6-1-1967

A Comparison of the Physical Fitness Levels of Athletes and Non-Athletes Over a Four Year Period at the University of North Dakota

Del Gab

A COMPARISON OF THE PHYSICAL FITNESS LEVELS OF ATHLETES AND NON-ATHLETES OVER A FOUR YEAR PERICD AT THE UNIVERSITY OF NORTH DAKOTA

by

Del Gab

B.S. in Business Education

Northern State College

1958

A Thesis

Submitted to the Faculty

of the

Graduate School

of the

University of North Dakota

in partial fulfillment of the requirements

for the Degree of

Master of Science

Grand Forks, North Dakota

June
1967

ACKNOWLEDGEMENTS

The author wishes to express his grateful appreciation to all persons whose time, interest, and effort made this study possible. Special recognition is extended to those University of North Dakota students who served as subjects for this investigation, for their time, effort, and cooperation.

The author is particularly indebted to Mr. Walter C. Koenig, director of this thesis, for his guidance, understanding, and helpful criticism which were essential in the successful completion of this study.

Appreciation is also extended to Dr. John Quaday for the helpful instruction, suggestions, and criticism.

Finally, for her assistance, encouragement, patience, and inspiration, the writer is deeply grateful to his wife, Joyce Rosiland.

This thesis, submitted by Del D. Gab in partial fulfillment of the requirements for the Degree of Master of Science at the University of North Dakota, is hereby approved by the committee under whom the work has been done.

W. C. Koenig
Chairman

John L. Quaday

John L. Rowe

Christopher J. Hamre
Dean of the Graduate School

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

The purpose of this study was to compare the physical fitness levels of athletes and non-athletes at the University of North Dakota.

This study was directly concerned with physical fitness as measured by the American Association for Health, Physical Education, and Recreation Modified Youth Fitness Test. The students involved were selected freshmen who had enrolled at the University of North Dakota in the fall of 1963, but did not participate in intercollegiate athletics in their four years of college and athletes who participated for three years in intercollegiate football, basketball, wrestling, cross country, and/or hockey at the University of North Dakota.

The test was administered in the fall of 1963 by the physical education staff at the University of North Dakota and both groups were retested by the writer during the first semester of 1966-67.

Comparisons were made between the mean differences within each group as indicated by the scores on the initial and final tests. The null hypothesis was assumed with respect to the differences within the groups on the initial and final tests. This hypothesis was tested with the "t" technique for the difference between means derived from correlated scores from small samples. Comparisons were also made between the experimental group and the control group by testing the significance of the difference between the mean differences found

within groups. The between group comparison used the "t" technique for uncorrelated data from small samples.

Some of the conclusions indicated by this study were:

1. The experimental group (athletes) showed significant improvement in all of the selected measures of physical fitness except the 50-yard dash and sit-ups at the criterion .01 level.

2. The control group (non-athletes) improved significantly in one selected measure of physical fitness, the shuttle run, at the criterion .01 level.

3. The control group showed a significant decrease in the 600-yard run-walk and sit-ups at the criterion .01 level.

4. In comparing the experimental and control groups, the writer found significant differences at the criterion .01 level in the following test items: sit-ups, shot put, pull-ups, standing broad jump, and the 600-yard run-walk.

CHAPTER I

THE PROBLEM AND ITS SCOPE

The Problem

The purpose of this study was to determine the changes elicited by a four-year intercollegiate athletic program with respect to the fitness values of the participants as compared to the changes in fitness values of a group who did not participate in any physical education program after the first two years in college.

The specific problems of this study were as follows:

1. To find the status of fitness of the control (non-athletes) group and the experimental (athletes) group in October of 1963.
2. To try to determine what effect, if any, a four year athletic program had upon the fitness levels of the participants.
3. To try to determine what effect, if any, a non-required program of physical activity had on the control group.

Need for the Study

Athletics play a very important part in the lives of many youngsters. Participation in sports is one of the foremost things in the minds of some children today. It seems that participation

in sports that are offered in the interscholastic athletic program should make the individuals into better physical specimens.

Through the administration of physical fitness tests, it is usually possible to obtain an evaluation of the effectiveness of the athletic program as concerns certain aspects of physical fitness. However, what can be done for those students who do not have the ability to compete in the intercollegiate athletic program? It is hoped that the results obtained from this study would justify an expansion of the intramural and physical education service programs at the University of North Dakota.

Delimitations

The delimitations of the study were as follows:

1. The experiment was limited to selected male freshman students who enrolled at the University of North Dakota in September of 1963.
2. The experimental group consisted of those male students who had participated in intercollegiate football, basketball, wrestling, cross country, and/or hockey for three years at the University of North Dakota.
3. The control group was limited to those male students who had enrolled at the University of North Dakota in the fall of 1963, but had never participated in any type of intercollegiate athletics.

4. The data obtained in this experiment were limited to the modified AAHPER Youth Fitness test administered in the fall of 1963, and the re-test which was given at the end of the four year period or in the first semester of 1966-67.
5. The investigator had no control over the non-athlete as far as participating in intramurals or working out on his own.

Definitions

Physical fitness is one phase of total fitness. The components of physical fitness are resistance to disease, muscular strength and endurance, cardio-vascular respiratory endurance, muscular power, flexibility, speed agility, coordination, balance and accuracy.[1]

The American Association of Health, Physical Education, and Recreation Youth Fitness Test as modified included sit-ups, pull-ups, shuttle run, standing broad jump, 50-yard dash, 600-yard run-walk, and the shot put (substituted in place of the softball throw).

The Control Group consisted of randomly selected male freshmen who had enrolled at the University of North Dakota in the fall of 1963, but did not participate in intercollegiate athletics in their four years of college. This group is referred to as the non-athlete group in this study.

The Experimental Group were athletes who participated for three years in intercollegiate football, basketball, wrestling, cross

---

[1]Thomas Kirk Cureton, Physical Fitness Appraisal and Guidance (St. Louis: The C. V. Mosby Company, 1947), p. 18.

country, and/or hockey at the University of North Dakota. This group included only male students who enrolled as freshmen in the fall of 1963 and were athletes.

Review of Related Literature and Research

A number of studies have been undertaken in the area of physical fitness and its relationship to the physical education activities program. However, very few studies have been done in the field of physical education covering a period of more than one year.

Helen Starr[2], in an article in fitness testing quoted Paul Hunsicker, chairman of the AAHPER Fitness Council, as follows:

> The physical performances tested in the youth fitness test include running, jumping, throwing, strength, agility, and endurance. These activities should be part of physical education programs and, within limits, an improvement in test scores should accompany continuous participation in physical education. If pupils are enrolled in physical education classes and fail to improve throughout the school year in all probability the program was not sufficiently vigorous.

The NCAA Fitness Committee[3] recommends urging every member institution and conference to increase, insofar as possible, the number of sports being conducted in intercollegiate and intramural activity and to increase the number of teams competing in all sports, such as junior varsity, freshman, and lightweight teams. If additional events could be added, thus giving impetus to participation of additional

---

[2]Helen M. Starr, "How to Fit in Fitness Testing," Journal of Health, Physical Education and Recreation, Vol. 30 (March, 1958), p. 19.

[3]Tom Hamilton, "Athletic Director Looks at Fitness," Journal of Health, Physical Education and Recreation, Vol. 28 (September, 1957), pp. 14-15.

competitors in such sports as soccer, handball, squash, volleyball, rugby, water polo, and field hockey, a better program would result.

David D. Rains[4] conducted a study to establish proof that athletes' growth was affected by participating in interscholastic sports. Rains attempted to determine if athletes were affected in their growth by participating in football, baseball, basketball, and track. The results were secured by testing secondary school boys by use of the Grid technique. Rains concluded that the quality of growth of athletes is inferior to non-athletes by the Grid standard during competition. After completion of competition, Rains found that the athletes exceeded the normal growth patterns, and attained a growth level with their peers.

Norman[5] used Cozen's short battery test of all-around ability and administered it to 11th and 12th grade participants in football, basketball, baseball, and track and field from four high schools. Tests were conducted at the end of the playing season. Football players possessed significantly less all-around athletic ability than those in the other three activities, possibly because the test stressed running. The other differences between groups were not significant.

---

[4]David D. Rains, "Growth of Athletes and Non-Athletes in Selected Secondary Schools as Assessed by Grid Technique," (unpublished Ph.D. dissertation, Indiana University, 1951).

[5]Gerald F. Norman, "Comparison of Athletic Achievement in Selected Activities Among Secondary School Athletes," Completed Research in Health, Physical Education and Recreation, Vol. 1 (1959). p. 28.

The purpose of De La Barre's[6] study was to determine the effect the five period physical education week had on fitness as compared to the two period physical education week. The experimental period was one semester in length.

De La Barre concluded that the five period physical education week did produce significant changes in pulse-ratio and in selected measures of strength except the right hand grip. The two period physical education week did not produce a significant change in pulse-ratio and produced significant changes only in the back strength test and in push-ups. The five period physical education week group showed greater increases than the two period physical education week group in each measure.

Kuhst Wieneke[7] completed a study comparing the physical development of freshman athletes and non-athletes during a one year period. All members of the freshman class were given a physical examination during the fall. These freshmen were classified into two groups: athletes (those out for varsity sports) and non-athletes (those in required physical education). The two groups were tested for strength of right and left grip, back and leg strength and vital (lung) capacity. The groups were equated by paired observation of the initial test. The test was again given at the close of the school year.

---

[6]Craig H. De La Barre, "A Study of the Fitness Values Derived from a Five Day Per Week Activity Compared with Like Values from a Two Day Per Week Activity," (unpublished Master's thesis, University of North Dakota, July, 1962).

[7]Kuhst Wieneke, "A Comparison of Certain Physical Developments of Freshman Athletes and Non-Athletes," Research Quarterly, Vol. 3 (May, 1932), pp. 223-234.

Wieneke concluded:

1. With respect to right and left grip and back and leg strength, the athletes exceeded non-athletes at the .02 level of confidence.

2. Test results showed no significant difference in lung capacity.

3. In general, participation in athletics seems to be related to certain physical developments not found in non-athletes.

4. Regular attendance in required physical education classes is not related to as great a change in certain physical developments as participation in athletics.

Vern B. Hoffman[8] tried to establish proof that participation in interscholastic athletics would tend to provide our nation with more physically fit individuals. Hoffman's study was to determine the relative values of participation in interscholastic athletics and physical education programs for the development of strength, as indicated by Rogers' Physical Capacity tests.

Hoffman secured his data by testing and retesting two groups of high school boys, one participating in interscholastic athletics and the other group participating in the physical education program.

