted at the Ohio State University. Their study modified previous step tests in that the work loads was divided into eighteen 50 second innings with a terminating point of 150 BPM so that even those in the poorest of physical conditions could be tested. At the same time, the last six innings were sufficiently strenuous to require the subjects to draw on their anaerobic reserves. Cotton modified the OSU Step Test for high school age boys in 1969 and found that the test was appropriate for mass testing of high school age boys.

Tuttle was among the first to conduct a study that was concerned with physical efficiency that did not necessitate expensive equipment, a great deal of time or an experienced administrator. In 1931 Tuttle developed the pulse-ratio technique to score his "stool climbing" test.

In 1943 Brouha designed a step test at Harvard University specifically to measure the general physical condition for hard work of normal, healthy men. The test was designed to classify unknown subjects into categories of excellent, good, average or poor. Approximately four months after the initial test, Brouha and Gallagher collaborated to modify the Harvard Step Test so that it would be appropriate for girls. Both tests yielded scores that were comparable to those obtained with more exacting techniques requiring expensive instrumentation.

In 1963 Skubic and Hodgkins conducted a study to find a test of cardiovascular fitness for girls and women which could be quickly and easily administered. They found that a three minute test was strenuous enough to be considered hard work for girls and women. After their initial study, Skubic and Hodgkins conducted two additional studies. In the first they established standards for college women and in the second they established norms for high school and junior high school age girls.

In 1966 Day completed a study that evaluated the reliability of the Rhyming Step Test which is a modification of the Harvard Step Test. The reliability coefficient as determined by the test-retest technique was .795.

Brizendine conducted a study in 1967 to determine the possibility of a test with a shorter stepping time and a lower bench yielding results similar to the original Harvard Step Test. Preliminary studies resulted in the selection of an eight inch bench and a one minute stepping time

(the B-1 test). The B-1 test distinguished between good, average and poor with an acceptable level of confidence.

In 1969 Myers conducted a study of the reliability of the Harvard Step Test as typically administered. A reliability coefficient of .84 was computed when the test was given individually by one test administrator. When the test was administered to a large group of eighth grade boys with the pulse rates counted by a partner, the reliability coefficient was .65.

Harvey and Scott conducted a study in 1970 to investigate the reliability and validity of the Kent State University Step Test and the relationship between digitally self-determined and stethoscopically determined pulse rates. The study indicated the KSU Step Test is both a valid and reliable measure of cardiovascular efficiency and that the self-palpitation method was sufficiently accurate to score the test.

The bicycle ergometer test used in the present study to determine the validity of the modified OSU Step Test was used by Adams and others in 1957 in a study to determine the Physical Working Capacity of 196 Swedish city and country children. The Sjostrand method was used to determine the PWC_{170}. The data was analyzed to study the relationship of height, weight, surface area, working capacity and heart volume to age and sex. It was found that the working capacity increases with age, height, weight, surface area, heart volume and the degree of physical fitness.

Alderman undertook a study in 1965 to study the fatigue curve for whole body work as a criterion for work capacity. The findings of the investigation were that the scores increased in reliability as the heart rate progressively increased from 100 to 160 BPM; the reliability

at 160 BPM was more reliable than that at 180 BPM; and as the work load increased at a steady progression, the heart rate increased in a linear fashion.

In 1968 McArdle, Zwiren and Magel conducted a study to determine the validity of the post-exercise heart rate as a means of estimating the rate during work. The average underestimation of the exercise heart rate increased to 5.7% in strenuous work and 13.5% in recovery from moderate exercise.

Chapter III presents the procedures used in the study.

CHAPTER III

PROCEDURES OF THE STUDY

Introduction

The purpose of the study was to determine the validity and the reliability of the modified Ohio State University Step Test as a measure of the submaximal working capacity of junior high school age girls. One hundred girls enrolled in Bedford Junior High School during the spring semester of the academic school year of 1970-1971 were randomly selected as subjects for the study. The subjects were given a test-retest on the modified step test, forty-eight hours apart, to establish the reliability of the test. Thirty subjects were selected randomly to be tested on the bicycle ergometer to establish the validity of the modified step test.

The procedures used in the development of the study are presented in this chapter under the headings of: Preliminary Procedures, Criteria for the Selection of the Instruments, Selection and Description of the Instruments, Selection and Description of the Subjects, Selection of the Student Assistants, Descriptions and Results of the Pilot Study, Procedures Followed in the Collection of the Data, Organization and Treatment of the Data, and Preparation of the Final Written Report.

Preliminary Procedures

A documentary analysis of information pertinent to all phases of the study was conducted before establishing definite procedures. Special emphasis was placed upon: (1) step tests as reliable and valid

measures of cardiovascular efficiency, (2) submaximal working capacity tests as reliable and valid measures of cardiovascular efficiency, (3) studies employing step tests with college men and high school boys, (4) studies employing step tests with college women, high school and junior high school girls, and (5) studies employing step tests with elementary school age children. Upon the basis of the documentary analysis and empirical judgement deemed acceptable by the thesis committee, the instruments were selected and the study was conducted.

Submaximal Working Capacity (SWC) has been defined by Kuruz, Fox and Mathews (1969) as "the inning of work in which the subject's heart rate reaches 150 BPM." Because of the limited information concerning the SWC for girls between the ages of twelve and sixteen, the present study was undertaken.

A Tentative Outline for the study was developed and presented in a Graduate Seminar conducted in the College of Health, Physical Education and Recreation at the Texas Woman's University in Denton, Texas, during the spring semester of the 1970-1971 academic school year. The Tentative Outline was revised in accordance with suggestions made by the members of the thesis committee. Upon approval, the Tentative Outline was filed in the form of a Prospectus in the office of the Dean of Graduate Studies.

Plans were formulated for the data to be collected. Permission was secured from the principal of Bedford Junior High School in Bedford, Texas, to conduct the study during the spring semester of the academic school year of 1970-1971. The cooperation of the physical education teachers at Bedford Junior High School was solicited and granted.

Criteria for the Selection of the Instruments

Prior to the selection of the instruments, criteria for the selection of each test was established. A survey of authoritative sources -- Willgoose (1961: 24), Mathews (1969: 24-28), Haskins (1971: 7-8), and Sheehan (1971: 47-54) -- indicated that the instruments should meet the following minimum criteria: validity, reliability, objectivity, and administrative feasibility. These were the criteria used in the selection of the instruments accepted for the present study.

