Summer 1984

Acquisition of Physical Fitness Knowledge by Central Washington University Physical Education Activity Students

Diana Dawne Rector
Central Washington University

ACQUISITION OF PHYSICAL FITNESS KNOWLEDGE BY CENTRAL WASHINGTON UNIVERSITY PHYSICAL EDUCATION ACTIVITY STUDENTS

A Thesis

Presented to

The Graduate Faculty

Central Washington University

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

---

by

Diana Dawne Rector

July, 1984

APPROVED FOR THE GRADUATE FACULTY

_____

Robert N. Irving Jr., COMMITTEE CHAIRMAN

_____

John A. Green

_____

James G. Nylander

ACQUISITION OF PHYSICAL FITNESS KNOWLEDGE BY CENTRAL WASHINGTON UNIVERSITY PHYSICAL EDUCATION ACTIVITY STUDENTS

by

Diana Dawne Rector

July, 1984

The purpose of this study was to determine what effect participation in activity classes has on the acquisition of physical fitness knowledge. No significant difference was found between activity and non-activity students. The performance of Central Washington University students did not differ significantly from national percentile norms. Certain topics of physical fitness knowledge were found to be significantly more difficult than others. A renewed effort in the construction and use of standardized knowledge tests in physical education was recommended.

ACKNOWLEDGEMENTS

The writer wishes to express her appreciation and gratitude to the following people:

To Dr. Robert Irving, Jr., committee chairman and originator of this project, whose dedication to the field of physical education was a source of inspiration.

To Dr. James Nylander and Dr. John Green for their assistance as members of the committee.

To my mom, for her endless support, encouragement, and love.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Chapter 1

INTRODUCTION

Many authorities have questioned the inclusion of physical education in public school education. Safrit captured the essence of this growing concern when she stated, "As the process of education becomes increasingly expensive, the public continues to question the cost of various educational programs, and to demand justification of their inclusion in the school curriculum" (39:5). Sizer, a former dean of Harvard's School of Education, proposed a reduced curriculum eliminating physical education because "much of what happens in schools today under that rubric is neither education . . . nor very physical" (55). Physical educators need to prepare a strategy to promote the worth of their program.

An examination of the traditional objectives of physical education, the reasons for its inclusion in the educational system initially, may provide the desired justification for its presence in schools today. These objectives were clarified by Hetherington, a pioneer of physical education, and summarized by Nash in 1948 as "organic power, neuromuscular development leading to skill

in performance, the ability to do interpretive thinking, and the guiding of the emotional urges" (31:59). The four objectives are interdependent, the success of each level of development depends on the existence of the ones below (see figure 1). The degree to which a physical education program satisfies these objectives can be measured by the assessment of student improvement in physical fitness, skill, knowledge, and attitude.

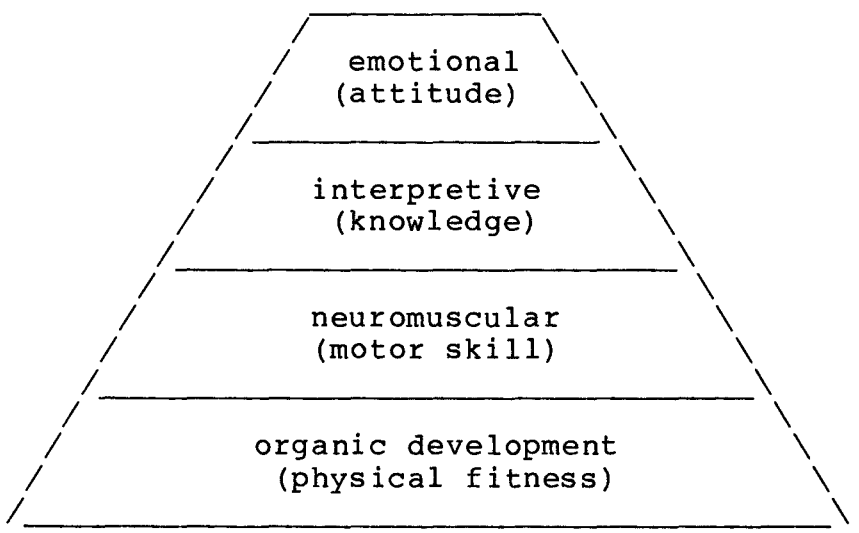

Figure 1
Levels of Development

The specific objective of "interpretive development", or simply the acquisition of knowledge, is a very critical area of consideration in light of the increasing emphasis on academia. "The physical educator is concerned not only with the development of motor skills

and abilities, but also the assimilation of knowledges" (39:166). This is evidenced by the amount of research performed in the past half-century to standardize knowledge tests for physical education activities. A good program of physical education can provide the opportunity for the students to increase their knowledge of important facts about physical activity (1:10), thus establishing a place for it in the academic world.

The popularity of physical fitness, an integral part of the physical education program, has mushroomed to nuclear proportions in the past decade. The importance of acquiring physical fitness knowledge places it as a primary concern and area of focus for physical educators to address in their classes.

This study was part of a joint effort to determine the effect of physical education activity classes at Central Washington University on the acquisition of physical fitness, motor skills, knowledge, and attitudes by the students enrolled. Mr. Jeff Zenisek, Mr. Dale Ehler, and Mr. Tim Clark researched the objectives related to fitness, skill, and attitudes, respectively. This researcher examined what effect activity classes had on the acquisition of physical fitness knowledge. Each of the four studies covered one of the four traditional objectives, using the same research subjects, during the same period of time.

Need for the Study

The objective of this study was to determine what effect physical education activity classes at Central Washington University have on the acquisition of physical fitness knowledge. The assessment of acquired knowledge has many beneficial purposes including student and teacher motivation (19:506), the evaluation of a program or a teacher's strengths and weaknesses (20:xi), an objective measure of student progress for grading (47:78), and a method of ability grouping or classification (16:74). For these reasons, and the fact that no study of this kind has been conducted in the physical education department at Central Washington University, the need was established.

Statement of the Problem

The problem studied in this paper was to determine the contribution that physical education activity classes make to the physical fitness knowledge of the students enrolled in the classes. Sub-problems considered in this research included the following:

1. To determine the general level of physical fitness knowledge possessed by Central Washington University students.

2. To determine the acquisition of physical fitness knowledge by students enrolled in four categories

of activity classes: fitness, team sports, individual/dual sports, and rhythm/dance.

3. To determine the specific topics of physical fitness knowledge with high or low levels of difficulty.

Delimitations

1. The study included a random sample of students in Tuesday-Thursday activity classes only.

2. The study included only college students at Central Washington University.

Limitations

1. All subjects did not complete the knowledge test under the same physical conditions.

2. An error was found in the reproduction of Form B of the Test of Physical Fitness Knowledge. Six items, fifty-eight through sixty-three, were inadvertently placed out of sequence. A high correlation (.984) was found between the scores of the subjects on form B with or without those six items. The items were not used in the analysis of topic difficulty.

3. The study included only six control group subjects at the end of the quarter.

Definitions

The following terms were used in this paper as

defined below:

Knowledge. "Those behaviors and test situations which emphasize the remembering either by recognition or recall, of ideas, materials, or phenomena" (46:21).

Physical Fitness. "The ability to carry out daily tasks with vigor and alertness, without undue fatigue, and with ample energy to enjoy leisure-time pursuits and to meet unusual situations and unforeseen emergencies" (5:12).

Hypotheses

The following hypotheses were investigated in this study:

1. There will be no significant difference in the physical fitness knowledge of college students enrolled in physical education activity classes from the beginning of the quarter to the end.

2. There will be no significant difference in the physical fitness knowledge of college students enrolled in physical education activity classes and those in the control group, not enrolled in activity classes.

3. There will be no significant difference between the physical fitness knowledge of students enrolled in any of the four categories of physical education classes (fitness, teams sports, individual/dual sports, rhythm).

4. There will be no significant difference between the physical fitness knowledge of Central Washington University students and the population sample from whom norms were constructed.

5. There will be no significant difference in the percentage of correct responses among the ten specific topics of physical fitness knowledge.