Hoffman arrived at the following conclusions:

1. The physical education and athletic programs contribute to the strength development of the individual.

---

[8]Vern B. Hoffman, "Strength Comparison of Athlete and Non-Athlete as Measured by Roger's Physical Capacity Test" (unpublished Master's thesis, University of Michigan, 1936).

2. Football in this study made the greatest contribution to strength development.

3. Football seems to be more conclusive in the development of arm and leg strength.

4. Basketball contributed to arm and leg strength.

5. Track contributed more to back strength than either of the other two sports.

6. On the average, the athlete showed a higher score both at the beginning and the end of the season than the non-athlete.

7. The winter program of physical education, which included gymnastics, swimming, wrestling, and basketball seemed to show a greater contribution toward gain in back strength than any one sport.

8. The athletes showed a gain of approximately twice that of non-athletes in all items tested except in grip strength.

9. Strength of arms, legs and back were the ones affected most by interscholastic athletics.

Michael L. Gaddie[9] conducted a study in the comparison of physical fitness levels of athletes and non-athletes at the University of North Dakota by use of the Harvard Step Test.

Gaddie concluded the following:

1. Fourteen athletes or twenty-eight per cent were considered in the excellent condition range as compared to one non-athlete.

2. Twenty-seven athletes finished the step test in good physical condition. Seven of these athletes missed the excellent group by one point with a mean score of eighty-nine.

---

[9]Michael L. Gaddie, "A Comparison of Athletes and Non-Athletes at the University of North Dakota as Measured by the Harvard Step Test" (unpublished Master's thesis, University of North Dakota, 1960).

3. Five non-athletes finished the step test in good physical condition and thirteen non-athletes were classified in high-average physical condition. These were the eighteen students participating in the intramural program.

4. Twenty-seven non-athletes, or fifty-four per cent, were found in the poor physical condition range with a mean score below fifty-five. The majority of this group failed to complete the step test.

5. The mean score for the athlete group was 86.96 compared to 54.80 for the non-athlete group. As a group, the athletes were lacking a mere three points of being considered in excellent physical condition. With a mean score of 54.80, the non-athlete group was considered in poor physical condition.

6. Even excluding those non-athletes failing to complete the step test, the athletes still had a much higher mean (86.96) score than the remaining non-athletes who had completed the test with a mean score of 77.18.

7. The entire group of athletes completed the five minute step test as compared to seventeen, or thirty-four per cent, of the non-athlete group. Of the seventeen non-athletes completing the test, fourteen participated in the intramural program.

8. The eighteen non-athletes participating in the intramural program had a mean score of 62.78 compared to the mean score of 50.31 for the thirty-two non-athletes with little or no physical activity. The group with little or no physical activity was in very poor physical condition as measured by the Harvard Step Test.

Utilizing the Physical Fitness Index (PFI) test scores, Clarke[10] reported that male students entering the University of Oregon with

---

[10]H. Harrison Clark, "Physical Fitness of University of Oregon Male Freshmen," Physical Fitness News Letter, University of Oregon, (March, 1955).

four years of high school physical education had a higher average than those who had two years or less. In another study by the same investigator,[11] the results of the Oregon Pilot Physical Fitness Project were presented. In this project, 100 boys and 100 girls in each of the 11 high schools that participated in this pilot project were tested with the Physical Fitness Index (PFI) test before and after a three months physical fitness program. The median PFI for all boys increased ten points, from 98 to 108; the median PFI increase for all girls was thirteen points, from 93 to 106.

Charles G. Campbell, Jr.[12] conducted an experimental study to determine the physical fitness status of the male students enrolled in freshman (110) and sophomore (220) physical education classes at North Carolina College at Durham, North Carolina, during the school year 1961-62. Campbell used the JCR test as the instrument of measurement. The test measured leg strength and power (vertical jump), arm and shoulder strength (dips), and speed, agility, and endurance (shuttle run). Campbell substituted dips for chins in the JCR test for his study. The test was given to the students enrolled in 110 physical education during the first semester and the 220 physical

---

[11]H. Harrison Clark, "Oregon Pilot Physical Fitness Project," The Physical Educator, (May, 1957), p. 55.

[12]Charles G. Campbell, Jr., "A Study to Determine the Physical Fitness Status of the Male Students Enrolled in 110 and 220 Physical Education Classes at North Carolina College, Durham, at Durham, North Carolina, During the School Year 1961-62," (unpublished Master's thesis, North Carolina College, Durham, North C rolina, 1962).

education students were tested near the end of the second semester. The 110 physical education class is the initial course in the required physical education program. The 220 physical education class is the final course in the required physical education program. An analysis of the results indicated that the male students enrolled in 220 physical education classes were better in all areas measured by the instrument employed by the investigator. Additionally, there was a significant difference at the .01 level of confidence for all areas measured except arm and shoulder strength. Indications were that the physical fitness status of the students enrolled in 220 physical education was much better than the physical fitness status of the male students enrolled in 110 physical education at North Carolina College at Durham, Durham, North Carolina, during the school year 1961-62.

Orlo A. Sundre[13] conducted an experimental study comparing two different physical education programs at the University of North Dakota. Sundre used four sections of regular physical education classes for the study. The participants themselves were not informed that they were being subjected to a research study program. Two sections took the regular introductory physical education program at the University of North Dakota and were referred to as the control group. The other two sections took the revised introductory physical

---

[13]Orlo A. Sundre, "A Comparative Study of Two Physical Education Programs for Male Students at the University of North Dakota" (unpublished Master's thesis, University of North Dakota, 1960).

education program which was designed for the study and stressed physical conditioning. These participants made up the experimental group. The revised program contained the same amount of time devoted to actual practice of the sport skills, but more time was spent on conditioning exercises. Comparisons were made by using the Physical Capacity test. The test was given to all participants at the beginning of the course and again at the completion of the course. The results after comparison of the two programs showed the revised introductory physical education program to be successful. The area of physical fitness, as measured by the Physical Capacity test, was better developed through the revised program.

In Virginia, Minnesota, Noel W. Olson[14] compared the effects of seasonal participation in selected interscholastic sports on motor fitness. He used boys who participated in the interscholastic sports of basketball, football, hockey, and swimming; and also a group of regular physical education class participants. The participants were tested at the beginning of their respective sport's season and again on the day after their season ended. The tests used were the side-step test of agility, the squat stand test of balance, the treadmill test of endurance, the vertical jump test of leg power, the fifty-yard dash for speed, and the pull-up test of arm strength. Olson was attempting to learn whether or not interscholastic athletic participation would develop greater motor fitness improvement.

---

[14]Noel W. Olson, "A Comparison of the Effects of Seasonal Participation in Selected Interscholastic Sports on Motor Fitness" (unpublished Master's thesis, University of North Dakota, 1961).

Olson concluded:

1. There appears to be more opportunity for motor fitness improvement through interscholastic sports participation than by participation in only the regular physical education program of Virginia Senior High School.

2. The basketball group had significant improvement in endurance and speed. In the areas of agility and power this group had such high initial mean scores that there was probably not much chance for significant increases.

3. The football group had significant improvement in the areas of agility, endurance, and power. The overall motor fitness improvement of this group was better than any other group; however, their initial mean scores on all tests were generally lower than the other athletic groups.

4. The hockey group had a significant mean improvement in endurance and evidenced the highest initial mean score in strength.

5. The swimming group had a significance of difference beyond the .02 level in agility. This group had high initial mean scores on all tests and relatively good improvement between the two tests, but the inconsistency of their scores affected the estimate of sampling error of the mean difference to the point where the increases could not be considered significant except in agility.

6. The non-athlete group made significant improvement in agility, endurance and speed, but this group did not achieve the highest mean increase in any area of motor fitness. Since their mean scores were relatively poor on the pre-season tests, it appeared that their prospects for improvement were better than for the groups whose initial mean performances were higher.

7. The experimental groups had the poorest results in the area of arm strength improvement which supports the theory that American athletics do little to improve arm strength.

8. The experimental groups appeared to improve most in the areas of agility and endurance. It should be noted that although the improvement in vertical jump was not highly significant, the mean scores were all quite high. This could be interpreted as evidence that previous athletic participation contributed to the unusual proficiency in leg power.

Floyd Boschee,[15] at the University of North Dakota, conducted a study on twenty members of the junior varsity and varsity football teams at Rugby High School. These boys were tested as to their physical fitness levels in accordance with the AAHPER Youth Fitness test. The three testing periods occurred before the initial practice, at the assumed peak during the season, and one month after the completion of the regular season. Comparisons were made in regard to the significance of difference between the first and second tests and the retention or loss between the second and third tests. The selected group made significant gains between the first and second tests in sit-ups, fifty-yard dash, and standing broad jump. The third test showed no significant loss. No significant gain was made between the first and second tests in the 600-yard run-walk, softball throw, pull-ups, and shuttle run. The losses were not significant in the third test except in the softball throw. Boschee concluded that football does significantly improve the physical fitness levels in sit-ups, fifty-yard dash, and the standing broad jump.

---

[15]Floyd Boschee, "A Comparison of the Physical Fitness Levels of Selected Participants in Interscholastic Football Before the Season, and One Month Later" (unpublished Master's thesis, University of North Dakota, 1961).

T. Erwin Blesh and Alfred E. Scholz[16] conducted a survey of the achievement of freshman students at Yale University on six tests of physical fitness which are administered as part of the regular program of physical education. Ten-thousand students took part in the ten-year survey. The physical fitness test included pull-ups, push-ups, fence vault, sit-ups, standing broad jump, and vertical jump. The results of the initial test in the ten-year survey had ranges from fifty-one per cent passing to twenty-nine per cent passing. After the initial test strenuous physical fitness classes were conducted for a period of twelve to fourteen weeks. At the conclusion of this period a retest was given. An average of eighty per cent of the students passed the retest. It was concluded that an individual's strength, agility, and co-ordination can be improved in a rather short period of time when concentrated effort is placed upon a particular factor.

Richard Vinger[17] conducted an experimental study in which he used 45 subjects from Rugby High School. In this study a control group and an experimental group were employed. The control group consisted of fifteen subjects which were compared with the experimental group of thirty subjects.

---

[16]T. Erwin Blesh and Alfred E. Scholz, "Ten Year Survey of Physical Fitness Tests at Yale University," Research Quarterly, Vol. 28 (December, 1957), pp. 321-326.

[17]Richard M. Vinger, "A Comparison of Physical Fitness Increases as the Result of a Selected Physical Education Program" (unpublished Master's thesis, University of North Dakota, 1964).

