

**THE COMPARATIVE EFFECTS OF TEST ADMINISTRATION METHODS
ON PHYSICAL FITNESS TEST**

GAIL WHITING

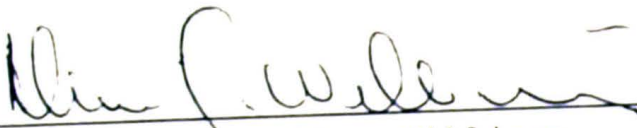
To the Graduate and Research Council:

I am submitting herewith a field study written by Gail Whiting entitled, "The Comparative Effects of test administration Methods on Physical Fitness Tests." I have examined the final copy of this paper for form and content, and I recommend that it be accepted in partial fulfillment of the requirements for the degree of Education Specialist, with a major in Education.

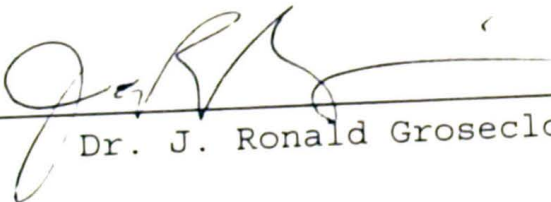


Dr. Joe Brown, Major Professor

We have read this field study
and recommend its acceptance:

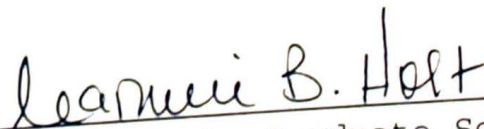


Dr. Allan Williams



Dr. J. Ronald Groseclose

Accepted for the Graduate
and Research Council



Dean of the Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this field study in partial fulfillment of the requirements for a Master's degree at Austin Peay State University, I agree that the Library shall make it available to borrowers under rules of the library. Brief quotations from this field study are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this field study may be granted by any major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this field study for financial gain shall not be allowed without any written permission.

Signature Gail White
Date 2/5/97

THE COMPARATIVE EFFECTS OF TEST ADMINISTRATION METHODS
ON PHYSICAL FITNESS TESTS

A Field Study

Presented to

the Graduate Council of
Austin Peay State University

In Partial Fulfillment
of the Requirements for the Degree
Education Specialist

Gail Whiting

May, 1997

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Joe Brown, for his guidance and patience. His assistance was invaluable. I would also like to thank the other committee members, Dr. Allan Williams for his encouragement and Dr. J. Ronald Groseclose for his comments and assistance.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
STATEMENT OF THE PROBLEM	2
SIGNIFICANCE OF THE STUDY	2
II. REVIEW OF RELATED LITERATURE	4
III. METHOD	15
Subjects	15
Procedures	16
IV. RESULTS	18
V. DISCUSSION	21
VI. REFERENCES	23

LIST OF TABLES

TABLE	PAGE
1. Two-Way ANOVA	18
2. Mean and Standard Devistion of the Three Groups. . .	19

LIST OF FIGURES

Figure	Page
1. Mean Comparisons on the Groups	20

CHAPTER I

INTRODUCTION

Physical education is an enjoyable class for many students. However, for the few students who simply dislike physical education, the annual fitness test is a particularly distasteful event. Unfortunately, the instructor sometimes inadvertently contributes to the students' distaste for testing. Accordingly, there are several teacher-controlled variables that can influence test performance. Preparation and test methods are two.

There is little research available that addresses the preparation of students for tests or how tests are administered. Academic standardized tests are given under very controlled conditions with classrooms, proctors, and sample tests being the same each time. Physical fitness test conditions, however, vary from instructor to instructor and facility to facility. Consequently, this procedure makes it unfair to compare scores against norms established under very different conditions. It is also reasonable to suggest that this unfairness might be a result

of the teacher's test administration protocol. For this reason, the present study offers some insight into what effect the test administration method has on physical fitness scores.

STATEMENT OF THE PROBLEM

The problem to be investigated in this study is the comparative effects of test administration methods on physical fitness tests. Specifically, are there significant differences produced by the traditional, competition or encouragement methods?