Chapter 2

REVIEW OF LITERATURE

To determine the amount of knowledge gained through physical education activity classes, it was necessary to select a reliable, valid, and objective evaluative tool that would effectively test knowledge applicable to any activity. The review of literature revealed the existence of dozens of standardized knowledge tests for physical education activities, ranging from archery to wrestling. Following will be an examination of several of the studies that were undertaken to develop knowledge tests in physical education over the past fifty-five years. Additionally, a thorough discussion of the two existing tests of physical fitness knowledge is included in an effort to focus on the selection of the measuring device in this study. The tests are broken into the following categories: (1) Team Sports, (2) Individual/dual Sports, (3) Miscellaneous Activities, (4) Tests of Multiple Activities, and (5) Physical Fitness/Health Tests.

Tests of Team Sports

Bradley (3) standardized a baseball knowledge test for male physical education majors in 1960. The curricular validity of the test was established by analysis of baseball textbooks and articles, and the judgements of thirty-three baseball instructors. Two preliminary forms of the test were administered to 759 men in twenty-seven colleges. Each test was comprised of eighty-seven four-option multiple choice questions. From those, eighty items that met selection criteria were used to develop the final test, which was administered to 1001 men. The resulting reliability coefficient was .847. The test was proven to be valid, reliable, and objective.

Two knowledge tests for basketball were reviewed. First, Schwartz (41), in 1937, developed knowledge and skills tests for high school girls. The 100 item test was administered to 1000 subjects. Schwartz's goal, in part, was to develop a standardized test so a teacher could identify instruction weaknesses, and pupils could be motivated to a greater degree. The range of scores for the test was thirty-three to 100, with a mean score of 69.9, and standard deviation of 9.48. She concluded that the test could be used as a reliable measure of basketball knowledge despite the fact that there were not enough items of average difficulty.

McCutcheon (25) constructed a basketball knowledge examination for college women in 1966. Fifty-four multiple choice questions made up the test, which was given to 303 subjects at four colleges. Item analysis, by Flanagan's method of using the upper and lower 29 percent of items with double weight to the extreme 9 percent, revealed a mean difficulty of 67 percent. The scores ranged from twenty to forty-five, with a mean score of thirty-three, and standard deviation of 8.13. McCutchoen utilized the Kuder-Richardson formula to determine the reliability coefficient of .852.

One test of field hockey knowledge was reviewed. Kelly and Brown (21) constructed the test for female physical education majors in 1952. The test items were selected after an analysis of eleven texts, in order to establish curricular validity. The test was developed in three stages: initial, pre-test, and revised. The initial test included 103 questions and was submitted for revision to a committee of experts. The resulting test of 116 items, the pre-test, was administered to thirty subjects. The revised test was created after item analysis was performed on the pre-test. The final form included 106 items and was administered to 209 women. The degree of difficulty ranged from five to 100 percent, but only four items fell outside the desired range of 10 to 90 percent. Flanagan's indices of discrimination were found

to be between -.23 and .63, with a mean index of .35. Eighteen items fell below the acceptable level of .19. Elimination of these items raised the mean index to .40 without affecting the curricular validity. Validity was established through the comparison of the scores of experts, majors, activity students, and lay persons; the correlation of test scores with field hockey experience; and the correlation of test scores with a rating of major student's field hockey competence. The Pearson product moment correlation was utilized, with the result being a coefficient of .60 between the written test and extent of experience, and .64 between the written test and instructors ratings. Test reliability of .89 was determined by split-halves, even versus odd, then corrected to .94 by the Spearman-Brown prophecy formula.

A knowledge test was discovered for touch football. Power (35) designed the test for junior high boys in 1960.

Two tests of lacrosse knowledge were reviewed. The first, by Hodges (18) in 1968, was designed for college women as part of a study to measure achievement in knowledge and skills of beginning lacrosse students. The forty-item multiple choice test was analyzed for difficulty and discrimination. All items had high indices of discrimination and adequate degrees of difficulty. The reliability coefficient was .68.

Warren's (53) test, constructed in 1972, also was designed as a measure of knowledge of women in beginning lacrosse classes, but at the high school level. The original form of the test had forty-six items and was given to 254 students. After item analysis, fourteen questions were omitted and the final form of thirty-two items was statistically analyzed. Though the reliability coefficient was low, .62, the test was determined to be an adequate measure of lacrosse knowledge if a specified course of instruction is followed.

A soccer test by Knighton (23), in 1930, was perhaps one of the first objective knowledge tests published. The test was formulated to give a better understanding of the game to beginners. The test items used were true-false and multiple choice. Though the test was not statistically analyzed, its importance as a model of one of the first attempts to measure knowledge of a physical education activity is significant.

Waglow and Stephens (51) constructed a softball knowledge test in 1955. The initial form of the test contained 100 questions and was first revised by a panel of experts. The first revision was administered to 116 students, then subjected to item analysis. Any item that fell outside the 10 to 90 percent range of difficulty, or was below the .20 index of discrimination, was rewritten. The second revised version was administered to 115

students. Reliability was calculated for both revisions, using odd versus even questions corrected to the whole test by the Spearman-Brown formula. The coefficient for the first revision was .71, and was corrected to .83. The second revision had a coefficient of .64, corrected to .78.

A speed-a-way knowledge test was completed by Palmer (32) in 1964. The knowledge test was administered to 279 women enrolled in fundamentals classes. Palmer analyzed the test using Flanagan's test for item validity, as well as tests for difficulty. The test was determined to be a satisfactory measure of speed-a-way knowledge.

Finally, a study by Rodgers (37) in 1939, described a procedure for the standardization and use of objective information tests for team sports. Specifically, Rodgers included a copy of a softball test that consisted of 100 true-false questions. The procedure of developing a test was the most significant part of this study, as the researcher illustrated each step of test construction with examples from her creation of a soccer test. Rodgers emphasized the need for objective knowledge tests as a means to eliminate personal bias in evaluation, for teachers to know how well they have taught, and to motivate teachers and pupils to a more meaningful physical education experience.

Tests of Individual/dual Sports

Blaess (2), in 1966, developed an archery knowledge test. She determined objectivity by the subjective evaluation of a jury of experts. Seventy-seven items, using Flanagan's index of discrimination, were validated, while reliability was found satisfactory using the Kuder-Richardson formula.

Three tests were created between 1941 and 1953 to test badminton knowledge. Scott (44) was first with a written and skills test battery in 1941. The initial form of the test had eighty-nine items, but was reduced to eighty because nine items had poor levels of difficulty. The final form contained forty-seven multiple choice and thirty-three true-false questions, and had reliability coefficients of .79 and .72 for the multiple choice and true-false questions respectively. Scott concluded that the written test was an adequate measure of badminton knowledge.

Phillips (34), in 1946, created a standardized badminton knowledge test for college women. She established curricular validity using analysis of the contents of courses of study, textbooks, and competent person's judgement. Nearly 1000 copies of the original test were sent to seventeen universities, with 648 returned. Reliability of the initial test form was

calculated at .879, with a range of scores from seventeen to 137. The test was reduced to 100 questions, forty-five multiple choice and fifty-five true-false. Phillips sent out copies of the final test and received 1471 completed from thirty colleges. Reliability of the test was established with the Kuder-Richardson formula (r=.921) and adequate levels of difficulty (7 to 93 percent). The test was determined to be valid, reliable, and objective.

Finally, Fox (10) devised a beginning badminton knowledge examination in 1953 to standardize and improve teaching and grading. The initial test contained ninety items and was given to eighty-nine beginners. The reliability, validity, and difficulty of each item was computed and the test revised to 107 items. This form was administered to 269 students at the end of the badminton course, then analyzed. The arithmetic mean of the scores was 25.05, with a standard deviation of 10.82. The reliability of the test was improved from .78 in the initial form, to .90 for the final test. Item validity revealed that 89.6 percent of the items discriminated at a high level, and item difficulty ranged from 2 to 69.67 percent.

Hardin (13), in 1960, constructed a knowledge test for college students enrolled in bowling classes. This test contained sixty alternate response and sixty multiple choice questions, and was administered to 446 students in

bowling classes at three colleges in Texas. The final examination was developed and was determined to adequately measure knowledge and understanding of bowling skills.

Two studies were reviewed of tests to measure golf knowledge. Murphy (29), in 1933, prepared a golf knowledge test to be used for the testing and grading of women in required freshman and sophomore classes at a Texas college for women. The items for the test were chosen from recognized texts in the field and based on their frequency of emphasis in those texts. Reliability of the test was found by self-correlation of the whole and half-test. The coefficient for the half-test was established at .76, and was .86 for the whole test. The study included the development of grading standards based on standard deviation, mean, and the normal probability curve.