The specific problems of this study were as follows:

1. To find the status of fitness of the control group and the experimental group at the beginning of the school year.

2. To determine the changes in physical fitness as the result of participating in the required physical education program.

3. To determine the changes made in fitness values of a control group who did not participate in physical education during the school year.

4. To try to determine what effect, if any, growth and maturation had upon fitness values during the experimental period.

The participants were tested as to their physical fitness levels in accordance with the AAHPER Youth Fitness test. The initial test was given to each group at the same time after the first week of the school year. The retests were given to each group during the last week of the school term. Comparisons were made between the mean differences within each group as indicated by the initial and final tests. Comparisons were also made between the experimental group and the control group by testing the significance of the difference between the mean differences found within the groups.

Vinger concluded:

1. The required physical education curriculum in which the experimental group was engaged did produce significant changes in all of the selected measures of physical fitness except the shuttle run at the criterion .01 level.

2. The control group who did not participate in any phase of the physical education program made no significant changes in any of the selected measures of physical fitness. The similarity between the

means of the initial test and the retest for the control group seems to indicate that the subjects, once they attain a level of physical fitness, lose very little of that level by not participating in the physical education program. However, they do not gain very much either.

3. The control group did not change significantly in any of the measures of physical fitness levels while the experimental group improved significantly in nearly all areas of physical fitness. This seems to indicate that the test-retest method of evaluating the effectiveness of a physical education program in meeting the objective of physical fitness is a satisfactory device. The data collected in this study for the control group indicates that growth and maturation have little effect on the physical fitness development of an individual. The physical educator who uses this method of evaluation could feel assured that any significant changes in physical fitness levels from the initial test to the retest period are due to the effectiveness of the program in attaining that objective and not to the growth and maturation of the individual.

4. The between group comparison indicates a significant difference in pull-ups and the softball throw between the groups in terms of changes occurring during the experimental period. The changes in the other measures of physical fitness between the two groups were not significant at the criterion .01 level.

The purpose of Moser's[18] study was to determine the effect of an entire season's participation in the interscholastic sports of basketball, wrestling, and hockey on physical fitness. This was done by comparing pre- and post-season test results. The participants were members of either the varsity or junior varsity basketball teams, the varsity hockey team, and the varsity wrestling team at Devils Lake High School. The

---

[18]Clifford J. Moser, "A Comparison of the Effects of Seasonal Participation in Selected Interscholastic Sports on Physical Fitness" (unpublished Master's thesis, University of North Dakota, 1961).

tests were the sit-ups, pull-ups, treadmill, shuttle run, agility dribble, and the standing broad jump. Moser was attempting to learn the amount of improvement in physical fitness of each group of athletes participating in the three different sports.

Moser made the following conclusions:

1. The basketball group was in the best physical condition at the time of the post test.

2. The basketball group was at the best pre-season physical fitness level.

3. All three of the athletic programs improved the physical fitness levels of the participants.

4. Based upon the results of this study, it seems apparent that the athletic programs at Devils Lake Central do very little to improve the participants in agility as measured by the shuttle run and agility dribble.

5. The wrestling group improved the most over the duration of the season, with a "t" value of 8.209.

6. According to the results of this study, it would seem obvious that the wrestling group did a lot of work to improve the abdominal muscles as indicated by a mean improvement in the raw score totals of 18.143 repetitions in sit-ups.

7. It is apparent that, though the basketball group ranked highest in physical fitness at the beginning of the experiments, they did not improve as much physically as did the wrestling group.

8. As a result of this study, it can be concluded that the wrestling group improved the most in physical fitness, as measured by the tests given, with a "t" value of 8.209. The basketball group was second with a "t" value of 5.809, and the hockey group was last with a "t" value of 4.469.

Summary of Related Literature

From the review of literature, it is evident that athletic participation appears to improve the physical fitness levels of the participants. Also, there is agreement among the researchers that required physical education programs over a four-year period contribute more to physical fitness than do elective programs.

Between groups comparisons of physical fitness levels involving participants in interscholastic athletic programs as against non-participants have consistently shown that groups of athletes make significantly greater improvement in fitness than do non-athletes.

CHAPTER II

PROCEDURE

The tests were administered in accordance with the recommendations and instructions of the American Association for Health, Physical Education, and Recreation Youth Fitness Test Manual.[1] The writer had to make one substitution in the test. At the University of North Dakota the shot put was used as a testing procedure for arm power instead of the softball throw. The method and procedure used in selection of groups, organization of facilities, and supervision of the testing have been presented in this chapter.

Selection of Groups

The selection of groups was accomplished by obtaining a list from the Dean of Men at the University of North Dakota of all male students who enrolled as freshmen in the fall of 1963. These freshmen were then divided into two groups, athletes and non-athletes. The athletes were considered the experimental group and consisted of those male students participating in intercollegiate football, basketball, wrestling, cross country and/or hockey for four years at the University of North Dakota. The non-athletes were considered the control group

---

[1]AAHPER, Youth Fitness Test Manual (Washington 6, D. C., The American Association for Health, Physical Education, and Recreation, 1958).

and consisted of randomly selected male students who had no inter-collegiate athletic participation while attending the University of North Dakota for four years.

Test Administration

An indoor gymnasium was used for the administration of the first part of the test which included the sit-ups, pull-ups, shuttle run, and standing broad jump. The second part was administered in the fieldhouse at the University of North Dakota two days later. This included the fifty yard dash, shot put, and six hundred yard run-walk. The initial tests had been given to each group by the physical education staff at the University of North Dakota in the fall of 1963. The re-tests were given to each group in the fall of 1966.

The subjects of both groups were given instructions on the execution of all phases of the test.

Sit-Ups

Equipment: Sit-ups were done on mats placed on the gym floor.

Procedure: The subject lay on his back with the knees bent at forty-five degrees. His hands were placed on the back of the neck with the fingers interlocked. A partner held the ankles down, the heels being in contact with the floor at all times.

The subject then sat-up, touching his elbows to the knees, and returned to the starting position.

Rules: 1. The fingers had to remain in contact behind the neck throughout the exercise.

2. The knees had to remain at a forty-five degree angle.

3. When returning to starting position, elbows had to be flat on the floor before sitting up again.

Scoring: One point was given for each complete movement of touching elbows to knees. No score was counted if the fingertips did not maintain contact behind the head, if the knees did not remain at a forty-five degree angle, or if the subject did not return correctly to the starting position. The maximum limit in terms of number of sit-ups was one hundred and thirty-one.

Pull-Ups

Equipment: A metal bar approximately one and one-half inches in diameter.

Procedure: The bar was high enough so that each subject could hang with his arms and legs fully extended and his feet free of the floor. The overhand grip was used. After assuming the hanging position, the subject raised his body by his arms until his chin could be placed over the bar and then lowered his body to a full hang as in the starting position. The exercise was repeated as many times as possible or until the subject reached the number of twenty-six.

Rules: 1. One trial was given to each subject unless it was obvious that the subject did not have a fair chance.

2. The knees could not be raised and kicking of the legs was not permitted.

3. The body could not swing during the execution of the movement. The pull could in no way be a snap movement. If the subject

started swinging he was checked by holding an extended arm across the front of his thighs.

Scoring: The number of completed pull-ups to the nearest whole number was recorded with a limit of twenty-six.

Standing Broad Jump

Equipment: Ruled markings of feet and inches on gym floor.

Procedure: The subject stood at a spot marked as the take-off line. Preparatory to jumping, the subject swung the arms backward and bent the knees. The jump was accomplished by simultaneously extending the knees and swinging the arms forward.

Rules: 1. Three trials were allowed.

2. The distance of the jump was measured from the take-off line to the heel or other part of the body that touched the floor nearest the take-off line.

3. The scorer stood to one side and observed the mark to the nearest inch.

Scoring: The best of three trials in feet and inches to the nearest inch was recorded.

Shuttle Run

Equipment: A stopwatch and two blocks of wood, 2 inches x 2 inches x 4 inches.

Procedure: Two parallel lines were marked on the floor thirty feet apart. The blocks of wood were placed behind one of the lines. The subject started from behind the other line on the signal "Ready?

Set! Go!" The subject ran to the blocks, picked one up, ran back to the starting line, and placed the block which he carried back across the starting line. The procedure was repeated for the second block.

Rules: 1. Two trials were allowed.

2. The blocks could not be thrown on the floor at the starting line. The subject had to set each block on the floor.

Scoring: The better of the two trials to the nearest tenth of a second was selected.

50-Yard Dash

Equipment: A stopwatch with a split second timer was used.

Procedure: The subject took his position behind the starting line and started on the commands "Ready? Set! Go!" The latter was accompanied by a downward swing of the starter's arm to give the timer a visual signal.

Rules: 1. The subject had to have his hands behind the starting line.

2. The score was the amount of time between the starter's signal and the instant the subject crossed the finish line.

Scoring: The seconds to the nearest tenth of a second were recorded.

Shot Put

Equipment: Shot put (12 pounds), lines marked in feet, and a tape measure.

Procedure: The area was marked according to the rules governing the shot put in track and field. The subject had to stand with his foot touching the toe board at the front edge of the circle. The subject was allowed to pivot if he so desired. His throw was marked according to where the ball landed. If his second throw was farther, this marking was used. The measurement was then taken and recorded.

Rules: 1. Only the pushing action was allowed.

2. Two trials were allowed.

3. The distance recorded was the distance from the point of landing to the front of the toe board in the ring.

Scoring: The best of the two throws to the nearest foot was recorded.

600-Yard Run-Walk

Equipment: Track marked accordingly and a stop watch.

Procedure: The subject started from a standing start. At the signal "Ready? Set! Go!" the subject started running the 600 yard distance.

Rules: Walking was permitted, but the object was to cover the distance in the shortest possible time.

Scoring: The time was recorded in minutes and seconds to the nearest second.

Following the collection of data, it became necessary to choose a statistical method that would test the significance of the differences between the two groups and within the groups.

Statistical Procedure

In order to determine the fitness levels of the two groups, tests were given on a test and retest basis. The means of each test item on the test and retest were determined. The means of the groups could then be compared to determine the actual differences between the means. The data was statistically analyzed to determine the significance of differences between the means within the groups. Comparisons were also made in the final individual test item scores to determine the significance of the mean differences between the two groups.