SIGNIFICANCE OF THE STUDY

It is hoped that the information in this study will make instructors aware of the different variables that may have an effect on the performance of their students during physical fitness testing. If true measurements are to be obtained, it is important that consistency be maintained. If we cheer for some students and quietly observe others, are we creating an advantage for those students? Is it fair one student is given three trials and another only two?

More importantly, such test administration variation can affect student motivation.

The most successful teachers and coaches at all levels have strong motivational skills. They are adept at bringing out the very best in their athletes. As evaluators, this can be a very valuable tool. Consequently, it is necessary that instructors continue to do all that they can to ensure that the tests are administered correctly and as equitably as possible.

CHAPTER II

REVIEW OF RELATED LITERATURE

Physical fitness tests were first instituted during the 1960s, and American youth have been taking them ever since. They remain today the most widely used form of assessment in physical education. A study by Hoppie and Graham (1995) has shown students view fitness tests as meaningless, do not like them and often become artful test dodgers. The tests are meaningless because students often do not understand what a test is measuring. The children in this study, when asked, could not explain the purpose of participating in the mile run. They felt it was a negative experience and avoided it whenever possible. Essentially, the children disliked the test because it was not fun. Instead, they wanted an activity that was more game like and more interesting. When motivation and interest are low, test scores will also be low. It appears likely that the test methods had an adverse influence on performance.

Nelson and Dorociak (1982) found that the scores of students who were allowed to practice were significantly higher than those that did not. There is often a paradox

between administrative standards and criteria for validity and reliability. For example, as a time saving technique, an instructor may allow fewer trials, and in the process possibly sacrifice test reliability. The authors' methods resulted in not only higher scores but more consistent scores on tests taken on different days. They further stated that student involvement added to validity, reliability, efficiency and a better understanding of their individual capabilities and progress.

Cohen (1986) found that some students take responsibility for their successes and failures while other students do not believe their actions are related to their achievements. Further, she states that research has shown the cycle of failure for these students can be broken through motivation. Thus, children who recognize the positive relationship between preparation and results through teacher feedback have overcome the helpless learner state. Their loci of control becomes more internal, and they are more easily motivated. Then, this higher motivation enables them to reach their true potential. Such action results in higher fitness test scores.

The importance of adhering to strict testing guidelines is supported by the Physical Fitness Training Manual for the U.S. Army. In an effort to maintain uniformity of tests, the army has developed exhaustive protocols that must be followed exactly. Testers are trained, and the testees are given a thorough knowledge of the tests and facilities. Further, the personnel and facilities must meet certain standards, and nothing is left to chance. Soldiers are told what to wear, and testers are told what to say. The facilities must be of a certain size and quality. Performance is closely monitored to ensure that each repetition is completely properly. The battery of tests are even taken in an assigned sequence. One can confidently say the test norms set by the U.S. Army are as valid and reliable as is possible.

The Department of Health and Human Services publishes the Get Fit Handbook (1995) for the Presidents Physical Fitness Test (1995). It includes test descriptions, norms and some preparatory information. It also gives guidelines for test conditions that can vary greatly. This variance in turn negates valid comparisons among students who take fitness tests.

Fox (1991) states that a child's interpretation of his/her physical activity becomes important when he/she decides to become more active. There is little motivational value in running when children perceive running to be pointless or unpleasant. When a child feels he/she is able to complete a task and that the task is worthwhile, his/her performance is enhanced. It is, then, vital that the instructor create a climate of efficacy.

Obtaining maximum performance from students during testing is an inherent problem. Fitness levels have been examined for several years and test scores have declined over this period. West (1987) believes it is reasonable that the feedback given during performance is not enough to motivate children. She identifies interest, motivation, relevant feedback and ease of administration as important factors in motivation. If a child does not think running for an extended time is fun, he/she will place little value in the effort. For example, knowing that being fit will help you to live longer has little meaning to someone young enough to still believe in their own immortality. Thus, a score at the 85th percentile is meaningless to most students.