Waglow and Rehling (50) created a golf knowledge test twenty years later. They used 100 true-false questions, with the items taken from prominent books to establish validity. The reliability of the test was determined using split-halves, odd versus even, for a coefficient of .82. Item analysis was performed and identified a range of difficulty from 17 to 100 percent, as well as indices of discrimination above the .19 level for eighty-five of the 100 items. Based on 100 men tested, the mean score was 56.14, with a range from

fourteen to eighty-one.

Four individuals undertook studies to develop tennis knowledge tests. Hewitt's test in 1937 had been a widely used test which he revised in 1964 by popular request (16,17). Hewitt emphasized a need to measure knowledge of an activity because knowledge increases the student's understanding and appreciation of an activity, and assists the instructor in the classification of students into instructional groups. Hewitt's initial test contained 200 questions and was given to 104 men and women enrolled in college tennis classes. Items were analyzed and the test was revised to 100 questions. The revised form of the 1937 test was administered to 105 students, then analyzed. Reliability was calculated by split-halves, even versus odd (r=.90), and for the whole test using the Spearman-Brown formula (r=.947). Also, as a check for reliability, the Minnesota Tennis Knowledge Test was given to fifty-four of the same subjects and the scores correlated with the Pearson product moment correlation (r=.808). Twenty-seven years later Hewitt updated the test, then compared it again to the Minnesota test (r=.81) and to Scott's Achievement Tennis Examination (r=.86). The reliability of the updated version was measured at .95 using the Spearman-Brown formula.

In 1941, Scott (43) developed achievement examinations for tennis classes, elementary and

intermediate levels. The final forms of these tests were used by Hewitt to compare to his updated version (17). Scott used true-false and multiple choice items and tested for difficulty. The final battery of items had reliability coefficients of .87 and .78 for the elementary and intermediate tests respectively.

Broer and Miller (4) also created achievement tests for beginning and intermediate tennis. The original tests of 100 questions were administered to eighty-seven subjects. The reliability coefficient, calculated by split-halves and corrected by the Spearman-Brown formula, was .84. The curricular validity was established as everything included was in the curriculum, and test item emphasis was based on the course of studies as well as analysis of available tests and textbooks. The initial test was revised to 128 items and was administered to 297 beginning and forty-six intermediate students. The reliability of the test was found to be .82 for beginners, .42 for intermediates, and .86 for combined scores.

In 1953, Miller (26) developed a battery of tests to measure achievement levels in tennis knowledge and skill for women physical education majors. The 326 item knowledge test was first taken by 381 subjects from twenty-seven colleges. Scores ranged from sixteen to 202, with a mean score of 110.25. The reliability coefficient of .93 was found with the use of the Kuder-Richardson

formula. Item validity was determined by the Votaw curve after which all invalid items were eliminated. The difficulty of the items ranged from 16.7 to 86.3 percent. The test was revised to a 100 item test and was administered to 612 subjects from forty-five institutions. The reliability coefficient for the final form was .900.

One test of track and field was reviewed. Parks (33) devised the test in 1960.

Finally, Kraft (24) constructed and standardized a wrestling knowledge test for male physical education majors. Two pre-test try-out forms of seventy-five items each were administered to majors at Western Michigan University. Two fifty-item tests were created from the initial forms and administered to 723 male physical education majors at twenty-one institutions. The final test of fifty questions was given to 339 majors in 1970, then analyzed for reliability and validity. The reliability coefficient was found by using the Kuder-Richardson formula (r=.85) and the Spearman-Brown reliability formula (r=.87). Item difficulty fell between .13 and .53, assuring that the final test form was valid and reliable.

Tests of Miscellaneous Activities

A limited number of knowledge tests have been created for activities other than team sports or

individual/dual sports. Following will be a review of five miscellaneous tests: boating, swimming, gymnastics, folk dance, and general physical education knowledge.

Reichle (36) constructed a test to measure knowledge in recreational boating in 1975. Two forms of the final test were created with fifty questions each. The two parallel forms were determined to be reliable and valid, and could be used to evaluate boating knowledge. Item analysis revealed that the two forms were indeed parallel. The mean difficulty of form A was .43 and for form B .41, with the mean index of discrimination of .32.

Shambaugh (45) wrote and analyzed an objective test of folk dance knowledge in 1935. The test was administered to a recreational group of students and to a group of physical education majors. The mean score of the recreational group was 55.6, and for the major group 64.5. The examination was tested for reliability, and achieved a .88 coefficient.

In 1957, Gershon (12) developed a test for college men that covered gymnastics knowledge. An experimental test of 150 questions was first created. The mean score was 81.45, with a standard deviation of 10.90. The reliability coefficient, using the Kuder-Richardson formula, was .69 before the true-false questions were corrected for guessing. After correcting, the test earned a .88 reliability coefficient. Seventy-five of the 150

items were valid according to Flanagan and Votaw procedures, and all ranged in difficulty from one to 96 percent. The final version of the test had 100 questions, with difficulty from eight to 91 percent, a mean score of 56.35, and a reliability coefficient of .92.

Scott (42) wrote and analyzed an achievement examination for swimming. She indicated three purposes to test knowledge in physical education: to measure levels of student understanding, to classify students, and to aid in grading. Scott hoped that the creation of written tests would "encourage the practice of expecting an increase in knowledge and understanding as well as the development of skill in college classes" (42:40). Scott sent out questionnaires to forty-two colleges in order to identify commonly taught materials, and then based test items on the resulting information. She used true-false and multiple choice questions, and computed reliability by correlating odd versus even questions, corrected to actual length by the Spearman-Brown formula. The coefficient of reliability for the tests were .888 and .867.

Finally, Walker (52) developed a general knowledge inventory test in physical education for college freshmen in 1950. His purposes were to determine the weaknesses of entering freshmen in general physical education knowledge and develop a resource syllabus for a foundation course. The test items were created from the following categories:

physiological principles, kinesiological principles, safety-first aid, sport knowledge, motor learning, philosophy, and objectives. Walker constructed two forms of the test, each with 100 items, and administered them to 550 first trimester freshmen at Florida State University. He analyzed the results for variability, central tendancy, reliability, item analysis, and equivalency. The reliability coefficient was found to be .637 for form A, and .662 for form B.

Tests of Multiple Activities

Shortly after the first physical education knowledge tests appeared, several researchers undertook the challenge to create a package of tests. The studies that follow were completed to standardize tests that contain information on more than one activity, or contain a group of individual tests that together measure the knowledge of several physical education activities.

In 1935, Snell (47) completed and tested a series of ten knowledge tests including the team sports of volleyball, field hockey, soccer, basketball, and baseball; individual/dual sports of archery, tennis, and golf; and horseback riding and hygiene. The tests were published in Research Quarterly in three parts from October 1935 to May 1936. Snell stated two major reasons for creating the tests: to serve as a pre-test to

classify students, and as a post-test for grading. The knowledge test constituted only one part of a group of four tests involving knowledge, posture, attitude, and motor ability, and was used in an experiment conducted to measure the effects of physical education on activity students at the University of Minnesota. Each constructed knowledge test originally contained seventy questions and was administered to 800 women. The number of questions for each test was then reduced to forty-five. Validity was established on the basis of expert opinion. Each test was correlated by split-halves, using even versus odd, then corrected to the whole test by the Spearman-Brown formula. All tests were found to be valid, objective, and reliable measures of knowledge for the specific activity.

Hemphill (14) created a group of knowledge tests in 1932, in an effort to develop and standardize test forms covering methods and techniques of the various phases of his physical education program. The tests were grouped into five categories: major sports (baseball, football, basketball), minor sports (volleyball, soccer, tennis, handball), health, self-defense, and recreational sports. Test items were taken from texts written by leading authorities and from established tests, such as Gates-Strang Health Knowledge Test. The items were formed into true-false and multiple choice questions for objectivity. Reliability was established with correlation

of correct scores on odd-numbered questions with correct scores on even-numbered questions, then corrected by the Spearman-Brown formula. The average reliability coefficient was .773. Individual tests scored as follows: baseball .773, football .780, basketball .666, self-defense .877, health .808, minor sports .847, and recreational sports .730. The tests were considered valid and reliable, and could be used with confidence.