This study assumed the null hypothesis with respect to differences between the means of the two groups. This hypothesis asserts that there is no true difference between the two mean scores and that the difference, if any, would actually be a chance occurrence of no significance.[2]

After investigating several techniques of testing the null hyopthesis, it appeared that the "t" technique for testing the significance of difference between means derived from correlated scores of small samples was the best to use in this study.[3]

The ratio thus derived is expressed as "t" and is checked for significance on table 29, found in Garrett's book on statistics.[4] The value of "t" is proportional to the degrees of freedom (N - 1) allowed

---

[2]Henry E. Garrett, Statistics in Psychology and Education (New York: Longmans, Green and Co., 1955).

[3]Quinn McNemar, Psychological Statistics (New York and London: John Wiley and Sons, Inc., 1962).

[4]Garrett, op. cit., p. 464.

in determining the relationship between the mean difference and the estimate of sampling error of the mean difference.[5]

It was also necessary to use the method outlined by Garrett for dealing with uncorrelated groups.[6] This procedure takes into consideration the standard error of the difference between two means to find the "t" value. The material was then treated as listed above using the "t" table to check for the significance of the difference between the mean differences of each group.

For the purpose of this study, the .01 level of confidence was selected.

Complete data including mean difference and raw and T scores, together with the details of the mathematical process employed in analysis for each testing area is presented in Appendix A, page 53.

---

[5]McNemar, op. cit.

[6]Garrett, op. cit., p. 209

CHAPTER III

ANALYSIS OF DATA

The purpose of the testing in this study was to discover whether there were any significant differences between fitness levels of the experimental group as compared to the fitness levels of the control group. The bases of comparison were results obtained through the use of the American Association for Health, Physical Education and Recreation Modified Youth Fitness Test.

It appeared that some uncontrolled bias was present as evidenced by the fact that, in all the initial test items, the mean of the control group was inferior to the mean of the experimental group except in the broad jump. This fact would appear to favor the control group as there was more opportunity to show improvement.

Results of Comparison

Sit-Ups

The control group had a mean score of 84.45 sit-ups in the initial test and a mean score of 53.20 sit-ups in the retest (see Table 1 on page 34).

The control group had a mean difference of 31.25 decrease in sit-ups between the initial test and the retest. The estimate of sampling error of the mean difference was 7.20. The "t" value of

-4.34 with 19 degrees of freedom indicated a significant difference at the criterion .01 level.

In the initial test, the experimental group had a mean score of 113.70 sit-ups, and, in the retest, this group had a mean score of 113.40 sit-ups.

The experimental group had a mean difference of .30 decrease in sit-ups between the initial test and retest. The estimate of the sampling error of mean difference was 4.008. The "t" value of -.075 with 19 degrees of freedom was not significant at the criterion .01 level.

Pull-Ups

The control group had a mean score of 7.85 pull-ups in the initial test and a mean score of 7.40 pull-ups in the retest (as shown in Table 1 on page 34). These tests measured arm and shoulder-girdle strength.

A mean difference of .45 decrease in pull-ups between the initial test and the retest was shown by the control group. The estimate of the sampling error of mean difference was .844. The "t" value of -.533 with 19 degrees of freedom was not significant at the criterion .01 level.

The experimental group had a mean score of 10.50 pull-ups in the initial test and a mean score of 13.25 pull-ups in the retest.

A mean difference of 2.70 increase in pull-ups between the initial test and the retest was shown by the experimental group.

The estimate of sampling error of mean difference was .764. The "t" value of 3.599 with 19 degrees of freedom indicated a significant improvement at the criterion .01 level.

Standing-Broad-Jump

The control group had a mean score of 90.95 inches in the standing broad jump in the initial test and a mean score of 90.90 inches in the retest (see Table 1 on page 34). These tests measured the explosive power of the legs.

A mean difference of .05 decrease between the initial test and the retest was shown by the control group. The estimate of sampling error of mean difference was .835. The "t" value of -.06 with 19 degrees of freedom was not significant at the criterion .01 level.

The experimental group had a mean score of 90.20 inches in the standing broad jump in the initial test and a mean score of 95.60 inches in the retest (see Table 1 on page 34).

A mean difference of 5.40 increase between the initial test and the retest was shown by the experimental group. The estimate of sampling error of mean difference was 1.35. The "t" value of 4.00 with 19 degrees of freedom indicated a significant improvement at the criterion .01 level.

Shuttle Run

The control group had a mean score of 9.925 seconds in the initial test and a mean score of 9.495 seconds in the retest (as shown in Table 1 on page 34). These tests measured agility and speed.

A mean difference of .430 seconds faster between the initial test and the retest was shown by the control group. The estimate of the sampling error of mean difference was .147. The "t" value of 2.925 with 19 degrees of freedom indicated a significant improvement at the criterion .01 level.

The experimental group had a mean score of 9.54 seconds in the initial test and a mean score of 8.765 seconds in the retest (see Table 1 on page 34).

A mean difference of .775 seconds faster between the initial test and the retest was shown by the experimental group. The estimate of sampling error of mean difference was .109. The "t" value of 2.86 with 19 degrees of freedom indicated a significant improvement at the criterion .01 level.

50-Yard Dash

The control group had a mean score of 6.48 seconds in the initial test and a mean score of 6.535 seconds in the retest (see Table 1 on page 34). These tests measured speed.

A mean difference of .055 seconds slower between the initial test and the retest was shown by the control group. The estimate of sampling error of mean difference was .07. The "t" value of -.786 with 19 degrees of freedom was not significant at the criterion .01 level.

The experimental group had a mean score of 6.27 seconds in the initial test and a mean score of 6.135 seconds in the retest (as shown in Table 1 on page 34).

A mean difference of .135 seconds faster between the initial test and the retest was shown by the experimental group. The estimate of sampling error of mean difference was .054. The "t" value of 2.50 with 19 degrees of freedom was not significant at the criterion .01 level.

Shot-Put

The control group had a mean score of 28.3 feet in the initial shot-put test and a mean score of 29.35 feet in the retest (see Table 1 on page 34). These tests measured the explosive power of the arm.

A mean difference of 1.05 feet increase between the initial test and the retest was shown by the control group. The estimate of sampling error of mean difference was .420. The "t" value of 2.50 with 19 degrees of freedom was not significant at the criterion .01 level.

The experimental group had a mean score of 31.65 feet in the initial test and a mean score of 34.95 feet in the retest (as shown in Table 1 on page 34).

A mean difference of 3.30 feet increase was shown by the experimental group in the initial test and the retest. The estimate of sampling error of mean difference was .343. The "t" value of 9.913 with 19 degrees of freedom indicated a significant improvement at the criterion .01 level.

600-Yard Run-Walk

The control group had a mean score of 102.85 seconds in the initial test and a mean score of 115.55 seconds in the retest (as

shown in Table 1 on page 34). These tests measured muscular and cardio-respiratory endurance.

A mean difference of 12.70 seconds slower between the initial test and the retest was shown by the control group. The estimate of sampling error of mean difference was 2.736. The "t" value of -4.642 with 19 degrees of freedom indicated a significant decline at the criterion .01 level.

The experimental group had a mean score of 99.35 seconds in the initial test and a mean score of 96.30 seconds in the retest (see Table 1 on page 34).

A mean difference of 3.05 seconds faster between the initial test and the retest was shown by the experimental group. The estimate of sampling error of mean difference was .869. The "t" value of 3.51 with 19 degrees of freedom indicated a significant improvement at the criterion .01 level.

As shown in Table 1 on page 34, the control group improved the mean scores in three test items: shuttle run, 50-yard dash and the shot-put. The experimental group had a mean score improvement in all test items except the sit-ups and this test item showed only a slight decline.

TABLE 1

MEAN SCORES IN TESTS OF SUBJECTS IN CONTROL GROUP

Name of test	Number	Initial Test	Retest
Sit-ups	20	84.45	53.20
Pull-ups	20	7.85	7.40
Standing Broad Jump	20	90.95	90.90
Shuttle Run	20	9.925	9.495
50-Yard Dash	20	6.48	6.535
Shot-Put	20	28.3	29.35
600-Yard Run-Walk	20	102.85	115.55

MEAN SCORES IN TESTS OF SUBJECTS IN EXPERIMENTAL GROUP

Name of test	Number	Initial Test	Retest
Sit-ups	20	113.70	113.40
Pull-ups	20	10.50	13.25
Standing Broad Jump	20	90.20	95.60
Shuttle Run	20	9.54	8.765
50-Yard Dash	20	6.27	6.135
Shot-Put	20	31.65	34.95
600-Yard Run-Walk	20	99.35	96.30

As shown by the analysis of data presented in Table 2 below, the experimental group exhibited significant improvement in all items except the 50-yard dash and sit-ups. In only one test item, the shuttle run, did the control group show any significant improvement. The control group showed a significant decline in the 600-yard run-walk and in sit-ups.

TABLE 2

"t" AND THE SIGNIFICANCE OF DIFFERENCE WITHIN GROUPS

Area of Comparison	"t" Value of Control Group	"t" Value of Experimental Group
Sit-ups	-4.34 Significant at .01 level	-.075 Not Significant
Pull-ups	-.533 Not Significant	3.599 Significant at .01 level
Standing Broad Jump	-.06 Not Significant	4.0 Significant at .01 level
Shuttle Run	2.925 Significant at .01 level	7.11 Significant at .01 level
50-Yard Dash	-.786 Not Significant	2.50 Not Significant
Shot-Put	2.50 Not Significant	9.913 Significant at .01 level
600-Yard Run-Walk	-4.642 Significant at .01 level	3.51 Significant at .01 level

Since both groups revealed changes between the means of the initial test and the retest, it was decided to test further for

possible differences between the two groups. The null hypothesis was assumed with respect to the differences between the two groups on values of mean differences found on the initial test and the retest between the two groups. The null hypothesis was tested in this case by the use of the "t" technique for uncorrelated data from small samples. The "t" values were checked with a table of "t" from Garrett[1] to ascertain whether the results were statistically significant. The results of the comparison of "t" and the significance of difference between the two groups are shown in Table 3 on page 40.

Sit-ups

The mean differences between the initial test and the retest were .30 decrease in sit-ups for the experimental group and 31.25 decrease in sit-ups for the control group. The difference between the mean differences of the two groups was 30.95 sit-ups. The estimate of the sampling error for the distribution of the differences between the mean differences was 8.24. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 3.756. With 38 degrees of freedom, this "t" value indicated a significant difference at the criterion .01 level. The decline on the part of the control group during the experimental period was significantly greater than that of the experimental group.

---

[1]Garrett, op. cit., p.449.

Pull-Ups

The mean differences between the initial test and the retest were 2.70 increase in pull-ups for the experimental group and .45 decrease in pull-ups for the control group. The difference between the mean differences of the two groups was 3.15 pull-ups. The estimate of the sampling error for the distribution of the differences between the mean differences was 1.14. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 2.80. With 38 degrees of freedom, this "t" value indicated a significant difference at the criterion .01 level.