Selection and Description of the Instruments

The modified Ohio State University Step Test was selected in accordance with the criteria as the measure of the submaximal working capacity of the one hundred subjects. A submaximal test was selected because of the ease of administration and because the required equipment could be made available to the investigator. The advantages of a submaximal test are: (1) motivation may be eliminated as a factor in physical fitness testing, (2) unfit or unconditioned subjects can be tested without possible hazards attendant upon a maximum work load, and (3) a well equiped laboratory and skilled technician are not necessary. (Mathews, 1969).

The Ohio State University Step Test was modified in several ways to adapt it from a test which was designed for college men to a test that would be appropriate for junior high school age girls. The first change that was made was in the height of the bench. An eighteen inch bench was used by the men; in the present study a bench of varying heights

was used that allowed the subjects to have an approximate right angle at both the hip and knee joints. In the original test, the cadance was increased in Phase II; in the modified test the cadance was not increased until Phase III because mechanical efficiency was lost at the faster cadance (Crowley, 1971). Instead of increasing the cadance in Phase II, the height of the bench was increased by two inches. Aside from these two modifications, the step test was administered in the same way as the step test conducted at the Ohio State University.

The Sjostrand method was selected to measure the physical working capacity (PWC) on the bicycle ergometer (Adams and others, 1961). Adams used the test extensively with school age children in California and Sweden. The test was shown to be a valid and reliable measure of the PWC at 170 heart beats per minute. The bicycle ergometer has the following advantages: (1) the work load is expressed in standard units of work thus allowing for work comparisons more easily than work on the treadmill; (2) measurements with various instruments can be taken easily because the subject's upper body remains relatively motionless (Adams, 1961). A detailed description of the test can be found in this chapter on pages 35 and 36.

Selection and Description of the Subjects

Two criteria were established for the selection of the subjects: each should be a female and each should be enrolled in Bedford Junior High School in Bedford, Texas, during the spring semester of the academic school year of 1970-1971.

A list was compiled of all the girls enrolled in Bedford Junior High School and a number was assigned to each girl. The lottery method

was used to randomly select one hundred subjects. Eleven subjects were eliminated from the study: seven withdrew from school before the study began and four were eliminated from the study for medical reasons. TABLE 1 presents this descriptive data in tabular form.

TABLE 1

DESCRIPTION OF THE SUBJECTS

VARIABLES	RANGE	MEAN
AGE IN YEARS	4 (12-16)	13.59
HEIGHT IN INCHES	16 (54-70)	62.47
WEIGHT IN POUNDS	148 (70-218)	112.76

The subjects ranged in age from twelve to sixteen with an average age of 13.59 years. The height of the subjects ranged from 54 inches to 70 inches with an average height of 62.47 inches. A wide range existed in the weight of the subjects: 70 pounds to 218 pounds. The average weight of the subjects was 112.76 pounds.

Selection of the Student Assistants

Ten student assistants were selected to aid the investigator in the administration of the step test. These girls had been selected at the beginning of the 1970-1971 school year as Office Help in the Physical Education Office. They had worked closely with the investigator on many occasions and had exhibited the ability to accept responsibility. A period of orientation was conducted for the assistants during which time they became familiar with the procedures of the test and the method of

palpitating and counting the pulse at the carotid artery. The responsibilities of the student assistants were to organize the subjects during each class period, determine the heart rate of the subjects and to record the heart rates on score cards.

Description and Results of the Pilot Study

A pilot study was conducted with the modified step test to determine the ability of the test to distinguish between good, average, and poor levels of cardiovascular efficiency. Eighteen students were purposively selected as subjects for the pilot study; they were not among the one hundred subjects used in the final investigation. The subjects were arbitrarily assigned to one of three groups: good, average and poor levels of cardiovascular efficiency on the basis of teacher observation. The equipment necessary for the administration of the pilot study was the same equipment that was used in the actual study: a tape recording of the test procedures and the cadance to be followed and a bench of varying heights. Each subject was tested on time only. The submaximal working capacity at the 150 BPM rate (SWC150) was recorded for each subject. TABLE 2 presents the results of the pilot study.

TABLE 2

RESULTS OF THE PILOT STUDY

GROUP	MEAN INNINGS COMPLETED
GOOD	7.4
AVERAGE	4.2
POOR	2.5

The scores of the pilot study were expressed in the number of innings of work that produced a heart rate of 150 BPM. The mean scores of the good group was 7.4; the mean score of the average group was 4.2; and the mean score of the poor group was 2.5 innings of work. On the basis of the pilot study, it was concluded that the modified step test differentiated between the levels of cardiovascular efficiency.

Procedures Followed in the Collection of Data

The investigator established procedures for use throughout the investigation. The pilot study was instrumental in the establishment of schedules and procedures to be used in the initial test and the retest.

Orientation

An orientation period was conducted for all the subjects. During the orientation meeting the subjects were informed of the purpose of the step test and how they were chosen as subjects. They listened to a tape recording of the exact procedures (see Appendix) and observed a student assistant taking the test. Time was allowed for any questions the subjects had about the test.

Step Test

The subjects were tested during their physical education class periods; the students who were in the band and were not enrolled in physical education were tested during their homeroom periods. Each subject was tested twice with a forty-eight hour period between the tests. Upon reaching the gymnasium, the subjects were asked to rest for at least five minutes before they were tested to establish a resting pulse rate. They were grouped by the student assistants so that three or four subjects

could be tested at one time. The limiting factors on the number who could be tested at one time were the height of the bench each subject used and the number of student assistants available during that class time.

At the start of the test, the subjects faced their bench and on the command "up" stepped up on the bench with one foot and then stood erectly on the bench with their other foot. When they stepped down, they did so with their lead foot first. They did this for the full working period of each inning. The innings were fifty seconds in duration: thirty seconds of work and twenty seconds of rest. During the rest period of twenty seconds, the pulse rate is counted for ten seconds. The palpitation method at the carotid artery was employed to count the pulse rate; the number obtained after the ten second counting period was multiplied by six to determine the pulse rate per minute. When the pulse rate reached 150 BPM, a notation was made on the score card. When the pulse rate reached 160 BPM, the test was terminated.

The eighteen innings were continuous. For the first twelve innings the cadance was twenty-four steps per minute as dictated by a metronome setting of ninety-six beats per minute. The subjects worked at the same bench for the first six innings. The bench they worked on for the first six innings allowed the subject to have an approximate ninety degree angle at both hip and knee joints when the foot was placed on the bench. After the sixth inning and before the seventh, the subjects were instructed to move in front of the next bench which was two inches higher than the original bench. After the twelfth inning and before the thirteenth, the subjects were instructed that the cadance would increase to thirty steps per minute, with a metronome setting of 120 beats per

minute for the last six innings of the test.