Hennis (15), in 1956, constructed seven knowledge tests, three for individual/dual sports (badminton, bowling, tennis) and four for team sports (volleyball, softball, field hockey, basketball). Hennis used four-choice multiple choice questions exclusively because of the adaptability of this type of question to most physical education subject matter, the objective scoring, the adaptable use of answer sheets, and ability to test a student's ability to eliminate incorrect answers as well as choose correct answers (15:302). The initial form of each test contained from forty to fifty-nine items and was administered to various numbers of people ranging in group size from 208 for the softball test, to 2291 students for the tennis test. Each test was analyzed and poor items were eliminated, thereby reducing each test to approximately thirty-five questions. Estimates of reliability were calculated by Augoff's Equation C, with softball gaining the highest coefficient of .81, and

badminton the lowest at .72. The indices of discrimination for all test items exceeded the acceptable level of .19.

In 1943, French (11) constructed knowledge tests in sixteen different physical education activities. The reliability coefficients of the tests ranged from .702 to .884, with eleven of the sixteen tests above .800 and five above .850. The tests were developed to test knowledge of women majors in the selected activities.

Hooks (19) developed four knowledge tests for college men with three purposes: motivation, classification, and evaluation. The tests were administered to 2832, 3513, 2740, and 4140 subjects in badminton, softball, tennis, and volleyball respectively. The level of difficulty for all four tests ranged from 7.9 to 17.2, and the indices of discrimination were in the acceptable range of .20 to .79. The reliability was determined by split-halves, even versus odd, and corrected by the Spearman-Brown formula. The coefficient for the four tests were .85 for badminton, .77 for softball, .81 for tennis, and .73 for the volleyball test. The tests were considered valid and reliable measures of achievement and understanding in all four areas of knowledge.

Farrow (8) developed objective, multiple-choice knowledge tests in four different individual/dual sports in 1972. Tests in the areas of archery, bowling, golf,

and tennis, were constructed and analyzed. The content emphasis of each test was based on the balance of content taken from inventories completed by instructors of each activity. The indices of discrimination for all items in all four tests were above the minimal .20 level, and the level of difficulty was in the acceptable 10 to 90 percent range. Reliability coefficients calculated by the Kuder-Richardson formula ranged from .85 to .91.

Wilson (54) directed a study at the University of Washington in 1969 that utilized an existing knowledge test. Wilson's objective was to evaluate instruments used to assess basketball knowledge and performance, swimming knowledge and performance, and gymnastics free exercise performance. All subjects used in the study were women activity students. To evaluate basketball assessment methods, she gave a previously developed, unpublished knowledge test and three skill tests to twenty-nine students. At the end of the quarter, three experienced judges used an established rating scale to rate the students, and a film evaluator rated the students from videotaped games. The sum of scores on the three skills tests and knowledge were correlated with the film evaluation. Reliability was predicted with the Spearman-Brown formula. There existed substantial relationship between the film evaluation and the knowledge test to conclude that the knowledge test could be used as

the single evaluative tool. Wilson also concluded that the correlation between the knowledge test and skill tests indicated no significant relationship to indicate that one could be used to determine the other. The swimming knowledge test was unique in that it contained only seven short answer type questions. A previous study determined that those seven questions could adequately measure swimming knowledge.

Tests of Physical Fitness/Health

A study by Murphy (28) in 1937 most closely resembled this writer's research. Murphy identified the gain in health knowledge of women in physical education activity classes. She made a statistical comparison of health knowledge between three groups of students: (1) women enrolled in physical education classes with no outside reading or discussion sessions, (2) women enrolled in physical education activity classes with reading and discussion groups, and (3) entering freshmen with no college class background. Murphy administered a standardized health knowledge test as the measuring device. In the first of two years of testing, 1005 women were given two tests, 188 items total. The second year, a pre-test was given in the fall and a post-test in the spring using the same test. Murphy found that the group of women enrolled in the physical education activity

classes that included reading and discussion groups showed a 11.17 percent higher gain than those in classes without the reading and discussion groups. The group of freshmen with no college background made the smallest gain.

Dearborn (7) administered his previously standardized knowledge test, the Dearborn College Health Knowledge Test, as a pre-test before instruction to determine the general level of health knowledge possessed by college freshmen. Twelve thousand freshmen were tested at fifteen colleges over a six-year period. A 25 percent sample, 3,000 cases, was selected for the study. The test included among other topics: nutrition and diet, exercise and body mechanics, fatigue and rest, and disease prevention and control. The range of difficulty of the questions was five to 89 percent, with a range of scores from thirteen to ninety-five. The results were analyzed in respect to the students type of college, four-year or junior college. The students from four-year colleges missed 37 percent of the questions, while those from junior colleges missed 65 percent. The mean score for four-year college enrollees was 54.4, and for junior college students 44.5. Dearborn indicated a need for greater emphasis in health instruction at the high school level and below.

One of two tests of physical fitness knowledge was constructed by Stradtman and Cureton (48) of the

University of Illinois in 1950. The two men felt that the emphasis on physical fitness had developed the need for a test that would determine the amount of knowledge of desirable practices in physical fitness possessed by the students (48:53). Stradtman and Cureton performed a review of literature prior to the construction of the test and found no other test of its kind. With a primary concern for validity, they submitted a preliminary test of 148 items to a group of experts in the fields of physical education, physical fitness, and physiology. The final form of the test contained 100 items validated by reference to reliable books and magazines and the judgements of experts. The men wrote the test at the sixth grade level to assure that the test would not be a test of vocabulary knowledge. The final test form was administered to 333 high school juniors and seniors, 153 boys and 180 girls. The reliability was determined by split-halves, then the Spearman-Brown formula, resulting in a highly favorable coefficient of .95. The test was considered valid and reliable. Finney and Daughterty (6,9) utilized this test in 1969 as a part of their examination of the physical education programs of selected Central Washington high schools.

The test selected for use in this study was developed by Mood (27) in 1970. Mood constructed and standardized two practical, valid, and reliable forms of a

knowledge test for physical fitness as an instrument to evaluate the undergraduate physical education curriculum. The test items were created after a thorough investigation of current written materials on physical fitness. The information was categorized into ten broad topic areas with the percentage of emphasis as follows:

1. Current status and promotion of physical fitness - 5%
2. Evaluation of physical fitness ---------- 10%
3. Kinesiological aspects of physical fitness ------- 10%
4. Nutritional aspects of physical fitness ---------- 10%
5. Physical fitness programs ---------- 12%
6. Physical fitness versus disease ---------- 5%
7. Physiological aspects of physical fitness -------- 20%
8. Psychological aspects of physical fitness -------- 10%
9. Sociological aspects of physical fitness --------- 8%
10. Miscellaneous concepts ---------- 10%

Three or more multiple choice items were created for each of the sixty facts selected from the topics, then submitted to the University of Iowa's faculty for correction and relevance. The items were revised and rewritten leaving 184 items for which four try-out test forms of forty-six items each were formed.

Over 1800 tests were administered. The 695 tests taken by senior physical education major students were used for item analysis. The index of difficulty averaged 52 percent, and the index of discrimination averaged .31 according to Flanagan's formula. From the try-out results, two sixty-item test were created. The two final forms were constructed by selecting and editing two of the

experimental items for each of the original sixty physical fitness facts. One of each of the pairs of items was placed into one of two piles, and interchanged until the level of difficulty and index of discrimination of both piles were equal.

A total of 11,949 copies of the final forms of the test were mailed to 255 institutions. At the time of analysis, 4167 tests had been returned properly completed. Mood tested for difficulty and found that forty items in form A and forty-five of the items in form B had middle difficulty.

Content validity was established from the steps taken during test construction. Construct validity was insured as the mean score of the physical education majors was significantly higher than non-majors. Reliability was determined by split-halves corrected by the Spearman-Brown prophecy formula and the Kuder-Richardson formula. Form A of the test was found to have a coefficient of .74, form B was calculated at .77 and .75.

Mood converted the raw scores to T-scores, which converted the raw score distribution into a normal distribution. This method was used because "measurement of growth was one the the major purposes for constructing two forms of the Test of Physical Fitness Knowledge" (40:60). The T-scores were based on the scores of the 2226 answer sheets completed by physical education majors,

the group the test was designed for. The T-scores were translated into percentile ranks, then set up into fifteen different norm groups depending on the year in college, sex, and major of the students tested.