Standing-Broad Jump

The mean differences between the initial test and the retest were 5.40 inches increase in the standing broad jump for the experimental group and .05 inches decrease in the standing broad jump for the control group. The difference between the mean differences of the two groups was 5.45 inches. The estimate of the sampling error for the distribution of the differences between the mean differences was 1.587. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 3.43. With 38 degrees of freedom, this "t" value indicated a significant difference at the criterion .01 level.

Shuttle Run

The mean differences between the initial test and the retest were .775 seconds faster in the shuttle run for the experimental group and .430 seconds faster for the control group. The difference between the mean differences of the two groups was .345 seconds. The estimate of the sampling error for the distribution of the differences between the mean differences was .183. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 1.89. With 38 degrees of freedom, this "t" value was not significant at the criterion .01 level.

50-Yard Dash

The mean differences between the initial test and the retest were .135 seconds faster in the 50-yard dash for the experimental group and .055 seconds slower for the control group. The difference between the mean differences of the two groups was .190 seconds. The estimate of the sampling error for the distribution of the differences between the mean differences was .0883. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 2.15. With 38 degrees of freedom, this "t" value was not significant at the criterion .01 level.

Shot-Put

The mean differences between the initial test and the retest were 3.30 feet increase in the shot-put for the experimental group and 1.05 feet increase for the control group. The difference between the mean differences of the two groups was 2.25 feet. The estimate of the sampling error for the distribution of the differences between the mean differences was .542. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 4.34. With 38 degrees of freedom, this "t" value indicated a significant difference at the criterion .01 level.

600-Yard Run-Walk

The mean differences between the initial test in the 600-yard run-walk and the retest were 3.05 seconds faster for the experimental group and 12.70 seconds slower for the control group. The difference between the mean differences of the two groups was 15.75 seconds. The estimate of the sampling error for the distribution of the differences between the mean differences was 2.87. The "t" value resulting from the relationship of the actual difference between the mean differences of the two groups and the estimate of the sampling error for the distribution of the differences between the mean differences was 5.49. With 38 degrees of freedom, this "t" value indicated a significant difference at the criterion .01 level.

As shown by the analysis of the data presented in Table 3 below, the comparison of the means between the two groups indicated that there was a significant difference at the criterion .01 level in all the test items except the 50-yard dash and the shuttle run.

TABLE 3

"t" AND SIGNIFICANCE OF DIFFERENCES

BETWEEN THE TWO GROUPS

Test Item Compared	"t" Value	.01 Level	Level Significance
Sit-ups	3.756	2.71	Significant at .01 level
Pull-ups	2.80	2.71	Significant at .01 level
Standing Broad Jump	3.43	2.71	Significant at .01 level
Shuttle Run	1.89	2.71	Not Significant at .01 level
50-Yard Dash	2.15	2.71	Not Significant at .01 level
Shot-Put	4.34	2.71	Significant at .01 level
600-Yard Run-Walk	5.49	2.71	Significant at .01 level

CHAPTER IV

DISCUSSION

Physical education is a very important part of the total educational process. Physical fitness is a very important area of the physical education program. With these thoughts in mind, the investigator decided to see what the trends in physical fitness levels would be for athletes and non-athletes at the University of North Dakota over a four year period.

Many questions have been asked as to reasons why students should be required to participate in physical education. Does an individual really need to become physically fit? There are few people who maintain a very high level of physical fitness, but would an individual even attempt to stay in any physical condition at all if this same person had not been required to participate in a physical education program.

This writer strongly suggests that physical education not only be required for the first two years in college, but that the program be expanded even further. Ideally, physical education classes should be conducted five days a week from the time a student enters school at the age of six until the completion of his education. A program of this nature could certainly influence many members of American society favorably with respect to interest in their physical welfare.

When the investigator decided to undertake this study comparing fitness levels of athletes to non-athletes, he realized immediately, due to his review of previous studies, that the athlete should be superior to the non-athlete. However, by comparing the two groups over a four year period the study took on a new significance. The investigator could now determine whether a required program of physical education for two years was sufficient for the non-athlete to maintain the physical fitness level he had achieved while participating in the required physical education program. Also, the writer could determine if an intercollegiate athletic program for the athlete would help this individual maintain the fitness level that the initial test results indicated.

The results of the study showed that the experimental group (athletes) improved significantly in all measures except the 50-yard dash and sit-ups. The control group (non-athletes) improved significantly in one test item, the shuttle run, and decreased significantly in the 600-yard run-walk and sit-ups. There was reason to believe that there may have been a significant improvement in sit-ups for the experimental group if a maximum limit, as to the number an individual could perform, had not been placed on the test. The test for sit-ups allowed each individual to do 131 sit-ups. When this maximum limit was reached, he was required to stop the test exercise. Test results showed that 12 athletes reached the maximum limit in the retest as compared to not one non-athlete. Even though the writer could not prove any significant

improvement for the experimental group between the initial test and the retest, results would indicate that there was improvement by the athletes.

The investigator had no control over the non-athlete group as far as their participation in any activities. It was impossible to determine how much time any of the participants had spent in maintaining a fitness level or trying to improve themselves. Test results showed that there was significant improvement at the criterion .01 level in the shuttle run. It is the opinion of this writer that physical maturity was a factor which produced this significance. Statistical results in Appendix A, page 53 show that the majority of the control group (non-athletes) improved in this area. Yet, in all of the other test items the majority of the participants did not improve. From these results, it is this writer's opinion that the shuttle run does not require a very high degree of physical fitness. In other words agility and fitness may not be highly correlated.

In the investigator's opinion there was considerable bias present in the selection of the non-athlete group. The investigator firmly believed that the non-athlete group which was tested was in the upper percentile in physical fitness for this group. The writer also believed that only those non-athletes who were at an above average level of physical condition accepted the invitation to be a part of this experiment. It is the investigator's opinion that the majority of the non-athletes participating in this study also participated in the intra-mural program.

After the investigator had administered the test to the individual participants of the control group, some of them had opinions as to why they decreased in the different test items.

Following are some of the statements made by the subjects:

"I just didn't realize how poor a shape I was really in until I took this test."

"It is impossible for me to stay in physical condition because I just don't have the time to work out."

"If they had more courses to offer in the physical education program, I know I would have been in better physical condition to take this test."

"Too many of us have taken for granted our physical condition. After seeing the results of this test, I, for one, am going to try to improve my condition."

As can be seen from some of the above quotations some individuals recognized that they have not maintained acceptable levels of physical condition. It was encouraging to the writer that these students recognized their lack of physical stamina and that many of them would try to improve themselves in this area.

The experimental group retained their physical fitness levels as indicated by the comparison of the results of the two tests. Results would indicate that a program of intercollegiate athletics can and does help the participants to maintain very high levels of physical fitness. The writer realizes that not all individuals

can be athletes and achieve such levels of physical condition, but can there be valid reasons as to why the non-athlete group should not work toward such a goal of physical fitness?

It is the opinion of this writer that the test results indicated that there was a gap in the physical education program for upper division college students. Somehow the physical education program is failing those students who do not participate in an intercollegiate athletic program. The need for more and better physical education programs in the elementary, secondary, and higher institutions is apparent. The program must be expanded to include every student who desires to achieve and maintain a minimal level of physical condition. It is the duty of educators to provide such a service to all of these students.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The forty subjects selected for this study were freshmen who enrolled at the University of North Dakota in September of 1963. The experimental group consisted of those male students who had participated in intercollegiate football, basketball, cross-country, wrestling, and/or hockey for three years at the University of North Dakota. The control group was limited to those male students who had enrolled at the University of North Dakota in the fall of 1963, but had never participated in any type of intercollegiate athletics. The physical fitness level of each group was determined by the results obtained from the American Association for Health, Physical Education, and Recreation Modified Youth Fitness Test. The test was administered in the fall of 1963 by the physical education staff at the University of North Dakota and both groups were tested by the writer during the first semester of 1966-67. The test results of the experimental group were compared to the test results of the control group to determine whether any significant changes were evident in the selected measures of physical fitness.

Comparisons were made between the mean differences within each group as indicated by the scores on the initial and final tests. The null hypothesis was assumed with respect to the differences within the group on the initial test and the final test. This hypothesis was tested with the "t" technique for the difference between means derived from correlated scores from small samples. Comparisons were also made between the experimental group and the control group by testing the significance of the difference between the mean differences found within the groups. The between group comparison used the "t" technique for uncorrelated data from small samples.

Conclusions

The following conclusions seem warranted on the basis of the data collected in this study:

1. The experimental group (athletes) showed significant improvement in all of the selected measures of physical fitness except the 50-yard dash and sit-ups at the criterion .01 level.

2. The control group (non-athletes) improved significantly in one selected measure of physical fitness, the shuttle run, at the criterion .01 level.

3. The control group showed a significant decrease in the 600-yard run-walk and sit-ups at the criterion .01 level.

4. In comparing the experimental and control groups, the writer found significant differences at the criterion .01 level in the following test items: sit-ups, pull-ups, standing broad jump, and the 600-yard run-walk.

5. The intercollegiate athletic program at the University of North Dakota seems to help maintain and improve the physical fitness levels of athletes.

Recommendations

The following recommendations are made relative to this study.

1. The physical education program at the University of North Dakota should be evaluated to see what improvements could be made to help students who do not participate in the intercollegiate athletic program in developing and maintaining physical fitness.

2. A study should be conducted comparing the control group at the University of North Dakota to other non-athletes in a required physical education program at another University.

3. A similiar study should be conducted comparing the two groups with tests other than the American Association for Health, Physical Education, and Recreation Modified Youth Fitness Test that the writer used in this study.

4. A study should be conducted testing the non-athletes who do not participate in any activity after the two year required program of physical education as compared to those non-athletes who participate in intramurals or some other form of activity.

5. It is further recommended that a follow-up study should be conducted involving these same two groups five years from the date of this study to see how they would compare in all the levels

of physical fitness by use of the American Association for Health, Physical Education, and Recreation Modified Youth Fitness Test.

BIBLIOGRAPHY

Books

Cureton, Thomas Kirk. Physical Fitness Appraisal and Guidance. St. Louis: The C. V. Mosby Company, 1947.

Garrett, Henry E. Statistics in Psychology and Education. New York: Longmans, Green and Co., 1955.

McNemar, Quinn. Psychological Statistics. New York and London: John Wiley and Sons, Inc., 1962.