The SWC_{150} and the SWC_{160} were determined for each subject. The rate of 150 BPM was the terminating pulse rate for the original Ohio State University Step Test. It was believed that the heart rate of 150 BPM might be too low and that the heart rate of 160 BPM might be a more valid measure. Corroborative evidence of this belief was found in the Mottinger study (1971) which determined that maximum Vo_2 values were obtained for women at the level of 162 BPM when the data were extrapolated from sub-maximal measures. Therefore, both heart rates were recorded and treated statistically to determine which might be a more valid measure.

Bicycle Ergometer Test

The Sjostrand method for determining the physical working capacity at 170 BPM (PWC_{170}) was selected as the test to establish the validity of the modified step test because the reliability and validity of that test for junior high school age girls had been established by Adams and others (1961). Thirty subjects were randomly selected by the use of the lottery method from among the one hundred subjects taking the modified Ohio State University Step Test and were asked to take the test on the bicycle ergometer. The bicycle ergometer test was administered in order to establish the validity of the modified step test. The equipment necessary for the administration of the test included: a Monark bicycle ergometer, a tape recording of the cadance to be followed, a stop watch, and a stethoscope.

Upon reaching the gymnasium, the subjects were asked to rest for at least five minutes before being tested so that the resting pulse rate could be recorded. The seat on the bicycle ergometer was adjusted for each subject so that full extension of the leg was possible while

they were seated. For the first minute the subject pedaled with no friction to the cadance of fifty revolutions per minute; the metronome was set to 100 beats per minute. After the first minute, the friction was set on 300 kilopond meters (KPM). The pulse rate was counted by means of the stethoscope for the last fifteen seconds of each minute and that number was multiplied by four to establish the heart rate per minute. When the subject's pulse stabilized for two consecutive minutes, the subject stopped and the heart rate was plotted on graph paper (see Appendix). The subject then rested until her pulse rate returned to its resting rate. The length of time required to rest varied for each subject. After resting, the subject returned to the bicycle ergometer and pedaled for one minute without friction to the same cadance - fifty revolutions per minute. After the first minute, the friction was set on 450 KPM. The heart rate was counted for the last fifteen seconds of each minute and that number was multiplied by four to establish the heart rate per minute. When the heart rate stabilized for two consecutive minutes, the subject was stopped and the heart rate was plotted on graph paper.

If the subject's heart rate reached 170 BPM during the exercise period, the test was terminated and the score was recorded as the amount of work, expressed in KPM, that produced the heart rate of 170 BPM. In the event that the subject's pulse rate did not reach 170 BPM during the exercise, the heart rates were plotted at both 300 and 450 KPM of work and a line was extrapolated to determine the amount of work that would have been necessary to produce the heart rate of 170 BPM.

Each subject was tested only one time on the bicycle ergometer. The resting time between the two work phases varied, but both work loads were given on the same day within minutes of each other.

Organization and Treatment of the Data Collected

The collected data were organized in a table of raw scores (see Appendix) under the heading of: Subject Number, Age, Height, Weight, Score for the Step Test at 150 BPM, Score for the Retest at 150 BPM, Score for the Step Test at 160 BPM, Score for the Retest at 160 BPM, and Score for the Bicycle Ergometer Test. The Pearson Product-Moment Correlation technique was utilized to make six comparisons between: Test 150 and Retest 150; Test 160 and Retest 160; Test 150 and Test 160; Retest 150 and Retest 160; Retest 150 and the Bicycle Ergometer Test; and Retest 160 and the Bicycle Ergometer Test.

Procedures Related to the Writing of the Final Written Report

Upon the completion of the statistical treatment of the data, the investigator summarized the report, stated a conclusion to the study, and discussed the implications of the study. The final procedures included making recommendations for further studies, compiling a bibliography and developing an appendix.

Summary

The procedures followed in the development of the study were outlined in this chapter. These procedures included preliminary procedures and those procedures directly concerned with the collection and treatment of the data. Preliminary procedures involved the selection of the instruments.

The purpose of the study was to determine the validity and the

reliability of the modified Ohio State University Step Test as a measure of the submaximal working capacity of junior high school age girls. The test-retest method was employed to establish the reliability of the step test and the Sjostrand method on the bicycle ergometer was employed to establish the validity. Both tests met the minimum requirements of validity, reliability, objectivity, and administrative feasibility. Subjects for the study were one hundred randomly selected girls enrolled in Bedford Junior High School in Bedford, Texas, during the spring semester of the academic school year of 1970-1971. Each subject was given the step test twice, forty-eight hours apart. Thirty of the subjects were randomly selected to be tested one time on the bicycle ergometer.

The Pearson Product-Moment Correlation was selected to study the relationships between: the Test 150 and the Retest 150; the Test 160 and the Retest 160; the Test 150 and the Test 160; the Retest 150 and the Retest 160; the Retest 150 and the Bicycle Ergometer Test; and the Retest 160 and the Bicycle Ergometer Test.

Chapter IV is the presentation and the interpretation of the findings.

CHAPTER IV

PRESENTATION AND INTERPRETATION OF THE FINDINGS

Introduction

The purpose of the present study was to modify the Ohio State University Step Test so that it would be a reliable and valid measure of the submaximal working capacity of junior high school age girls. The subjects were one hundred girls between the ages of twelve and sixteen who were enrolled in Bedford Junior High School in Bedford, Texas, during the spring semester of the academic school year of 1970-1971. Data were collected on the modified step test and a retest. A record was kept of the work required to reach a heart rate of 150 BPM. Work was terminated when the heart rate reached 160 BPM. Thirty subjects were randomly selected to be tested one time on the bicycle ergometer to establish the Physical Working Capacity at 170 BPM for the purpose of determining validity coefficients of correlation.

Descriptive Data of the Test-Retest of the Step Test at Both 150 and 160 BPM

The researcher employed the test-retest method to establish the reliability of the modified Ohio State University Step Test at both 150 and 160 BPM. A coefficient of correlation was also made between the test at 150 and the test at 160 and between the retest at 150 and the retest at 160 BPM. Each subject was monitored by a student assistant to note the time (in innings) required to reach a heart rate of 150 BPM.

Each subject was allowed to continue until she reached 160 BPM at which time the test was terminated.