Mood concluded that there was evidence for content, concurrent, and construct validity sufficient to indicate that the test measured physical fitness knowledge. The reliability coefficients obtained for the final forms are such that the test can be used with confidence to discriminate accurately between performance groups. Finally, the two final forms of the test were found to be essentially parallel.

Chapter 3

PROCEDURES

The purpose of this study was to determine the acquisition of physical fitness knowledge by Central Washington University physical education activity students. The subjects that were evaluated consisted of a group of randomly selected students from Tuesday-Thursday activity classes and a group of students not enrolled in a physical education activity class during Winter, 1984.

Method

The experimental randomized groups pre-test post-test design was used in this study. An experimental group was drawn from Central Washington University physical education activity classes, and a control group from the general student body not enrolled in an activity class. Both groups were administered a pre-test at the beginning of the quarter, and a post-test at the end. Approximately eight weeks elapsed between testing, which allowed for sixteen, forty-minute teaching sessions in the physical education activities for the experimental subjects.

Sampling Procedure

Experimental Group

The proportional stratified random sample technique was used to select students from Tuesday-Thursday physical education activity classes for the experimental group. Subjects were drawn from four of the five categories of classes offered at Central Washington University: (1) physical education fitness (PEF), (2) physical education team sports (PETS), (3) physical education individual/ dual (PEID), and physical education rhythm (PER). Physical education aquatics classes, the fifth category of activity classes, were not held during Winter, 1984 due to the rennovation of the college swimming pool.

During Winter Quarter, 1984, there were 795 students enrolled in Tuesday-Thursday activity classes: 33.7 percent in PEF, 27.4 percent in PEID, 20.5 percent in PETS, and 18.3 percent in PER. An arbitrary decision to choose a sixteen percent sample required that 127 students be selected. The subjects were drawn from the four categories of classes in proportion to the total percent enrolled in each. Forty-three subjects were chosen from PEF classes (33.8 percent), thirty-five from PEID classes (27.5 percent), twenty-six from PETS classes (20.4 percent), and twenty-three from PER classes (18.1

percent).

After the sample size was established, it was necessary to determine a method of selecting subjects from each of the four groups in a random manner. The subjects were selected on the basis of the last two numbers in their social security numbers, which are assigned randomly by the federal government. Individuals who had 00 as the last two numbers were chosen first, then 01, 02, and so forth until the targeted number of subjects were accumulated for each category.

Control Group

The systematic method of sampling was used to gather individuals for the control group. An alphabetized list of all students enrolled at Central Washington University Winter, 1984 was acquired. Every tenth student was selected to be contacted. It was necessary to determine first, whether or not the student was enrolled in a physical education activity class, in which case the student could not be in the control group. Secondly, it was necessary to determine whether the individual would participate in the study. Only three of the first 120 called were willing to participate. Twenty additional subjects were recruited for the control group at random from the Samuelson Union Building at Central Washington University.

Administration of the Test

Mood's Test of Physical Fitness Knowledge (39,40) was administered to each subject in this study. Each form of the test included sixty, four-choice multiple choice items. The knowledge test was one of four tests (knowledge, attitude, fitness, and skill) given within the same time frame.

The first testing session, at the beginning of the quarter, occured on the Saturday following the add-drop deadline. Form A was given to all participants in the study as the pre-test. The subjects were given written and verbal instructions on how to properly fill out the answer form, and were allowed as much time as they needed to complete the test. Along with the knowledge test, each individual was handed an attitude test to complete, after which the subjects were administered a series of physical tests. The lack of participation in the Saturday session necessitated a second day of testing. Five days later students were taken directly out of their activity classes to take the series of tests. Both sessions of knowledge testing took place in a classroom setting.

The post-test was administered approximately eight weeks after the pre-test. Form B of Mood's test was utilized as the testing device. Because of the lack of cooperation from instructors on the first series of tests,

and the time it took to administer the physical tests, it was necessary to allow the subjects to take the knowledge test out of the building to complete on their own time.

Statistical Analysis of the Data

The tests were machine scored and analyzed by Central Washington University's Testing and Evaluative Services. Tests were also scored by hand to be assured of proper correction. The testing service provided individual scores, arithmetic means, variances, and item analysis (indices of discrimination and level of difficulty for each test item). Further analysis was performed on the Texas Instrument 99/4A home computer using programs adapted from VanTassel (49).

The t-test for significance of difference between the means of related groups was used to determine any significant difference between the test scores of physical fitness knowledge of the experimental subjects at the beginning of the quarter and the end. Thirty-three subjects received both forms of the test. The critical t-value, for the degrees of freedom of thirty-two and a level of significance of .05, was ± 2.042.

The t-test for significance of difference between the means of two independent groups was used to determine any significant difference between test scores of the control group and the experimental group at the beginning

of the quarter. Fifty-one experimental subjects and nineteen control group subjects took form A of the test. With the resultant degrees of freedom of sixty-eight, and the level of significance at .05, the critical t-value was $\pm$2.000.

The t-test for two independent groups was used to determine any significant difference between the test scores of the control group and the experimental group at the end of the quarter. Six control subjects and thirty-three experimental subjects took form B of the test at the end of the quarter. The degrees of freedom of thirty-seven and level of significance of .05, required a critical t-value of $\pm$2.042.

A one-way analysis of variance (38) was used to identify any significant difference in the physical fitness knowledge of students enrolled in each of the four categories of activity classes. An F-ratio of 2.93 was needed for significance at the .05 level.

The chi-square test of goodness of fit (38) was used to determine any significant difference between the test scores of Central Washington University students and the national percentile norms for both forms of the test. Raw scores were translated to standard T-scores, then converted to percentile ranks. Expected cell values for all ten cells were one-tenth of the total number of subjects to complete each test. Each cell contained ten

percentile ranks, from zero to nine in the first cell, to ninety through ninety-nine in the tenth cell. The chi-square value for significance at the .05 level was 16.919.

Finally, a one-way analysis of variance was used to determine any significant difference in the difficulty of the ten physical fitness topics. The percentage of physical education activity students to answer each item correctly was found, then grouped with the percentages of the other items from the same topic area. The ten topic areas, the specific item numbers from each form of the test, and the percent of students to answer each question correctly, may be found in appendix C. The critical f-ratio, for the degrees of freedom of 9/104 at the .05 level of significance, was 2.04. The Scheffe test was used to examine pairs of topics in order to identify the specific areas of knowledge that were significantly different.

Chapter 4

RESULTS OF THE STUDY

The analysis of data was performed on the results of test scores for fifty-one experimental subjects and nineteen control group subjects at the beginning of the quarter, and thirty-three experimental subjects and six control group subjects at the end of the quarter. A complete breakdown of the subjects according to sex, major, and class standing may be found in appendix A. The fifty-one experimental subjects at the beginning of the quarter consisted of: eighteen from physical education fitness (PEF) classes, nine from physical education team sports (PETS), eleven from physical education individual/dual sports (PEID), and thirteen from physical education rhythm classes (PER). The thirty-three at the end of the quarter were distributed as follows: fourteen in PEF, six in PEID, six in PETS, and seven in PER. The study experienced a mortality rate of 35 percent for the experimental group (22 percent PEF, 33 percent PETS, 45 percent PEID, 40 percent PER) and 68 percent for the control group.

The t-test for two independent samples was used to

determine any significant difference between knowledge acquisition of students in Central Washington University physical education activity classes and those not enrolled in activity classes. The t-test was performed on data collected at the beginning of the quarter, and on scores obtained at the end of the quarter.

The t-test for two related samples was used to determine any significant difference between physical fitness knowledge from the beginning of the quarter to the end for the experimental subjects, the control group, and subjects from each of the four categories of activity classes.

A one-way analysis of variance was used to determine any significant difference in physical fitness knowledge among the subjects from the four categories of activity classes at the beginning of the quarter and at the end.

A chi-square test was used to determine any significant difference between the performance of Central Washington University students to the established national norms for each form of the test.

Finally, a one-way analysis of variance and the Scheffe test were used to determine any significant difference between the subject's knowledge of the ten specific topics of physical fitness knowledge.

The results of the hypotheses are presented in the

following sections.

Hypothesis one: There will be no significant difference in the physical fitness knowledge of college students enrolled in physical education activity classes from the beginning of the quarter to the end.

The t-test for two related samples, pre-test to post-test scores, for each category of physical education activity class and for the experimental group as a whole, are shown in table 1. The t-values ranged from -1.080 to 0.646. No values were significant.