Articles and Periodicals

AAHPER. Youth Fitness Test Manual. Washington 6, D. C.: The American Association for Health, Physical Education, and Recreation, 1958.

Blesh, T. Erwin, and Scholz, Alfred E. "Ten Year Survey of Physical Fitness Tests at Yale University," Research Quarterly, Vol. 28 (December, 1957), 321-326.

Clarke, H. Harrison. "Physical Fitness of University of Oregon Male Freshmen," Physical Fitness News Letter, University of Oregon, (March, 1955).

Clarke, H. Harrison. "Oregon Pilot Physical Fitness Project," The Physical Educator, (May, 1957), 55.

Hamilton, Tom. "Athletic Director Looks at Fitness," Journal of Health, Physical Education and Recreation, Vol. 28 (September, 1957), 14-15.

Norman, Gerald F. "Comparison of Athletic Achievement in Selected Activities Among Secondary School Athletics," Completed Research in Health, Physical Education and Recreation, Vol. 1 (1959), 28.

Starr, Helen M. "How to Fit in Fitness Testing," Journal of Health, Physical Education and Recreation, Vol. 30, (March, 1958), 19.

Wieneke, Kuhst. "A Comparison of Certain Physical Developments of Freshman Athletes and Non-Athletes," Research Quarterly, Vol. 3, (May, 1932), 223-234.

Unpublished Material

Boschee, Floyd. "A Comparison of the Physical Fitness Levels of Selected Participants in Interscholastic Football Before the Season, and One Month Later." Unpublished Master's thesis, Department of Physical Education, University of North Dakota, 1961.

Campbell, Charles G. Jr. "A Study to Determine the Physical Fitness Status of the Male Students Enrolled in 110 and 220 Physical Education Classes at North Carolina College, Durham, at Durham, North Carolina, During the School Year 1961-62." Unpublished Master's thesis, Department of Physical Education, North Carolina College, Durham, North Carolina, 1962.

De La Barre, Craig H. "A Study of the Fitness Values Derived from a Five Day Per Week Activity Compared with Like Values from a Two Day Per Week Activity." Unpublished Master's thesis, Department of Physical Education, University of North Dakota, July, 1962.

Gaddie, Michael L. "A Comparison of Athletes and Non-Athletes at the University of North Dakota as Measured by the Harvard Step Test." Unpublished Master's thesis, Department of Physical Education, University of North Dakota, 1960.

Hoffman, Vern B. "Strength Comparison of Athlete and Non-Athlete as Measured by Roger's Physical Capacity Test." Unpublished Master's thesis, Department of Physical Education, University of Michigan, 1936.

Moser, Clifford J. "A Comparison of the Effects of Seasonal Participation in Selected Interscholastic Sports on Physical Fitness." Unpublished Master's thesis, Department of Physical Education, University of North Dakota, 1961.

Olson, Noel W. "A Comparison of the Effects of Seasonal Participation in Selected Interscholastic Sports on Motor Fitness." Unpublished Master's thesis, Department of Physical Education, University of North Dakota, 1961.

Rains, David D. "Growth of Athletes and Non-Athletes in Selected Secondary Schools as Assessed by Grid Technique." Unpublished Ph. D. dissertation, Indiana University, 1951.

Sundre, Orlo A. "A Comparative Study of Two Physical Education Programs for Male Students at the University of North Dakota." Unpublished Master's thesis, Department of Physcial Education, University of North Dakota, 1960.

Vinger, Richard M. "A Comparison of Physical Fitness Increases as the Result of a Selected Physical Education Program." Unpublished Master's thesis, Department of Physical Education, University of North Dakota, 1964.

INITIAL TEST AND RETEST OF CONTROL GROUP IN SIT-UPS

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	69	41	-28	784
2.	65	25	-40	1600
3.	51	60	9	81
4.	100	45	-55	3025
5.	60	35	-25	625
6.	131	35	-96	9216
7.	131	118	-13	169
8.	110	60	-50	2500
9.	53	40	-13	169
10.	71	90	19	361
11.	84	58	-26	676
12.	58	45	-13	169
13.	83	47	-36	1296
14.	75	70	- 5	25
15.	91	54	-37	1369
16.	75	70	- 5	25
17.	101	30	-71	5041
18.	75	40	-35	1225
19.	75	57	-18	324
20.	131	44	-87	7569
	1689	1064	-625	39249

Mean Score of Initial Test	84.45
Mean Score of Retest	53.20
Sum of Differences	-625
Sum of Dif. Squared	10390

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Sit-Ups GROUP Control

N = 20

D = -625

D^2 = 39249

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$ $= \frac{S_D}{\sqrt{N}} =$

$$\sqrt{\frac{D^2 - \frac{(D)^2}{N}}{N-1}} \qquad \sqrt{\frac{39249 - \frac{(-625)^2}{20}}{19}}$$

$$\frac{}{\sqrt{N}} \qquad \frac{}{\sqrt{20}}$$

$S_{\bar{D}}$ = 7.20

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{-625}{20} = 31.25$

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{-31.25}{7.20} = -4.34$

df = N - 1 = 19

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF CONTROL GROUP IN PULL-UPS

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	5	4	-1	1
2.	11	9	-2	4
3.	1	5	4	16
4.	9	6	-3	9
5.	8	6	-2	4
6.	5	5	0	0
7.	8	9	1	1
8.	7	7	0	0
9.	7	5	-2	4
10.	7	16	9	81
11.	11	7	-4	16
12.	10	6	-4	16
13.	4	6	2	4
14.	15	10	-5	25
15.	8	9	1	1
16.	7	14	7	49
17.	7	4	-3	9
18.	11	8	-3	9
19.	5	6	1	1
20.	11	6	-5	25
	157	148	-9	275

Mean Score of Initial Test 7.85
Mean Score of Retest 7.40
Sum of the Differences -9
Sum of Dif. Squared 275

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Pull-Ups GROUP Control

N = 20

D = - 9

D^2 = 275

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\sqrt{\frac{D^2 - \frac{(D)^2}{N}}{N - 1}} \qquad \sqrt{\frac{275 - \frac{(-9)^2}{N}}{19}}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}}$ = .844

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{-9}{20} = -.45$

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{-.45}{.844} = -.533$

df = N - 1 = 19

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF CONTROL GROUP IN BROAD JUMP

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	90	92	2	4
2.	87	92	5	25
3.	85	84	-1	1
4.	84	85	1	1
5.	93	90	-3	9
6.	87	88	1	1
7.	99	101	2	4
8.	94	96	2	4
9.	86	88	2	4
10.	91	92	1	1
11.	97	106	9	81
12.	96	94	-2	4
13.	85	86	1	1
14.	91	86	-5	25
15.	100	97	-3	9
16.	96	97	1	1
17.	89	84	-5	25
18.	86	86	0	0
19.	93	85	-8	64
20.	90	89	-1	1
	1819	1818	-1	265

Mean Score of Initial Test	90.95
Mean Score of Retest	90.90
Sum of Differences	-1
Sum of Dif. Squared	265

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Broad Jump GROUP Control

$N = 20$

$D = -1$

$D^2 = 265$

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= S_{\frac{D}{\sqrt{N}}} =$

$$\sqrt{\frac{D^2 - \frac{(D)^2}{N}}{N - 1}} \Big/ \sqrt{N} \qquad \sqrt{\frac{265 - \frac{(-1)^2}{N}}{19}} \Big/ \sqrt{20}$$

$S_{\bar{D}} = .835$

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{-1}{20} = .05$

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{-.05}{.835} = -.06$

$df = N - 1 = 19$

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF CONTROL GROUP IN SHUTTLE RUN

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	10.9	10.1	+ .8	.64
2.	9.4	9.3	+ .1	.01
3.	10.3	10.2	+ .1	.01
4.	10.4	9.6	+ .8	.64
5.	11.0	9.2	+1.8	3.24
6.	9.6	9.1	+ .5	.25
7.	10.1	8.6	+1.5	2.25
8.	10.0	8.8	+1.2	1.44
9.	10.8	10.1	+ .7	.49
10.	9.1	9.5	- .4	.16
11.	9.4	8.8	+ .6	.36
12.	8.8	9.4	- .6	.36
13.	10.0	10.0	.0	.0
14.	9.8	9.5	+ .3	.09
15.	9.9	9.3	+ .6	.36
16.	9.4	9.2	+ .2	.04
17.	10.0	9.8	+ .2	1.44
18.	10.0	9.7	+ .3	.09
19.	10.4	10.6	- .2	.04
20.	9.2	9.1	+ .1	.01
	198.5	189.9	+8.6	11.92

Mean Score of Initial Test	9.925
Mean Score of Retest	9.495
Sum of the Difference	8.6
Sum of Dif. Squared	11.92

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Shuttle Run GROUP Control

$N = 20$

$D = 8.6$

$D^2 = 11.92$

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$)

$$= S_{\bar{D}} = \frac{\sigma_D}{\sqrt{N}} =$$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N - 1} \qquad \frac{\sqrt{11.92 - \frac{(8.6)^2}{20}}}{19}$$

$$\frac{}{\sqrt{N}} \qquad \frac{}{\sqrt{20}}$$

$S_{\bar{D}} = .147$

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{8.6}{20} = .43$

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{.43}{.147} = 2.925$

$df = N - 1 = 19$

"t" at .01 level = 2.86

Significant at the .01 level

INITIAL TEST AND RETEST OF CONTROL GROUP IN 50 YARD DASH

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	6.7	6.8	-.1	.01
2.	6.2	6.6	-.4	.16
3.	7.1	6.9	.2	.04
4.	6.1	6.4	-.3	.09
5.	6.0	6.4	-.4	.16
6.	6.6	6.4	.2	.04
7.	6.2	6.5	-.3	.09
8.	6.4	6.3	.1	.01
9.	6.2	6.5	-.3	.09
10.	6.4	6.4	.0	.00
11.	6.0	6.0	.0	.00
12.	6.2	6.4	-.2	.04
13.	7.1	6.3	.8	.64
14.	6.4	6.5	-.1	.01
15.	6.3	6.4	-.1	.01
16.	6.9	6.5	.4	.16
17.	6.8	6.9	-.1	.01
18.	6.4	7.0	-.6	.36
19.	7.0	7.0	.0	.00
20.	6.6	6.5	.1	.01
	129.6	130.7	-1.1	1.93

Mean Score of Initial Test	6.48
Mean Score of Retest	6.535
Sum of the Difference	-1.1
Sum of Dif. Squared	1.93

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST 50-Yard Dash GROUP Control