TABLE 1

DESCRIPTIVE DATA OF THE TEST-RETEST OF THE STEP TEST AT BOTH 150 AND 160 BPM (N=100)

	RANGE	MEAN	S	SE_m	R_{tt}
TEST 150	11(1-12)	3.42	2.69	0.27	.55
RETEST 150	11(1-12)	4.43	2.66	0.26	
TEST 160	12(1-13)	5.75	2.99	0.30	.55
RETEST 160	13(1-14)	6.63	3.32	0.33	
TEST 150	11(1-12)	3.42	2.69	0.27	.79
TEST 160	12(1-13)	5.75	2.99	0.30	
RETEST 150	11(1-12)	4.43	2.66	0.26	.77
RETEST 160	13(1-14)	6.63	3.32	0.33	

A study of TABLE 1 reveals that the range in scores on both test and retest at 150 BPM was eleven innings of work. The extreme scores, indicating the number of innings required to reach 150 BPM, in both cases were the first and twelfth innings of work. The range of scores for the initial test at 160 BPM was twelve innings of work; the range for the retest at 160 BPM was thirteen innings. The extreme scores for the retest at 160 BPM were the first and fourteenth innings while the extreme scores for the initial test at 160 BPM were the first and thirteenth innings of work.

The mean score for the initial test at 150 BPM was 3.42 innings

and the mean score for the retest was 4.43 innings of work. An improvement of 1.01 innings of work was made from the initial test to the retest. The mean scores for the initial test at 160 BPM was 5.75 innings and the mean score for the retest was 6.63. The improvement from the test to the retest was .88 innings of work.

The standard deviations for all of the tests were greater than the ideal one-sixth of the range which indicates a leptokurtic curve suggesting homogeneity of the group. The standard deviation for the initial test at 150 BPM was 2.69. The standard deviation for the retest at 150 BPM was 2.66. The standard deviation for the initial test at 160 BPM was 2.99 and the standard deviation for the retest at 160 BPM was 3.32. The standard error of the mean for all the tests appear to be relatively small and may be considered normal.

The coefficient of correlation as determined by the Pearson Product-Moment Technique for the test-retest at 150 BPM was .55. This is considered a poor coefficient of correlation by Barrow and McGee (1971: 38). The coefficient of correlation for the test-retest at 160 BPM was also .55. The coefficient of correlation for the initial test at 150 BPM and continuing the test until 160 BPM was reached was .79 which is considered good by Barrow and McGee. This is the relationship on the initial test of the number of innings required to reach the heart rates of 150 and 160 BPM without a rest period. This actually may be considered to be a poor correlation coefficient for a test that is continuous with two points during the work correlated. It would be assumed that a very high correlation existed. The correlation between the retest at 150 and 160 BPM was .77 which is also considered good by Barrow and McGee. The higher correlation for the continuous test that

measure the heart rates of 150 and 160 BPM should be expected since the heart rate increases in a linear fashion during exercise. The coefficients of correlation, while acceptable, are considered too low to be of pracital value.

The fact that the subjects were not allowed to practice the step test before it was administered may have been the contributing factor to the low correlation between the test and the retest at both 150 and 160 BPM. There appears to be a considerable improvement between the test and the retest which were administered forty-eight hours apart. The improvement may be the result of loss of any anxiety the subjects experienced the first time they were tested. Perhaps the increased familiarity with the test during the second measurement eliminated the anxiety factor which would account for the higher coefficients of correlation obtained.

Descriptive Data of the Retest 150, Retest 160 and the Bicycle Ergometer Test

The bicycle ergometer test (Sjostrand Method) was administered to thirty randomly selected subjects who were part of the one hundred subjects who participated in the step test. The subjects were tested one time within two weeks after retest on the step test for the purpose of establishing the validity of the modified step test. A comparison was made statistically between the retest at both 150 and 160 BPM and the bicycle ergometer test. The retests were selected for comparison because they were thought to be more reliable measures of the subjects submaximal working capacity than the initial tests because of the familiarity and less anxiety of the subjects.

TABLE 2

DESCRIPTIVE DATA OF THE RETEST 150, RETEST 160 AND THE BICYCLE ERGOMETER TEST

	RANGE	MEAN	S	SE_m	R
STEP TEST 150 BPM (in innings)	11(1-12)	3.42	2.69	.27	.019
BICYCLE ERGOMETER (in Kpm)	900(300-1200)	734.5	188.95	34.50	
STEP TEST 160 BPM (in innings)	13(1-14)	6.63	3.32	.33	.167
BICYCLE ERGOMETER (in Kpm)	900(300-1200)	734.5	188.95	34.50	

The statistical treatment of the retests of the modified step test at both 150 and 160 BPM has previously been discussed (pp 40-41). A study of TABLE 2 reveals that the range of scores on the bicycle ergometer test was 900 kilopond meters with the extreme scores being 300 and 1200 Kpm of work. The mean score was 734.50 Kpm per minute. The S of 188.95 was larger than the ideal one-sixth of the range and resulted in four standard scores: two positive and two negative. The coefficients of correlation between the step retests and the bicycle ergometer test as determined by the Pearson Product-Moment Technique were very poor. The coefficient of validity between the step test retest at 150 BPM and the bicycle ergometer test was only .019 and the coefficient of validity between the step test retest at 160 BPM and the bicycle ergometer test was .167.

The fact that the subjects were not allowed to practice on the

bicycle ergometer before they were measured might account for the low coefficients between the bicycle ergometer and the step test retests. The subjects may have experienced anxiety concerning the test or they may have been unfamiliar with the exercise required. If the bicycle ergometer test, as it was administered, was a reliable and valid measure of the physical working capacity of the junior high school age girls, the low coefficient must be interpreted to indicate that the modified step test was not a valid and reliable measure of the submaximal working capacity of junior high school age girls.

Summary

In this chapter the researcher presented the descriptive data for the modified step test and a retest with terminal heart rates of 150 and 160 BPM and the results of the bicycle ergometer test in tabular and narrative form. The range, mean, standard deviation, standard error, and the coefficients of reliability and validity were presented. The data collected were interpreted to suggest that due to low reliability and validity, the modified Ohio State University Step Test is not a valid and reliable measure of the submaximal working capacity of junior high school age girls.

Chapter V includes a summary of the investigation, the conclusion of the study, and recommendations for further studies appropriate to junior high school age girls.

CHAPTER V

SUMMARY, FINDINGS, CONCLUSION, AND RECOMMENDATIONS FOR FURTHER STUDIES

Summary of the Investigation

The purpose of the study was to modify the Ohio State University Step Test in order that a valid and reliable measure of submaximal working capacity of junior high school girls might be obtained. One hundred girls enrolled in Bedford Junior High School in Bedford, Texas, during the spring semester of the academic school year of 1970-1971 were randomly selected as subjects. The test-retest protocol on a modified step test was used to establish the reliability of the measure and thirty of the initial subjects were tested one time on the bicycle ergometer in the manner of Sjostrant to establish a coefficient of validity.