Table 1

T-test Between Pre-test and Post-test Scores of the Experimental Group (points)

	N	mean A	mean B	diff of means	SE	t
PEF	14	28.9	27.9	-1.0	1.254	-0.797
PETS	6	30.8	28.7	-2.2	2.007	-1.080
PEID	6	27.5	27.7	0.2	2.677	0.062
PER	7	29.6	31.0	1.4	2.213	0.646
TOTAL	33	29.1	28.6	-0.5	0.911	-0.532

Note: T-ratios of 2.160, 2.571, 2.447, and 2.042 were needed at the .05 level of significance.

Hypothesis two: There will be no significant difference in the physical fitness knowledge of college students enrolled in physical education activity classes and those in the control group, not enrolled in activity classes.

The t-test for two independent samples, experimental and control groups pre-test and post-test results, are shown in tables 2 and 3. The t-ratios of 1.131 for the pre-test and -.219 for the post-test were not significant.

Table 2

T-test Between the Experimental and Control Groups Pre-test Means (points)

Group	N	mean	df	SE	diff	t
EXP	51	28.3	68	1.7767	2.0	1.131
CTL	19	26.3				

Note: T-ratio of 2.000 needed for significance at the .05 level.

Table 3

T-test Between the Experimental and Control Groups Post-test Means (points)

Group	N	mean	df	SE	diff	t
EXP	33	28.6	37	3.148	-0.7	-0.219
CTL	6	29.3				

Note: T-ratio of 2.042 needed for significance at the .05 level.

Hypothesis Three: There will be no significant difference between the physical fitness knowledge of students enrolled in any of the four categories of physical fitness classes (fitness, team sports, individual/dual sports, rhythm).

The results of the one-way analysis of variance for the pre-test and the post-test scores of the subjects when grouped in the four categories of physical education activity classes, are shown in Table 4. The F-ratios of .186 for the pre-test and .348 for the post-test were not significant.

Table 4

Analysis of Variance for Significance of Difference Among the Four Categories of Activity Classes (MSb= mean of squares between, MSw= mean of squares within)

Test	df	MSb	MSw	F-ratio
Pre-test	3,29	11.92	64.06	0.186
Post-test	3,29	17.75	50.98	0.348

Note: F-ratio of 2.04 needed for significance at the .05 level of confidence.

Hypothesis four: There will be no significant difference between the physical fitness knowledge of Central Washington University students and the population sample from whom norms were constructed.

The chi-square test of goodness of fit was used to compare the scores of Central Washington students with national norms, as shown in tables 5 and 6. The values of

14.0 and 3.82 were not significant.

Table 5

Chi-Square Test of Central Washington University Students as Compared to National Percentile Norms (Pre-test)

Percentile	0	10	20	30	40	50	60	70	80	90
Expected n	7	7	7	7	7	7	7	7	7	7
Observed n	7	6	7	9	3	15	5	7	4	7
difference	0	-1	0	2	-4	8	-2	0	-3	0

Chi-square = 14.0

Note: Chi-square needed for signicance at the .05 level was 16.919.

Table 6

Chi-square Test of Central Washington University Students as Compared to National Percentile Norms (Post-test)

Percentile	0	10	20	30	40	50	60	70	80	90
Expected n	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9
Observed n	4	4	2	3	3	3	6	4	4	6
difference	.1	.1	-1.9	-.9	-.9	-.9	2.1	.1	.1	2.1

Chi-square = 3.82

Note: Chi-Square needed for significance at the .05 level was 16.919.

Hypothesis five: There will be no significant difference in the percentage of correct responses among the ten specific topics of physical fitness knowledge, based on the test scores of physical education activity students.

A one-way analysis of variance was used to determine any significant difference among the levels of difficulty of the ten topics of physical fitness. The F-ratio of 2.67, as shown in table 7, was significant. Further analysis using the Scheffe test showed no specific topics of significant difference.

Table 7

Analysis of Variance for Significance of Difference Among the Percentages of Correct Responses for the Ten Topics (MSb= mean of squares between, Msw= mean of squares within)

Degrees of freedom	MSb	MSw	F-ratio
9,104	695.2	260.6	2.67*

* Significant at the .05 level of confidence.

Chapter 5

SUMMARY AND CONCLUSIONS, IMPLICATIONS, RECOMMENDATIONS

Summary and Conclusions

The purpose of this study was to determine the relationship between the acquisition of physical fitness knowledge and attendence in a physical education activity class. The two parallel forms of Mood's Test of Physical Fitness Knowledge (28) were used as the testing instruments. A random sample of students from Tuesday-Thursday activity classes comprised the experimental group, while a selection of willing non-activity students were recruited at random for a control group. The t-test for both related and independent samples, a chi-square test of goodness of fit, and a one-way analysis of variance were used to determine statistical significance in data analysis.

The following hypotheses were utilized in order to provide direction to this study. The conclusions drawn from the results of the testing will be given in the following sections. All tests were used to determine significance at the .05 level.

Hypothesis One: There will be no significant difference in the physical fitness knowledge of college students enrolled in physical education activity classes from the beginning of the quarter to the end.

This hypothesis was retained. One group (physical education rhythm) showed an improvement from 29.57 to 31.00, however, most of the improvement was due to the twelve point gain of one subject. One other group showed little improvement, while the other two had lower means at the end. Only twelve of the thirty-three subjects made any gain, three received the same scores, and eighteen scored lower on the second test.

Hypothesis Two: There will be no significant difference in the physical fitness knowledge of college students enrolled in physical education activity classes and those in the control group, not enrolled in activity classes.

This hypothesis was retained. Several factors may have contributed to the failure to find a significant difference. These include the small sample size, the invalid content of the testing instrument for evaluating activity students, the focus of activity classes not being on knowledge development, and the possibility of improper or ineffective teaching.

Hypothesis Three: There will be no significant difference between the physical fitness knowledge of students enrolled in any of the four categories of physical education classes (fitness, rhythm, team sports, individual/dual sports).

This hypothesis was retained. The difference of

means for the pre-test was greatest between team sports and individual/dual sports (3.33), yet the F-ratio was not significant. The greatest difference of means for the post-test was between individual/dual and rhythm classes (3.33), but again the F-ratio was not significant.

Hypothesis Four: There will be no significant difference between the physical fitness knowledge of Central Washington University students and the population sample from whom norms were constructed.

This hypothesis was retained. The chi-square for the pre-test was very close to the significant level because over twice as many scores as expected were located in the percentile range of 50-59, while half as many as expected were in the 40-49 percentile range. Other than those deviations, the scores were distributed fairly evenly among the cells for both tests.

Hypothesis Five: There will be no significant difference in the percentage of correct responses among the ten specific topics of physical fitness knowledge.

This hypothesis was rejected, as the F-ratio found using a one-way analysis of variance was significant. Further analysis utilizing the Scheffe test failed to reveal specific differences among the topics. The mean percentage between certain topics, however, was great enough to indicate that significant differences existed.

Implications

The majority of the studies discussed in the review of literature were performed over twenty years ago. Very little research has been performed in physical education knowledge testing since. The need exists for renewed efforts in the use and construction of standardized tests for physical education.

The fact that students in activity classes made no significant gain in physical fitness knowledge is no surprise. Physical education classes have evolved to the point where little or no teaching of knowledge facts takes place. Even so, one can not expect the teaching of physical fitness knowledge facts in a bowling class or a class in ballroom dancing. It is difficult to speculate on the intellectual benefit of physical education activity classes using a test on one specific, though universally important, aspect of physical education.

Central Washington University students did perform as expected on the tests when compared with national norms established by Mood. This would indicate that the subjects selected for testing were not significantly above or below average.

The most significant results of this study involved the ten topics of physical fitness. The analysis of the percentage of correct responses brought out the

significant difference. The Scheffe test was applied to the pairs of topics with the greatest difference of means to identify the cause of the significant difference found by a one-way analysis of variance. The Scheffe tests showed no significance. This is not unusual since the Scheffe test is not as powerful a statistical tool as an analysis of variance. Since the analysis of variance showed significance, then the pairs of topics with the greatest difference of means may, nevertheless, be considered significantly different (38:421). The greatest difference existed between topics six and one (34.6), followed by nine and one (28.8), six and three (26.7), and nine and three (20.9). These results indicated a general lack of knowledge in the areas of "current status and promotion of physical fitness" (topic one) and "kinesiological aspects of physical fitness" (topic three), while the students were generally knowledgeable about "physical fitness versus disease" (topic six) and "sociological aspects of physical fitness" (topic nine).