N = 20

D = -1.1

D^2 = 1.93

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N - 1} \qquad \frac{\sqrt{1.93 - \frac{(-1.1)^2}{20}}}{19}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}}$ = .07

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{-1.1}{20} =$ -.055

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{-.055}{.07} =$ -.786

df = N - 1 = 19

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF CONTROL GROUP IN SHOT-PUT

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	25	27	2	4
2.	27	28	1	1
3.	20	21	1	1
4.	28	30	2	4
5.	31	31	0	0
6.	29	30	1	1
7.	35	35	0	0
8.	32	30	-2	4
9.	25	26	1	1
10.	29	32	3	9
11.	28	29	1	1
12.	31	32	1	1
13.	31	29	-2	4
14.	24	30	6	36
15.	27	28	1	1
16.	26	29	3	9
17.	28	30	2	4
18.	34	36	2	4
19.	28	26	-2	4
20.	28	28	0	0
	566	587	21	89

Mean Score of Initial Test	28.3
Mean Score of Retest	29.35
Sum of the Differences	21
Sum of Dif. Squared	89

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Shot Put GROUP Control

$N = 20$

$D = 21$

$D^2 = 89$

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \dfrac{S_D}{\sqrt{N}} =$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N-1} \qquad \frac{\sqrt{89 - \frac{(21)^2}{20}}}{19}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}} = .420$

$\bar{D}$ (Mean Difference) $= \dfrac{D}{N} = \dfrac{21}{20} = 1.05$

$t = \dfrac{\bar{D}}{S_{\bar{D}}} = \dfrac{1.05}{.420} = 2.50$

$df = N - 1 = 19$

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF CONTROL GROUP IN 600 YARD - RUN - WALK

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	106	125	-19	361
2.	104	108	- 4	16
3.	102	119	-17	289
4.	94	109	-15	225
5.	84	113	-29	841
6.	116	120	- 4	16
7.	94	98	- 4	16
8.	107	104	3	9
9.	103	128	-25	625
10.	102	109	- 7	49
11.	96	104	- 8	64
12.	106	114	- 8	64
13.	104	111	- 7	49
14.	101	116	-15	225
15.	106	119	-13	169
16.	113	103	10	100
17.	117	165	-48	2304
18.	97	113	-16	256
19.	102	116	-14	196
20.	103	117	-14	196
	2057	2311	-254	6070

Mean Score of Initial Test	102.85
Mean Score of Retest	115.55
Sum of the Differences	-254
Sum of Dif. Squared	6070

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST 600 Yard Run-Walk GROUP Control

N = 20

D = -254

D^2 = 6070

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\sqrt{\frac{D^2 - \frac{(D)^2}{N}}{N - 1}} \qquad \sqrt{\frac{6070 - \frac{(-254)^2}{20}}{19}}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}}$ = 2.736

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{-254}{20} =$ -12.70

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{-12.70}{2.736} =$ -4.642

df = N - 1 = 19

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN SIT-UPS

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	131	131	0	0
2.	131	131	0	0
3.	131	131	0	0
4.	131	131	0	0
5.	131	131	0	0
6.	131	131	0	0
7.	102	112	10	100
8.	100	63	-37	1369
9.	110	85	-25	625
10.	59	70	11	121
11.	68	90	22	484
12.	130	89	-41	1681
13.	80	82	2	4
14.	131	131	0	0
15.	127	131	4	16
16.	131	131	0	0
17.	127	131	4	16
18.	102	105	3	9
19.	90	131	41	1681
20.	131	131	0	0
	2274	2268	- 6	6106

Mean Score of Initial Test	113.70
Mean Score of Retest	113.40
Sum of the Differences	- 6
Sum of Dif. Squared	6106

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Sit-ups GROUP Experimental

N = 20

D = - 6

D^2 = 6106

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N - 1} \Big/ \sqrt{N} \qquad \frac{\sqrt{6106 - \frac{(-6)^2}{N}}}{19} \Big/ \sqrt{20}$$

$S_{\bar{D}}$ = 4.008

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{-6}{20} =$ -.3

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{-.3}{4.008} =$ -.075

df = N - 1 = 20

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN PULL-UPS

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	15	18	3	9
2.	14	23	9	81
3.	12	13	1	1
4.	13	18	5	25
5.	9	16	7	49
6.	12	18	6	36
7.	10	11	1	1
8.	4	6	2	4
9.	10	14	4	16
10.	11	13	2	4
11.	14	15	1	1
12.	12	10	-2	4
13.	3	9	6	36
14.	12	20	8	64
15.	7	8	1	1
16.	19	20	1	1
17.	6	8	2	4
18.	8	12	4	16
19.	11	7	-4	16
20.	8	6	-2	4
	210	265	55	373

Mean Score of Initial Test	10.50
Mean Score of Retest	13.25
Sum of Differences	55
Sum of Dif. Squared	373

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Pull-ups GROUP Experimental

$N = 20$

$D = 55$

$D^2 = 373$

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N - 1} \Big/ \sqrt{N}$$

$$\frac{\sqrt{373 - \frac{(55)^2}{20}}}{19} \Big/ \sqrt{20}$$

$S_{\bar{D}} = .764$

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{55}{20} = 2.75$

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{2.75}{.764} = 3.599$

$df = N - 1 = 19$

"t" at .01 level = 2.86

Significant at the .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN BROAD JUMP

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	82	92	10	100
2.	83	95	12	144
3.	93	101	8	64
4.	94	90	-4	16
5.	95	104	9	81
6.	90	98	8	64
7.	82	96	14	196
8.	86	94	8	64
9.	93	96	3	9
10.	100	98	-2	4
11.	107	106	-1	1
12.	96	101	5	25
13.	92	98	6	36
14.	81	90	9	81
15.	96	101	5	25
16.	89	99	10	100
17.	80	90	10	100
18.	95	86	-9	81
19.	91	89	-2	4
20.	79	88	9	81
	1804	1912	108	1276

Mean Score of Initial Test 90.20
Mean Score of Retest 95.60
Sum of Differences 108
Sum of Dif. Squared 1276

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Broad Jump GROUP Experimental

$N = 20$

$D = 108$

$D^2 = 1276$

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \dfrac{S_D}{\sqrt{N}}$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N-1} \qquad \frac{\sqrt{1276 - \frac{(108)^2}{20}}}{19}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}} = 1.35$

$\bar{D}$ (Mean Difference) $= \dfrac{D}{N} = \dfrac{108}{20} = 5.40$

$t = \dfrac{\bar{D}}{S_{\bar{D}}} = \dfrac{5.40}{1.35} = 4.0$

$df = N - 1 = 19$

"t" at .01 level = 2.86

Significant at the .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN SHUTTLE RUN

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	9.3	8.8	.5	.25
2.	9.6	9.0	.6	.36
3.	9.5	8.3	1.2	1.44
4.	10.0	9.3	.7	.49
5.	9.6	8.6	1.0	1.00
6.	9.6	8.8	.8	.64
7.	9.6	8.6	1.0	1.00
8.	9.9	8.7	1.2	1.44
9.	10.0	8.6	1.4	1.96
10.	8.5	8.1	.4	.16
11.	9.5	8.5	1.0	1.00
12.	8.9	8.3	.6	.36
13.	9.4	8.7	.7	.49
14.	10.0	8.9	1.1	1.21
15.	8.3	8.3	0	.00
16.	10.4	8.7	1.7	2.89
17.	9.7	8.5	1.2	1.44
18.	10.0	9.4	.6	.36
19.	9.2	9.4	- .2	.04
20.	9.8	9.8	0	.00
	190.8	175.3	15.5	16.53

Mean Score of Initial Test	9.54
Mean Score of Retest	8.765
Sum of Differences	15.50
Sum of Dif. Squared	16.53

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Shuttle Run GROUP Experimental

N = 20

D = 15.50

D^2 = 16.53

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{\sqrt{N}} \qquad \frac{\sqrt{16.53 - \frac{(15.50)^2}{20}}}{\sqrt{20}}$$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N - 1} \qquad \frac{\sqrt{16.53 - \frac{(15.50)^2}{20}}}{19}$$

$S_{\bar{D}}$ = .109

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{15.50}{20} =$ .775

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{.775}{.109} =$ 7.11

df = N - 1 = 19

"t" at .01 level = 2.86

Significant at the .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN 50-YARD DASH

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	6.4	6.4	0	0
2.	6.3	6.1	.2	.04
3.	6.2	5.8	.4	.16
4.	6.0	6.5	-.5	.25
5.	5.9	5.5	.4	.16
6.	6.0	5.9	.1	.01
7.	6.3	6.3	0	0
8.	6.1	5.9	.2	.04
9.	6.1	6.2	-.1	.01
10.	6.2	5.8	.4	.16
11.	5.8	6.0	-.2	.04
12.	6.4	6.1	.3	.09
13.	6.5	5.9	.6	.36
14.	6.4	6.3	.1	.01
15.	6.0	5.7	.3	.09
16.	6.9	6.5	.4	.04
17.	6.5	6.4	.1	.01
18.	6.3	6.2	.1	.01
19.	6.3	6.4	-.1	.01
20.	6.8	6.8	0	0
	125.4	122.7	2.7	1.49

Mean Score of Initial Test	6.27
Mean Score of Retest	6.135
Sum of Differences	2.7
Sum of Dif. Squared	1.49

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

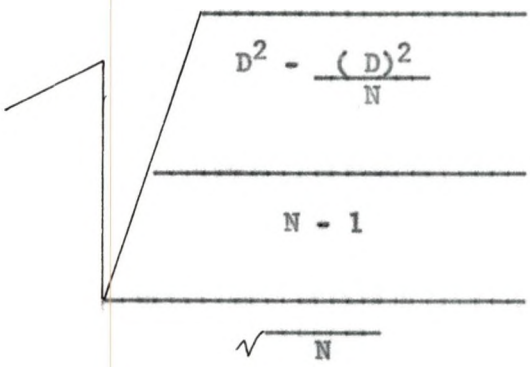
TEST 50 Yard Dash GROUP Experimental

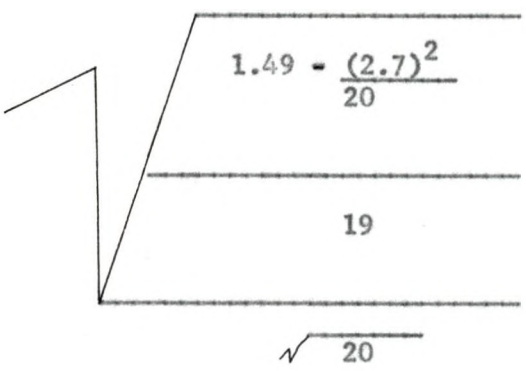
N = 20

D = 2.7

D^2 = 1.49

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$





$S_{\bar{D}}$ = .054

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{2.7}{20} =$ .135

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{.135}{.054} =$ 2.50

df = N - 1 = 19

"t" at .01 level = 2.86

Not significant at .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN SHOT PUT