The literature indicates that Tuttle was among the first to conduct a study concerned with physical efficiency that did not necessitate expensive equipment, a great deal of time or an experienced administrator. In 1931 Tuttle developed the pulse-ratio rechnique to score his "stool climbing" test. In 1943 Brouha designed a step test to measure the general physical condition for hard work of normal, healthy men. The test was designed to classify unknown subjects into categories of excellent, good, average, or poor. Approximately four months after the initial test, Brouha and Gallagher collaborated to modify the Harvard Step Test so that it would be appropriate for girls. Both tests yielded

scores that were comparable to those obtained with more exacting techniques requiring expensive instrumentation.

In 1963 Skubic and Hodgkins conducted a study to find a test of cardiovascular fitness for girls and women which could be quickly and easily administered. They found that a three minute step test on an eighteen inch bench was strenuous enough to be considered hard work for girls and women. After their initial investigation, Skubic and Hodgkins reported two additional studies. In the first, they established standards for the step test previously discussed for college women and in the second, they established norms for high school and junior high school age girls on the same test.

Alderman undertook a study in 1965 to study the fatigue curve for total body work as a criterion for work capacity. The findings of the investigation were that the scores increased in reliability as the heart rate progressively increased from 100 to 160 BPM; the reliability at 160 BPM was more reliable than that at 180 BPM; and as the work load increased at a steady progression, the heart rate increased in a linear fashion.

In 1966 Day completed a study that evaluated the reliability of Rhyming Step Test which is a modification of the Harvard Step Test. Fifty-two male volunteers from the University of Oregon served as subjects for the study. The reliability coefficient as determined by the test-retest technique was .795. Brizendine conducted a study in 1967 to determine the possibility of a test with a shorter stepping time and a lower bench yielding results similar to the original Harvard Step Test. The subjects for the study were twenty-nine students from Andrews University. Preliminary studies resulted in the selection of an

eight inch bench and a one minute stepping time (the B-1 test). The B-1 test was reported to distinguish between the good, average and poor with an acceptable level of confidence.

In 1969 Myers conducted a study of the reliability of the Harvard Step Test as typically administered. A reliability coefficient of .84 was computed when the test was given individually by one test administrator. When the test was administered to a large group of eighth grade boys with the pulse rates counted by a partner, the reliability coefficient diminished to .65.

In 1968 McArdle, Zwiren and Magel conducted a study to determine the validity of the post-exercise heart rate as a means of estimating the heart rate during work. The average underestimation of the exercise heart rate increased from 5.7% in recovery from strenuous work to 13.5% in recovery from moderate work.

Harvey and Scott conducted a study in 1970 to investigate the reliability and validity of the Kent State University (KSU) Step Test and the relationship between digitally self-determined and stethoscopically determined pulse rates. The subjects for the study were thirty-two freshman women and fourteen women athletes from Kent State University. The study indicated that the KSU Step Test is both a valid and a reliable measure of cardiovascular efficiency and that the self-palpitation method was sufficiently accurate to score the test.

The present study is based on the research of Kuruz, Fox and Mathews conducted at the Ohio State University. Their study modified previous step tests in that the work loads were divided into eighteen innings of fifty second duration with three work loads that progressed in intensity. The work was terminated when the subject's pulse reached

150 BPM. The exercise commences so gradually that even those in the poorest of physical conditions can be tested and, at the same time, the last six innings are believed to be sufficiently strenuous to require the subjects to draw on their anaerobic reserves. A coefficient of reliability of .94 was determined by the test-retest technique and the coefficient of validity of .94 was determined by correlating the scores of thirty of the subjects on the OSU Step Test with the Balke Treadmill Test. Cotton, modified the Ohio State University Step Test for high school age boys and found that the test was appropriate for mass testing. The bicycle ergometer test used in the present study to determine the validity of the modified Ohio State University Step Test was the Sjostrand Methods used by Adams et. al. in a study conducted in 1957 to determine the physical working capacity of 196 Swedish children between the ages of ten and twelve.

The previous research has indicated that a step test may be a reliable and valid measure of cardiovascular efficiency and that the results compare satisfactorily with results obtained by more rigorous techniques. A submaximal step test possess the additional advantages of eliminating to a great extent the motivational factor in the subject as well as providing a method of testing those in poor physical conditions without unnecessary strain.

The general purpose of the study was to modify the Ohio State University Step Test to determine if it could be valid and reliable measure of the submaximal working capacity of junior high school age girls. The null hypothesis tested was that the modified Ohio State University Step Test is not a reliable and valid measure of the submaximal working capacity of junior high school age girls.

One hundred female students who were enrolled in Bedford Junior High School during the spring semester of the academic school hear of 1970-1971 were randomly selected as subjects. An orientation period was held for all subjects. A test-retest protocol was employed to establish the reliability of the step test and the Sjostrand procedure was employed to establish the validity. Each subject was given the modified step test twice, forty-eight hours apart. Thirty subjects were randomly selected to be tested one time on the bicycle ergometer. In the original Ohio State University Step Test the terminating point of the exercise was 150 BPM; on the basis of a study by Mottinger and the findings of Alderman, the terminating point for the modified test was 160 BPM. However, the 150 BPM was noted and then recorded for each subject. The heart rates were determined by the palpitation method at the carotid artery by student assistants who had been trained in the palpitation method.

The Pearson Product-Moment Correlation Technique was selected to study the relationships between: the Test 150 and the Retest 150; the Test 160 and the Retest 160; the Test 150 and the Test 160; the Retest 150 and the Retest 160; the Retest 150 and the bicycle ergometer test; and the Retest 160 and the bicycle ergometer test. The purpose of the correlations between the test and retest at both 150 and 160 BPM was to establish the reliability of the modified step test. The purpose of the correlations between the test at 150 and 160 BPM and the retests at 150 and 160 BPM was to determine which terminating pulse rate was more useful. The purpose of the correlations between the retests at both 150 and 160 BPM and the bicycle ergometer test was to determine the validity of the modified step test.

Findings of the Study

The following statements reveal the findings of the study:

1. The coefficient of correlation (r) between the initial test at 150 and the retest at 150 BPM was .55.
2. The r between the initial test at 160 BPM and the retest at 160 BPM was .55.
3. The r between the initial test at 150 BPM and the initial test at 160 BPM was .79.
4. The r between the retest at 150 BPM and the retest at 160 BPM was .77.
5. The r between the retest at 150 BPM and the bicycle ergometer test was .019.
6. The r between the retest at 160 BPM and the bicycle ergometer test was .167.