It is interesting to note some of the specific questions that were answered incorrectly by an extremely high percentage of the subjects, especially those where a general misconception may exist. This was approached in two ways: an examination of questions that were answered correctly by less than 25 percent of the activity students, and an examination of items where a large

percentage of the subjects chose the same incorrect response.

On form A, five items were answered correctly by less than 25 percent of the subjects. On item thirteen, 53 percent of the subjects incorrectly chose "carbohydrates" as the substrate that "provides the greatest caloric yield." Apparently, there is a misconception since twice as many subjects chose "carbohydrate," than those who answered "fat" correctly. Item thirty-one dealt with the "best index of an individual's physical fitness level." Only 10 percent answered "maximal oxygen consumption" correctly, while 45 percent incorrectly chose "maximal heart rate." On item sixteen, 55 percent of the subjects chose "endurance and speed" as the make up of power, while only 22 percent correctly chose "speed and strength." Sixty-one percent of the subjects incorrectly chose "agility" as the "neuromuscular phenomenon" used in the jump-turn in skiing and a 100-yard dash. Only 20 percent correctly answered "muscular power."

On form B, five items were answered correctly by less than 25 percent of the subjects. Item fifteen asked for the part of the body that fat is removed from during general weight loss. With the current diet and exercise fad, it was a surprise that 45 percent of the subjects chose "the areas being exercised most often," rather than

the correct response, "equally from the fat deposits of the entire body." On item twenty-three, 64 percent answered that "mental practice is as effective as physical practice," while in fact "mental practice is effective, but less so than physical practice." A substantial 61 percent of the subjects, on item thirty-five, chose "too much stress . . . on the value of athletic competition" as the reason for participants in traditional physical education classes to not continue with vigorous activity. The correct response, "knowledge of physiological consequences of inactivity has not been taught," was selected by only 21 percent of the subjects. Only 15 percent of the subjects knew the characteristics of lordosis, asked in item thirty-one. On question fifty-one, 58 percent answered that muscle becomes shorter and wider during isometric exercise, while only 21 percent correctly answered "the muscle length does not change." A misconception apparently exists involving the effect of cholesterol on the body. On question fifty-five, 52 percent blamed it as having the "most pronounced inhibitory effect on neuromuscular performance." Only 15 percent chose the correct response "hydrogen ions."

Recommendations

1. It is recommended that new research be conducted into the construction and use of knowledge tests

for physical education.

2. It is recommended that further investigation be made into the physical fitness knowledge of Central Washington University students, but with the following differences:

(a) the study should be done independently of other studies,

(b) the study be completed over a period of one year rather than one quarter, and

(c) the study involve much larger groups, possibly a random selection of entire classes of students.

3. It is recommended that this study be repeated at Central Washington University using only physical education majors.

4. It is recommended that a current study be made of the level of physical fitness knowledge of high school physical edcuation students.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Annarino, Anthony A., Charles C. Cowell, and Helen W. Hazelton. Curriculum Theory and Design in Physical Education. St. Louis: C.V. Mosby, 1980.

2. Blaess, Nancy. "The Development of an Archery Knowledge Test." Completed Research in Health Physical Education, and Recreation, 8:95, 1966.

3. Bradley, William. "Standardization of a Baseball Knowledge Test for College Men Majoring in Physical Education." Completed Research in Health Physical Education and Recreation, 2:38, 1960.

4. Broer, Marion R. and Donna M. Miller. "Achievement Tests for Beginning and Intermediate Tennis." Research Quarterly, 21:303-321, 1950.

5. Clarke, H. Harrison. Application of Measurement to Health and Physical Education. New Jersey: Prentice-Hall, Inc., 1976.

6. Daugherty, B.R. Discriminative Study of Boy's Physical Education in High Schools in the Big Eight Conference. Master of Education Thesis, Central Washington University, 1969.

7. Dearborn, Terry H. "Personal Health Knowledge of College Students Before Instruction." Research Quarterly, 29:154-159, 1958.

8. Farrow, Andrea. "Skill and Knowledge Proficiencies for Selected Activities in the Required Program at Memphis State University." Completed Research in Health Physical Education and Recreation, 14:163, 1972.

9. Finney, C.D. Discriminative Study of Girl's Physical Education in High Schools of Big Eight Conference. Master of Education Thesis, Central Washington University, 1969.

10. Fox, Katherine. "Beginning Badminton Written Examination." Research Quarterly, 24:135-140, 1953.

11. French, Ester. "The Construction of Knowledge Tests in Selected Professional Courses in Physical Education." Research Quarterly, 14:406-424, 1943.

12. Gershon, Ernest. "Apparatus Gymnastics Knowledge Test for College Men in Professional Physical Education." Research Quarterly, 28:332-339, 1957.

13. Hardin, Ruby. "The Construction of an Information Examination for College Students Enrolled in Bowling Classes." Completed Research in Health Physical Education and Recreation, 4:72, 1962.

14. Hemphill, Fay. "Information Tests in Health and Physical Education for High School Boys." Research Quarterly, 3:83-96, 1932.

15. Hennis, Gail M. "Construction of Knowledge Tests in Selected Physical Education Activities for College Women." Research Quarterly, 27:301-309, 1956.

16. Hewitt, Jack. "Comprehensive Tennis Knowledge Test." Research Quarterly, 8:74-84, 1937.

17. _____. "Hewitt's Comprehensive Tennis Knowledge Test Forms A and B Revised." Research Quarterly, 35:147-155, 1964.

18. Hodges, Carolyn. "Construction of an Objective Knowledge Test and Skills Tests in Lacrosse for College Men." Completed Research in Health Physical Education and Recreation, 10:61, 1968.

19. Hooks, Edgar. "Hook's Comprehensive Knowledge Test in Selected Physical Education Activities for College Men." Research Quarterly, 37:506-514, 1966.

20. Johnson, Ralph. Knowledge and Understanding in Physical Education. American Alliance for Health, Physical Education and Recreation, Washington D.C., 1969.

21. Kelly, Ellen, and Jane Brown. "The Construction of a Field Hockey Test for Women Physical Education Majors." Research Quarterly, 23:322-329, 1952.

22. Kirkendall, Don, Joseph Gruber and Robert Johnson. Measurement and Evaluation for Physical Educators. Dubuque:Wm. C. Brown, 1980.

23. Knighton, Marion. "Soccer Questions." Journal of Health and Physical Education, 1:29,60, 1930.

24. Kraft, George. "The Construction and Standardization of a Wrestling Knowledge Test for College Men Majoring in Physical Education." Abstracts of Research Papers 1972, Washington D.C.:AAHPER, 1972.

25. McCutcheon, Sallie. "The Construction of an Objective Basketball Knowledge Examination for College Women." Completed Research in Health Physical Education and Recreation, 8:69-7, 1966.

26. Miller, Wilma. "Achievement Levels in Tennis Knowledge and Skill for Women Physical Education Major Students." Research Quarterly, 24:81-90, 1953.

27. Mood, Dale. "Test of Physical Fitness Knowledge: Construction Administration and Norms." Research Quarterly, 42:423-430, 1971.

28. Murphy, Mary. "Gain in Health Knowledge of Two Groups of Women Students Classified in Physical Education." Research Quarterly, 8:78-88, 1937.

29. _____. "Criteria for Judging a Golf Knowledge Test." Research Quarterly, 4:81-88, 1933.

30. _____. "Grading Student Achievement in Golf Knowledge." Research Quarterly, 5:83-90, 1934.

31. Nash, Jay B. Physical Education: Interpretation and Objectives New York:A.S. Barnes, 1948.

32. Palmer, Wendall. "An evaluation of a Speed-a-way Knowledge Test." Completed Research in Health Physical Education and Recreation, 6:51-52, 1964.

33. Parks, Jack. "Comprehensive Knowledge in Track and Field." Completed Research in Health Physical Education and Recreation 2:30, 1960.

34. Phillips, Marjorie. "Standardization of a Badminton Knowledge Test for College Women." Research Quarterly, 17:48-63, 1946.

35. Power, William. "A knowledge Test in Touch Football for Junior High Boys." Completed Research in Health Physical Education and Recreation, 2:30, 1960.