When testing cannot be completed in a relatively short time, it is not uncommon for instructors and students to tire of the task and become careless and inconsistent in test administration. As a result, administration styles will vary from individual to individual. Plowman (1983), therefore, suggests that workshops should be developed where questions can be answered and instructors can become familiar with test procedures. In addition, attention should be given to the changing concept of fitness and time must be provided for the practice of test skills. Also, record keeping and the reporting of results present special problems. To alleviate these problems and to save time, the author recommended an optical scan form be used to store data.

The climate in which a test is taken can greatly influence how well or how poorly students perform. Petry (1989) has made several recommendations to assist in the preparation of efficient test administration and to provide for an appropriate testing atmosphere. She begins by stating that one should have a thorough knowledge of the test manual and understand all directions. Moreover, accommodations should be made for students with special

needs. Additionally, time availability, the order of presenting the test items and sufficient practice time should be addressed. Another important factor is assuring the child that the goal of physical fitness testing is to provide information and stimulate improvement.

Probably, the most popular method for organizing classes is the station method and pullout method. Scheduling concerns include class size, length of class period, facilities, equipment and staff available. When these procedures are followed, students will be properly prepared for testing (Petry, 1989).

According to the American Alliance for Health, Physical Education, Recreation and Dance, the fitness level of American children has been steady since the 1950s.

Pangrazzi (1993) states that a significant amount of test performance can be directly attributed to hereditary and maturation. The environment and nutritional lifestyle can also make a difference. For example, heat, humidity and the quality of the student's diet all play a part in how well a student will perform. The child who eats nutritious meals regularly can be expected to perform better than the child who does not. All children benefit from physical activity

but do so at varying rates. The more mature child will probably do better than the less mature child. This means some children will require more encouragement and feedback than others.

A reward reinforces the end product of fitness. It does nothing to encourage participation when students realize their performance is what is really valued. Nonetheless, performance awards usually motivate only the child who believes that he/she has the ability to earn them (Corbin, 1988). In fact, most of the students will find fitness tests difficult, if not impossible, and others will feel that trying is futile. Consequently, it would be wiser and more beneficial if rewards were based upon regular participation and activity; thereby, presenting all students with the opportunity for receiving an award as well as creating the habit of being active. In other words, effort should count as much as performance.

Effort is critical to one's performance and anything that a teacher can do to encourage it will yield results. Students have many different reasons for wanting to accomplish set goals and require different motivational techniques (Veal, 1995). Thus, instructors must seek the

most effective methods, because what works with one child may not work with others. Veal states that by giving feedback instructors can raise skill levels. Therefore, students must recognize that effort and performance are both important and an increase in effort will increase performance.

The concept of effort plus performance is utilized effectively by health clubs. As an illustration, physical fitness testing has become an integral part of the health club program. Tests include an evaluation of aerobic endurance, strength, and flexibility. It is also very important that the person administering the tests is qualified and trained so individualized programs can be adjusted to reach set goals (Rees, 1994).

There are seemingly countless variables that impact results of fitness tests. Some lie within the limitations of the physical education program such as lack of time, class size, limited funds and facilities (Hill, 1992). Hence, it is very challenging to develop a program which leads to improved test scores with large numbers of students for short periods of time in inadequate facilities. It is not uncommon for two, even three, teachers to simultaneously

share a gymnasium. Researchers have also concluded that students who are involved in the development of the program tend to quickly attain fitness better than those students whose plan is dictated to them. For this reason, Corbin (1987) suggests that the emphasis be placed on fitness components and not on athletic talent or competition among the students.

Unfortunately, there still exists today a surprisingly high number of students who do not participate in a daily physical education program. As few as 36% had a daily physical education program and still others (37%) only had physical education once or twice a week (Nelms, 1990). Seidenberg (1989) found that students who participated in physical education programs twice a week had significantly lower test scores than students who had daily physical education. Blazer, Leeds, McSwegin and Petry (1989) suggest the length of a class should range between 20 and 45 minutes.

Since fitness plays a major influence on life long health status, year round physical activity is needed because fitness maintenance is an ongoing process. Hence, students must be taught the how, why and what of fitness so

they can make wise choices for the rest of their lives (McSwegin, Pemberton, & Petry, 1989). One important choice is quality and organization. The quality and the organization of a program may have an effect on fitness levels of children. Although not definitive, some research has shown that by supplementing the regular program with specific fitness activities, scores can be raised (Deal, Updyke, & Gallahue, 1993).