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	25	29	4	16
2.	26	29	3	9
3.	33	37	4	16
4.	32	32	0	0
5.	42	49	7	49
6.	33	35	2	4
7.	34	38	4	16
8.	31	32	1	1
9.	27	27	0	0
10.	43	46	3	9
11.	29	33	4	16
12.	29	35	6	36
13.	38	43	5	25
14.	29	31	2	4
15.	34	38	4	16
16.	29	33	4	16
17.	30	33	3	9
18.	35	38	3	9
19.	27	30	3	9
20.	27	31	4	16
	633	699	66	276

Mean Score of Initial Test	31.65
Mean Score of Retest	34.95
Sum of Differences	68
Sum of Dif. Squared	276

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST Shot Put GROUP Experimental

$N = 20$

$D = 68$

$D^2 = 276$

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\frac{\sqrt{D^2 - \frac{(D)^2}{N}}}{N - 1} \qquad \frac{\sqrt{276 - \frac{(68)^2}{20}}}{19}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}} = .343$

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{68}{20} = 3.40$

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{3.40}{.343} = 9.913$

$df = N - 1 = 19$

"t" at .01 level = 2.86

Significant at the .01 level

INITIAL TEST AND RETEST OF EXPERIMENTAL GROUP IN 600 YARD-RUN-WALK

	Initial Test	Retest	Sum of Difference	Difference Squared
1.	98	92	6	36
2.	94	94	0	0
3.	103	104	-1	1
4.	99	94	5	25
5.	97	93	4	16
6.	96	95	1	1
7.	100	100	0	0
8.	98	94	4	16
9.	89	87	2	4
10.	106	103	3	9
11.	94	92	2	4
12.	99	92	7	49
13.	120	109	11	121
14.	107	102	5	25
15.	88	90	-2	4
16.	95	87	8	64
17.	96	100	-4	16
18.	103	102	1	1
19.	107	98	9	81
20.	98	98	0	0
	1987	1926	61	473

Mean Score of Initial Test	99.35
Mean Score of Retest	96.30
Sum of Differences	61
Sum of Dif. Squared	473

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM CORRELATED SCORES FROM SMALL SAMPLES

TEST 600 Yard Run-Walk GROUP Experimental

N = 20

D = 61

D^2 = 473

$S_{\bar{D}}$ (estimate of sampling error of $\bar{D}$) $= \frac{S_D}{\sqrt{N}} =$

$$\sqrt{\frac{D^2 - \frac{(D)^2}{N}}{N-1}} \qquad \sqrt{\frac{473 - \frac{(61)^2}{20}}{19}}$$

$$\sqrt{N} \qquad \sqrt{20}$$

$S_{\bar{D}}$ = .869

$\bar{D}$ (Mean Difference) $= \frac{D}{N} = \frac{61}{20} =$ 3.05

$t = \frac{\bar{D}}{S_{\bar{D}}} = \frac{3.05}{.869} =$ 3.51

df = N - 1 = 19

"t" at .01 level = 2.86

Significant at the .01 level

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM UNCORRELATED SCORES FROM SMALL SAMPLES

TEST: Sit-Ups

Experimental Group $\bar{D}$ = -.3 Control Group $\bar{D}$ = 31.25

Experimental Group $S_{\bar{D}}$ = 4.008 Control Group $S_{\bar{D}}$ = 7.20

$S_{D_M D}$ (the estimate of the sampling error for the distribution of differences between the mean differences)

$$\sqrt{S_{\bar{D}_1}^2 + S_{\bar{D}_2}^2} \qquad \sqrt{(4.008)^2 + (7.20)^2}$$

$S_{D_M D}$ = 8.24

$D_{\bar{D}} = \bar{D}_1 - \bar{D}_2 = (-.3) - (-31.25) =$ 30.95

$$t = \frac{D_{\bar{D}}}{S_{D_M D}} = \frac{30.95}{8.24} = 3.756$$

df = $(N_1 - 1) + (N_2 - 1) = 19 + 19 =$ 38

"t" at .01 level = 2.71

Significant at .01 level

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM UNCORRELATED SCORES FROM SMALL SAMPLES

TEST: Pull-Ups

Experimental Group $\bar{D}$ = 2.75 Control Group $\bar{D}$ = -.45

Experimental Group $S_{\bar{D}}$ = .764 Control Group $S_{\bar{D}}$ = .844

$S_{D_M D}$ (the estimate of the sampling error for the distribution of differences between the mean differences)

$$\sqrt{\left(S_{\bar{D}_1}\right)^2 + \left(S_{\bar{D}_2}\right)^2} \qquad \sqrt{(.764)^2 + (.844)^2}$$

$S_{D_M D}$ = 1.14

$\frac{D}{D}$ = $\bar{D}_1 - \bar{D}_2$ = 2.75 - (-.45) = 3.20

$$t = \frac{\frac{D}{D}}{S_{D_M D}} = \frac{3.20}{1.14} = 2.80$$

df = $(N_1 - 1) + (N_2 - 1)$ = 19 + 19 = 38

"t" at .01 level = 2.71

Significant at the .01 level

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM UNCORRELATED SCORES FROM SMALL SAMPLES

TEST: Shuttle Run

Experimental Group $\bar{D}$ = .775 Control Group $\bar{D}$ = .43

Experimental Group $S_{\bar{D}}$ = .109 Control Group $S_{\bar{D}}$ = .147

$S_{D_M D}$ (the estimate of the sampling error for the distribution of differences between the mean differences)

$$\sqrt{\left(S_{\bar{D}_1}\right)^2 + \left(S_{\bar{D}_2}\right)^2} \qquad \sqrt{(.109)^2 + (.147)^2}$$

$S_{D_M D}$ = .183

$\frac{D}{D}$ = $\bar{D}_1 - \bar{D}_2$ = .775 - .43 = .345

$$t = \frac{\frac{D}{D}}{S_{D_M D}} = \frac{.345}{.183} = 1.89$$

df = $(N_1 - 1) + (N_2 - 1)$ = 19 + 19 = 38

"t" at .01 level = 2.71

Not significant at the .01 level

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM UNCORRELATED SCORES FROM SMALL SAMPLES

TEST: 50-Yard Dash

Experimental Group $\bar{D}$ = .135 Control Group $\bar{D}$ = -.055

Experimental Group $S_{\bar{D}}$ = .054 Control Group $S_{\bar{D}}$ = .07

$S_{D_M D}$ (the estimate of the sampling error for the distribution of differences between the mean differences)

$$\sqrt{\left(S_{\bar{D}_1}\right)^2 + \left(S_{\bar{D}_2}\right)^2} \qquad \sqrt{(.054)^2 + (.07)^2}$$

$S_{D_M D}$ = .0883

$$\frac{D}{D} = \bar{D}_1 - \bar{D}_2 = .135 - (-.055) = .190$$

$$t = \frac{\frac{D}{D}}{S_{D_M D}} = \frac{.190}{.0883} = 2.15$$

$$df = (N_1 - 1) + (N_2 - 1) = 19 + 19 = 38$$

"t" at .01 level = 2.71

Not significant at the .01 level

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM UNCORRELATED SCORES FROM SMALL SAMPLES

TEST: Shot Put

Experimental Group $\overline{D}$ = 3.40 Control Group $\overline{D}$ = 1.05

Experimental Group $S_{\overline{D}}$ = .343 Control Group $S_{\overline{D}}$ = .420

$S_{D_M D}$ (the estimate of the sampling error for the distribution of differences between the mean differences)

$$\sqrt{\left(S_{\overline{D}_1}\right)^2 + \left(S_{\overline{D}_2}\right)^2} \qquad \sqrt{(.343)^2 + (.420)^2}$$

$S_{D_M D}$ = .542

$\underset{D}{D}$ = $\overline{D}_1 - \overline{D}_2$ = 3.40 - 1.05 = 2.35

$$t = \frac{\underset{D}{D}}{S_{D_M D}} = \frac{2.35}{.542} = 4.34$$

df = $(N_1 - 1) + (N_2 - 1)$ = 19 + 19 = 38

"t" at .01 level = 2.71

Significant at the .01 level

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

DERIVED FROM UNCORRELATED SCORES FROM SMALL SAMPLES

TEST: 600-Yard Run-Walk

Experimental Group $\overline{D}$ = 3.05 Control Group $\overline{D}$ = 12.70

Experimental Group $S_{\overline{D}}$ = .869 Control Group $S_{\overline{D}}$ = 2.736

$S_{D_M D}$ (the estimate of the sampling error for the distribution of differences between the mean differences)

$$\sqrt{\left(S_{\overline{D}_1}\right)^2 + \left(S_{\overline{D}_2}\right)^2} \qquad \sqrt{(.869)^2 + (2.736)^2}$$

$S_{D_M D}$ = 2.87

$\underline{D}_D = \overline{D}_1 - \overline{D}_2 = 3.05 - (-12.70) = 15.75$

$$t = \frac{\underline{D}_D}{S_{D_M D}} = \frac{15.75}{2.87} = 5.49$$

$df = (N_1 - 1) + (N_2 - 1) = 19 + 19 = 38$

"t" at .01 level = 2.71

Significant at .01 level

APPENDIX B

TABLE 4

RANK ORDER OF "t" FOR CONTROL GROUP

Area of Comparison	"t" Value
Shuttle Run	2.925
Shot-Put	2.50
Standing Broad Jump	-.06
Pull-Ups	-.533
50-Yard Dash	-.786
Sit-Ups	-4.34
600-Yard Run-Walk	-4.642

TABLE 5

RANK ORDER OF "t" FOR EXPERIMENTAL GROUP

Area of Comparison	"t" Value
Shot-Put	9.913
Shuttle Run	7.11
Standing Broad Jump	4.00
Pull-Ups	3.599
600-Yard Run-Walk	3.51
50-Yard Dash	2.50
Sit-Ups	-.075

TABLE 6

RANK ORDER OF "t" FOR BETWEEN GROUPS

Area of Comparison	"t" Value
600-Yard Run-Walk	5.49
Shot-Put	4.34
Sit-Ups	3.756
Standing Broad Jump	3.43
Pull-Ups	2.80
50-Yard Dash	2.15
Shuttle Run	1.89