36. Reichle, Marvin. "Development of an Instrument to Measure Knowledge in Recreational Boating." Completed Research in Health Physical Education and Recreation, 1976.

37. Rodgers, Elizabeth. "The Standardization and Use of Objective Type Information Tests in Team Game Activities." Research Quarterly, 10:102-112, 1939.

38. Roscoe, John T. Fundamental Research Statistics for the Behavioral Sciences. New York:Holt, Rinehart, and Winston, Inc., 1969.

39. Safrit, Margaret. Evaluation in Physical Education. New Jersey: Prentice Hall, 1973.

40. _____. Teacher's Manual Evaluation in Physical Education. New Jersey:Prentice-Hall, 1980.

41. Schwartz, Helen. "Knowledge and Achievement Tests in Girls Basketball on the Senior High Level." Research Quarterly, 8:143-156, 1937.

42. Scott, Gladys. "Achievement Examinations for Elementary and Intermediate Swimming Classes." Research Quarterly, 11:100-111, 1940.

43. _____. "Achievement Examinations for Elementary and Intermediate Tennis Classes." Research Quarterly, 12:40-49, 1941.

44. _____. "Achievement Examinations in Badminton." Research Quarterly, 12:242-253, 1941.

45. Shambaugh, Mary. "The Objective Measurement of Success in the Teaching of Folk Dancing to University Women." Research Quarterly, 6:33-58, 1935.

46. Shick, Jacqueline. "Written Tests in Activity Classes." Journal of Health Physical Education and Recreation, pp. 21-22,83, April, 1981.

47. Snell, Catharine. "Physical Education Knowledge Tests." Research Quarterly, 6:78-94, 7:73-82, 7:77-91, 1935, 1936.

48. Stradtman, Alan, and T.K.Cureton. "A Physical Fitness Knowledge Test for Secondary School Boys and Girls." Research Quarterly, 21:53-57, 1950.

49. VanTassel, Dennie. BASIC-Pack Statistics Programs for Small Computers. New Jersey:Prentice-Hall, 1981.

50. Waglow, I.F., and C.H. Rehling. "A Golf Knowledge Test." Research Quarterly, 24:463-470, 1953.

51. Waglow, I.F., and F. Stephens. "A Softball Knowledge Test." Research Quarterly, 26:234-243, 1955.

52. Walker, William. "The Development of a General Knowledge Inventory Test and a Research Syllabus for a Foundation Course in Physical Education for College Freshmen." Dissertation Abstract International, 26, 1965, 4431.

53. Warren, Margaret. "Construction of a Knowledge Test in Lacrosse." Completed Research in Health Physical Education and Recreation, 14:118, 1972.

54. Wilson, Ruth. Determination of Evaluative Devices for Adequate Assessment of Levels of Competence in Certain Physical Education Activities. Washington D.C.: Office of Education (DHEW), September, 1969.

55. Yakima Herald Republic, February 9, 1984.

APPENDIX A

BREAKDOWN OF RESEARCH SUBJECTS

BREAKDOWN OF RESEARCH SUBJECTS

Experimental Group -- Pre-test Subjects

	Females (n=24)				Males (n=27)				
	Fr	S/J	Se	Gr	Fr	S/J	Se	Gr	TOTALS
PE maj	1	-	-	-	-	-	1	-	2
PE min	-	1	-	-	2	-	-	-	3
NON-PE	11	8	2	1	5	15	4	-	46
TOTALS	12	9	2	1	7	15	5	0	51

Experimental Group -- Post-test Subjects

	Females (n=14)				Males (n=19)				
	Fr	S/J	Se	Gr	Fr	S/J	Se	Gr	TOTALS
PE maj	-	-	-	-	-	-	1	-	1
PE min	-	1	-	-	2	-	-	-	3
NON-PE	7	4	1	1	4	10	2	-	29
	7	5	1	1	6	10	3	0	33

Control Group -- Pre-test Subjects

	Females (n=8)				Males (n=11)				
	Fr	S/J	Se	Gr	Fr	S/J	Se	Gr	TOTALS
PE maj	-	-	1	-	-	-	-	-	1
NON-PE	1	5	-	1	1	6	4	-	18
TOTALS	1	5	1	1	1	6	4	0	19

Control Group -- Post-test Subjects

	Females (n=4)				Males (n=2)				
	Fr	S/J	Se	Gr	Fr	S/J	Se	Gr	TOTALS
PE maj	-	-	1	-	-	-	-	-	1
NON-PE	-	2	-	1	1	1	-	-	5
TOTALS	0	2	1	1	1	1	0	0	6

APPENDIX B

RAW SCORES

RAW SCORES

Experimental Subjects								Control Group	
PEF		PETS		PEID		PER			
A	B	A	B	A	B	A	B	A	B
33	28	33	29	17	19	31	27	28	31
19	18	35	27	29	22	38	37	35	35
36	31	20	21	9	18	28	25	23	-
20	20	34	30	29	33	25	32	23	-
28	20	23	29	40	32	30	30	29	-
33	29	40	36	41	42	26	25	30	-
17	17	27	-	29	-	29	41	30	-
31	40	31	-	20	-	11	-	24	-
26	28	26	-	31	-	26	-	20	22
41	35			26	-	33	-	22	-
24	30			22	-	32	-	33	-
37	39					33	-	28	-
28	26					25	-	17	-
31	29							29	-
25	-							29	-
27	-							26	-
22	-							30	38
38	-							23	17
								21	33

APPENDIX C

TEN TOPICS OF PHYSICAL FITNESS KNOWLEDGE

TEN TOPICS OF PHYSICAL FITNESS KNOWLEDGE
Item Number (percent correct)

1. Current Status and Promotion of Physical Fitness (29.67%)
   Form A - 25(33%), 30(25%), 36(18%)
   Form B - 35(24%), 40(33%), 49(45%)

2. Evaluation of Physical Fitness (50.73%)
   Form A - 8(53%), 15(63%), 21(61%), 31(10%), 38(59%), 56(35%)
   Form B _ 32(73%), 44(55%), 45(52%), 46(61%),54(36%), 58(42%)*

3. Kinesiological Aspects of Physical Fitness (37.58%)
   Form A - 19(47%), 23(55%), 43(35%), 44(20%), 53(37%), 61(31%)
   Form B - 7(48%), 11(39%), 24(52%), 27(39%), 28(33%), 31(15%)

4. Nutritional Aspects of Physical Fitness (50.55%)
   Form A - 5(76%), 13(27%), 32(43%), 55(75%), 60(39%), 62(29%)
   Form B - 8(76%), 15(36%), 20(55%), 38(73%), 52(27%), 62(45%)*

5. Physical Fitness Programs (46.86%)
   Form A - 12(43%), 17(51%), 27(73%), 28(86%), 42(25%), 48(57%), 59(39%)
   Form B - 5(67%), 12(58%), 22(33%), 37(27%), 51(21%), 53(15%), 56(61%)

6. Physical Fitness Versus Disease (64.33%)
   Form A - 6(76%), 18(69%), 46(43%)
   Form B - 6(73%), 10(52%), 13(73%)

7. Physiological Aspects of Physical Fitness (45.05%)
   Form A - 10(39%), 11(24%), 14(39%), 20(39%), 24(31%), 26(29%), 29(57%), 39(39%), 41(49%), 47(63%), 52(39%), 64(47%)
   Form B - 18(70%), 19(55%), 21(61%), 29(52%), 36(61%), 47(64%), 55(15%), 57(52%), 59(58%)*, 60(36%)*, 63(48%)*, 64(21%)

8. Psychological Aspects of Physical Fitness (48.25%)
   Form A - 7(71%), 22(51%), 33(53%), 35(37%), 37(47%), 49(59%)
   Form B - 16(55%), 23(30%), 26(61%), 30(58%), 33(39%), 34(18%)

9. Sociological Aspects of Physical Fitness (58.50%)
   Form A - 45(20%), 50(75%), 51(63%), 54(55%), 63(80%)
   Form B - 17(70%), 41(42%), 42(52%), 48(61%), 50(67%)

10.Miscellaneous Concepts of Physical Fitness (49.27%)
Form A - 9(60%), 16(22%), 34(41%), 40(55%), 57(51%), 58(59%)
Form B - 9(58%), 14(36%), 25(55%), 39(45%), 43(55%), 61(36%)*

* Item not used in comparing levels of difficulty due to an error in test reproduction.