Test of Hypothesis

The null hypothesis stated in the first chapter was tested through the application of the Pearson Product-Moment technique. It was found that . . .

> The modified Ohio State University Step Test is not a reliable and valid measure of the submaximal working capacity of junior high school age girls.

The data collected for this study provided sufficient information for the investigator to accept the null hypothesis.

Conclusion of the Study

It may be concluded that the modification of the Ohio State University Step Test as it was administered in this study and for this particular sample utilized was not a valid or reliable measure of the submaximal working capacity of junior high school age girls. The results may be due to the modifications made from the original Ohio State University Step Test or to the inappropriateness of the test for junior high school age girls.

Limitations of the Study

During the conduct of this study procedures outlined by Adams, et. al. were followed. Adams used an electric bicycle ergometer that has a wheel circumference smaller than that of the Monark Bicycle. This necessitates a difference in pedal speed to equate work loads on the Monark Bicycle Ergometer. This was not done. The data collected is stated in kilopond meters, but is not exact. The effect of this error on the validity coefficient of correlation is unknown.

Recommendations for Further Studies

The following suggestions have been recommended for further investigation.

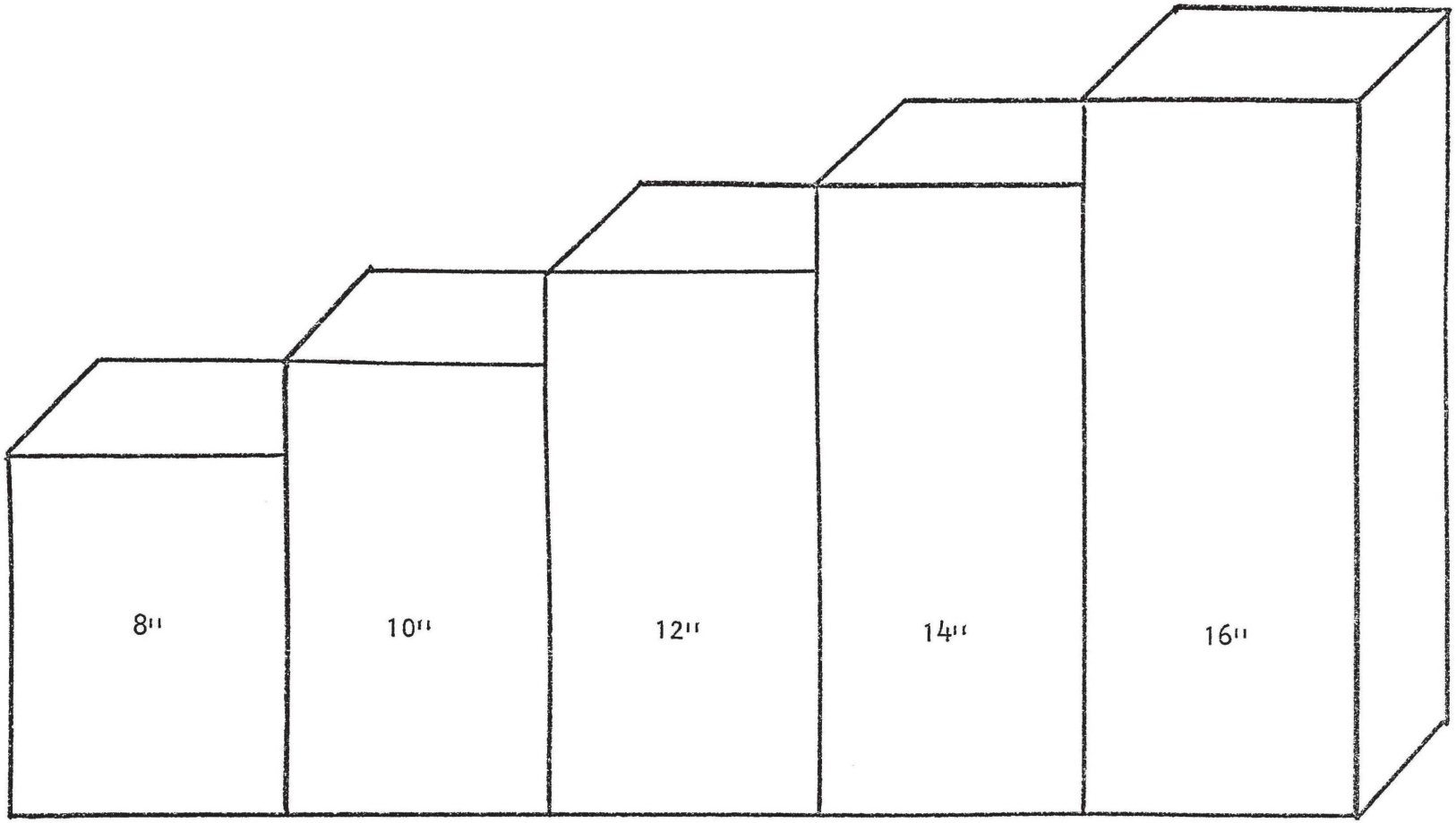
1. The conduct of a similar study with only one person determining the pulse rates.
2. The conduct of a similar study allowing the subjects to practice the exercise until they are thoroughly familiar with the procedures so that learning can not

be considered to play any part in the test results.

3. The modification of the Ohio State University Step Test for junior high school age girls with only the height of the bench being modified to allow the subject an approximate 90 degree angle at the hip and knee joints when stepping up onto the bench.

APPENDIX

ILLUSTRATION OF THE STEP



TAPE RECORDING OF SUBJECT ORIENTATION

You have been chosen to take a cardiovascular step test. This test will measure the response of your heart to exercise. There are no standards for this test for girls in your age group, so there is no way you can fail. Your score will simply be the amount of work required to reach a heart rate of 160 beats per minute. The work you will do is stepping up and down on a step to the beat of a metronome. Be sure that in your stepping that you straighten your knees completely on top of the step and stand erectly - no stopping. You will step for thirty seconds and then rest for twenty seconds. The fifty seconds of work and rest comprise one inning. During the twenty seconds of rest, sit down on your step and the student assistant will take your pulse at the carotid artery (show the students the location) for ten seconds. If you wish, you may take your own pulse rate at the same time. You will continue to work until your pulse reaches twenty-seven for the ten second counting period which would be 160 beats per minute. When your heart reaches 160 BPM your test will be finished.

A student assistant will now demonstrate two innings of work. (demonstration) If you have any questions, please ask them at this time. (pause) Now, you are ready to take your step test. Stand and face your step and begin.

RAW DATA

SUBJECT NUMBER	AGE (in years)	HEIGHT (in inches)	WEIGHT (in pounds)	STEP TEST-150 (in innings)		STEP TEST-160 (in innings)		BICYCLE ERGOMETER TEST (in KPM)
				TEST	RETEST	TEST	RETEST	
1	13	65.5	107	4	5	5	8	-
2	12	64.5	124	1	1	5	5	550
3	15	66	126	3	4	4	10	-
4	12	61	89	3	3	5	8	-
5	13	58	97	2	3	3	12	-
6	13	65.5	107	5	9	8	9	-
7	12	64.5	115	1	6	5	11	-
8	12	62	115	3	5	4	6	-
9	12	61.5	103	4	2	5	5	-
10	13	63	158	1	1	3	3	1200
11	13	64	102	3	5	4	6	-
12	13	64.5	123	1	5	1	9	550
13	13	62	117	1	2	5	5	-
14	12	59	95	3	8	5	12	600
15	13	62	108	4	1	6	4	-
16	13	61	82	1	5	5	12	-
17	13	63.5	112	6	4	8	4	-
18	13	59.5	80	1	1	5	6	-
19	12	68	147	1	1	1	1	-
20	12	56.5	127	9	12	8	8	850
21	13	64.5	119	5	5	8	6	-
22	12	64.5	114	2	5	5	7	950
23	13	62	98	5	5	11	12	-

SUBJECT NUMBER	AGE	HEIGHT	WEIGHT	STEP TEST-150		STEP TEST-160		BICYCLE ERGOME TEST
				TEST	RETEST	TEST	RETEST	
24	13	63.5	83	6	8	9	12	750
25	13	58	79	2	2	2	2	300
26	13	60	85	5	5	8	5	-
27	12	59.5	108	1	4	11	5	-
28	13	63.5	106	3	8	6	13	800
29	14	68	194	1	5	5	11	-
30	14	61	120	8	5	9	8	-
31	13	60	96	12	12	12	13	700
32	13	63	100	3	4	5	6	775
33	14	59.5	99	1	8	5	9	-
34	13	61	92	2	2	5	4	-
35	13	55	107	11	8	12	9	-
36	13	62	97	4	10	8	10	550
37	12	60	79	11	12	12	12	-
38	13	55	75	11	12	13	14	-
39	12	62	104	5	3	8	5	-
40	13	61	130	1	1	2	1	-
41	13	60.5	108	1	1	3	3	-
42	12	54	70	3	5	5	8	-
43	12	54	91	2	1	3	1	500
44	13	64.5	135	1	1	3	1	-
45	13	63	104	4	3	5	8	-
46	14	65	136	2	2	5	3	-
47	13	60	95	1	2	1	3	-
48	14	61	109	1	2	1	5	725
49	14	62	95	12	8	12	11	825
50	13	62	96	3	6	5	10	-

SUBJECT NUMBER	AGE	HEIGHT	WEIGHT	STEP TEST-150		STEP TEST-160		BICYCLE ERGOME TEST
				TEST	RETEST	TEST	RETEST	
51	14	60	103	9	8	9	12	-
52	14	65	96	1	9	3	13	900
53	14	70	218	1	1	2	1	-
54	13	61.5	122	1	4	2	6	-
55	14	63.5	111	3	5	5	8	760
56	15	66.5	196	2	2	3	2	-
57	14	61.5	107	8	6	9	9	800
58	14	61	89	5	5	8	8	-
59	12	62.5	100	3	1	5	2	-
60	13	61	95	6	5	6	8	600
61	15	65	128	3	3	5	5	-
62	14	64.5	129	2	1	3	5	-
63	15	63.5	97	5	5	9	9	-
64	14	65.5	106	3	4	5	4	850
65	14	63	104	3	6	8	9	-
66	15	55.5	112	1	4	5	4	700
67	13	67.5	126	5	2	5	3	-
68	14	61	133	5	2	5	5	850
69	15	64.5	132	1	2	3	5	-
70	15	61.5	109	1	4	3	5	-
71	14	67.5	137	2	3	2	3	500
72	14	60	104	3	5	12	6	700
73	14	65	120	4	1	5	3	-
74	15	65.5	136	5	5	8	8	700
75	14	62.5	110	3	6	5	9	-
76	16	64.5	128	6	8	8	10	-
77	15	64.5	123	3	2	5	5	1050

SUBJECT NUMBER	AGE	HEIGHT	WEIGHT	STEP TEST-150		STEP TEST-160		BICYCLE ERGOME TEST
				TEST	RETEST	TEST	RETEST	
78	15	63	115	2	3	5	5	-
79	15	60	100	9	5	12	8	1075
80	15	63	106	1	1	3	3	-
81	15	65	100	1	1	3	4	-
82	15	59.5	135	6	5	6	8	800
83	15	64.5	118	5	3	9	5	-
84	15	63	124	2	3	5	3	-
85	14	66	116	1	1	2	3	-
86	15	63.5	126	3	4	5	5	-
87	15	66	125	7	5	8	5	-
88	15	64	115	1	5	5	9	-
89	14	65	115	2	3	3	5	-
90	14	65	123	3	6	5	10	-
91	13	62.5	107	8	12	9	13	-
92	12	63	162	1	2	5	3	550
93	15	63	110	1	2	4	4	-
94	15	61	129	3	5	5	5	-
95	14	61.5	108	6	6	8	9	-
96	15	66	120	1	1	5	2	-
97	15	62	103	5	6	8	11	700
98	13	61	82	1	8	1	9	-
99	13	64	128	5	5	8	8	-
100	15	63	109	1	8	11	8	575

STEP TEST SCORE CARD

NAME: _____ NUMBER _____

BENCH: _____ AGE: _____ HEIGHT: _____ WEIGHT: _____

PHASE I PHASE II

	1	2	3	4	5	6	7	8
TEST								
RETEST								

PHASE III

	10	11	12	13	14	15	16	17	18
TEST									
RETEST									

Resting Pulse _____ Score _____

BICYCLE ERGOMETER SCORE CARD

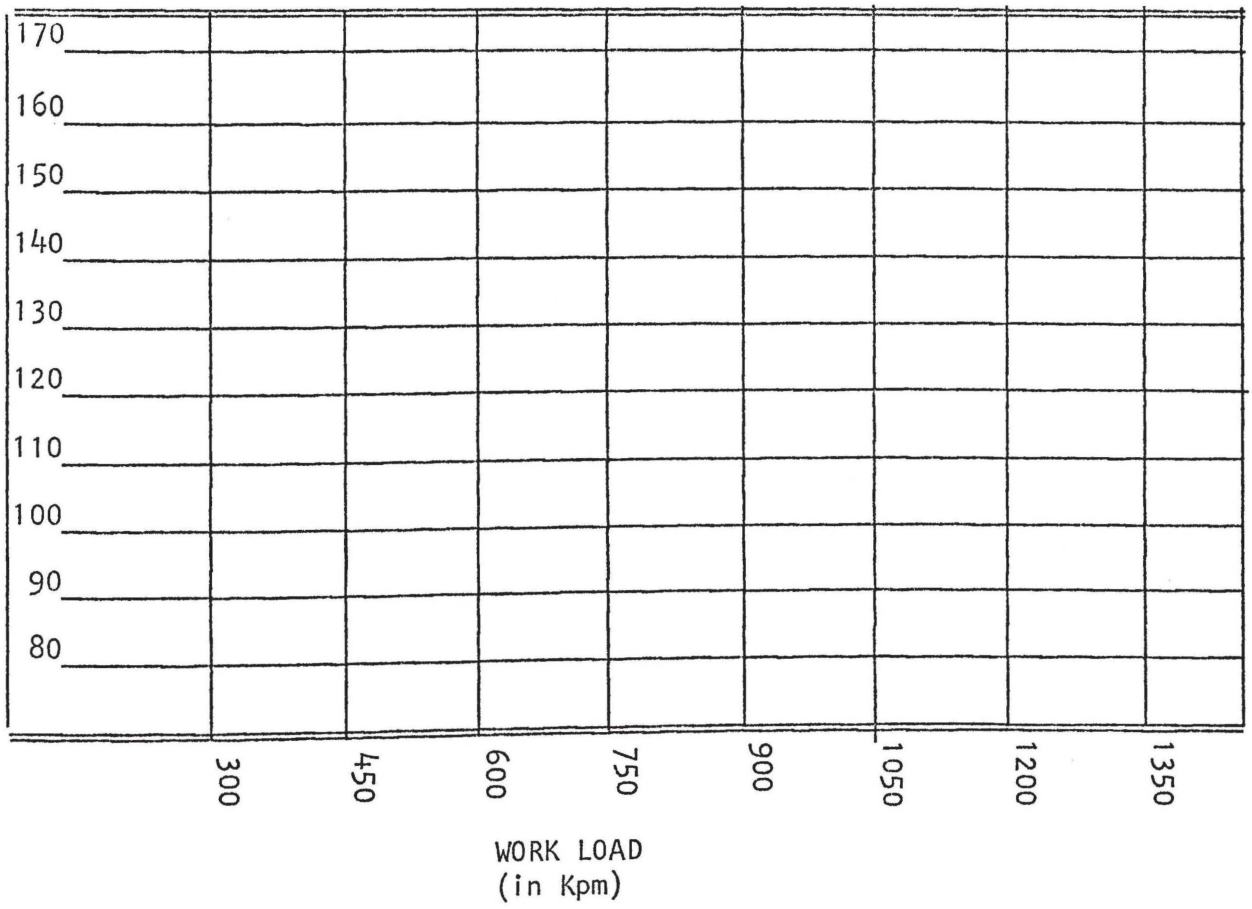
NAME_____ SUBJECT NUMBER _____

SCORE_____ RESTING PULSE RATE _____

PULSE RATES

MINUTES	1	2	3	4	5	6
300 Kpm						
450 Kpm						
600 Kpm						

HEART RATE



BIBLIOGRAPHY

BIBLIOGRAPHY

BOOKS:

Barrow, Harold M. and McGee, Rosemary. A Practical Approach to Measurement in Physical Education. 2nd Ed. Philadelphia: W. B. Saunders Company, 1971.

Haskins, Mary Jane. Evaluation in Physical Education. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1971.

Mathews, Donald K. Measurement in Physical Education. Philadelphia: W. B. Saunders Company, 1969.

Sheehan, Thomas J. An Introduction to the Evaluation of Measurement Data in Physical Education.

Willgoose, Carl E. Evaluation in Health Education and Physical Education. New York: McGraw-Hill Book Company, Inc., 1961.

PERIODICALS:

Adams, Forest H., Bengtsson, Ellas, Berven, Hans, and Wegelius, Carl. "The Physical Working Capacity of School Age Children." Pediatrics, XXVIII (1961), 244-251.

Alderman, Richard B. "Reliability of Individual Differences in the 180 Heart Rate Response Test in Bicycle Ergometer Work." Research Quarterly, XXXVII (1966) 429-434.

Brouha, Lucien. "The Step Test: A Simple Method of Measuring Physical Fitness for Muscular Work in Young Men." Research Quarterly. XIV (1943) 31-35.

Brouha, Lucien and Gallagher, J. Roswell. "A Functional Fitness Test for High School Girls." Journal of Health, Physical Education and Recreation. XIV (1943) 517 and 550.

Day, James A. P. "A Statistical Investigation of the Rhyming Step Test." Research Quarterly, XXXVIII (1966) 539-542.

Cotton, Doyce J. "A Modified Step Test for Group Cardiovascular Testing." Research Quarterly, XXXXII (1971) 91-95.

Harvey, Virginia P. and Scott, Gwendolyn D. "The Validity of a One Minute Step Test for Women." Journal of Sports Medicine, (1970) 185-190.

Kuruz, Robert L., Fox, Edward L., and Mathews, Donald K. "Construction of a Submaximal Step Test." Research Quarterly, XXXX (1969) 115-120.

McArdle, William D., Zwiren, Linda, and Magel, John R. "Validity of the Postexercise Heart Rate as a Means of Estimating Heart Rate During Work of Varying Intensities." Research Quarterly, XXXX (1969) 525-528.

Myers, Carlton R. "A Study of the Reliability of the Harvard Step Test." Research Quarterly, XXXX (1969) 423.

Skubic, Vera and Hodgkins, Jean. "Cardiovascular Efficiency Test for Girls and Women." Research Quarterly, XXXIV (1962) 191-197.

Skubic, Vera and Hodgkins, Jean. "Cardiovascular Efficiency Test Scores for College Women in the United States." Research Quarterly, XXXIV (1963) 454-461.

Skubic, Vera and Hodgkins, Jean. "Cardiovascular Efficiency Test Scores for Junior and Senior High School Girls in the United States." Research Quarterly, XXXV (1964) 184-190.

Tuttle, W. W. "The Use of the Pulse-Ratio Test for Rating Physical Efficiency." Research Quarterly, II (1931) 5-17.

THESIS:

Brezendine, S. F. "A Comparitive Study of the Harvard Step of Varied Times and Heights." Master of Arts Thesis. Michigan State University, 1967.