Physical education has reached a "fork in the road." One path is the traditional sports and games curriculum and the other path is the health and fitness curriculum. Traditionally, students were taught the skills they needed to effectively play team sports such as basketball and volleyball. With the health and fitness curriculum, the emphasis has shifted toward activities that can be pursued into adult life with an emphasis on maintaining a healthy way of living. The physical fitness tests administered today such as Physical Best and the President's Physical Fitness Test are health related tests. Yet, these tests are not administered the same way for everyone.

It is clear that the traditional approach to physical education is an ineffective method for improving physical

fitness scores (Cicerno, Gilliam, & Ray, 1988). It is also clear that administration methods of these tests are not consistent from place to place or time to time. If as suggested by the literature, test administration does make a difference in performance, a case for standardizing test protocols would be a significant contribution to the profession. Therefore, this study was undertaken to determine the influence of different test administration methods on physical fitness test.

CHAPTER III

METHOD

Subjects

The subjects were third grade students enrolled at St. Bethlehem Elementary School, Clarksville, Tennessee. They included both males and females eight years of age. The students were divided into three groups. Group one was the traditional method. Students were told "You are being tested to see how fit you are. Try to do your best." The instructor remained quiet during the performance of the test. Group two was the competition group. Students were told "You are being tested to see how fit you are. Your scores will be compared to your classmates and to students at other schools. Try to do your best." The instructor remained quiet during the performance of the test. Group three was the encouragement group. Students were told "You are being tested to see how fit you are. Try to do your best." During the performance of the test, the instructor gave constant encouragement, motivation and praise to the student.

Each of the three instructors chose one test method to administer to their classes. Class assignments were made by the principal at the beginning of the school year. All students were given time to practice and become familiar with the test items and procedures prior to testing. Each group was tested in a separate area to avoid contamination of method. The tests were administered in accordance with the following guidelines.

40 Yard Dash - The student attempted to run the 40-yard dash as fast as possible. The child stood behind a starting line, ran through a 40-yard distance and past a finish line. Scores were recorded to the nearest tenth of a second.

Curl-ups - The student was supine with knees bent, feet on the floor and arms crossed. The feet were held by a fellow student. The child kept his/her elbows against the chest and rose up so his/her elbows touched the thighs. The number of completed curl-ups was recorded as his/her score.

Shuttle Run - For the shuffle-run, two parallel lines were marked on the floor twenty feet apart. Two chalk board erasers were placed behind one line. The student ran from one line to the opposite line, placed the eraser behind the

line and then repeated the same action with the second eraser. The scores were recorded to the nearest tenth of a second.

Flexed Arm Hang - The student was lifted up to the horizontal bar so his/her chin was above the bar and his/her grip was an overhand one. The student held this position until his/her chin dropped below the bar. The time was recorded to the nearest tenth of a second.

CHAPTER IV

RESULTS

All of the raw data were converted to t-scores to facilitate a comparison among differing measurement methods. The data were analyzed by a Two-way Analysis of Variance. The results showed that there was no difference among the traditional, competition or encouragement methods (see Table 1).

TABLE 1

TWO-WAY ANOVA

Source	Sum of Sqr.	DF	R-Ratio	Prob F
Administration Method	0.04	2	0.00	0.9998
Fitness Items	0.28	3	0.00	
Administration X Fitness	48382.55	472		
Error	48383.47			
Total	48383.47	483		

A comparison of the means and the standard deviations showed the groups were very similar (see Table 2).

TABLE 2
MEAN AND STANDARD DEVIATION OF THE THREE GROUPS

Groups	N	Mean	St. Dev.
Encouragement	164	50.03	10.02
Traditional	166	50.02	9.99
Competition	154	50.05	10.08

These data are graphically displayed in Figure 1. It is clear from this graphic data that neither test administration method was more effective than the others. Therefore, based on this study and within its limitations, it may be concluded that the test administration method did not significantly influence the test performance of these subjects.

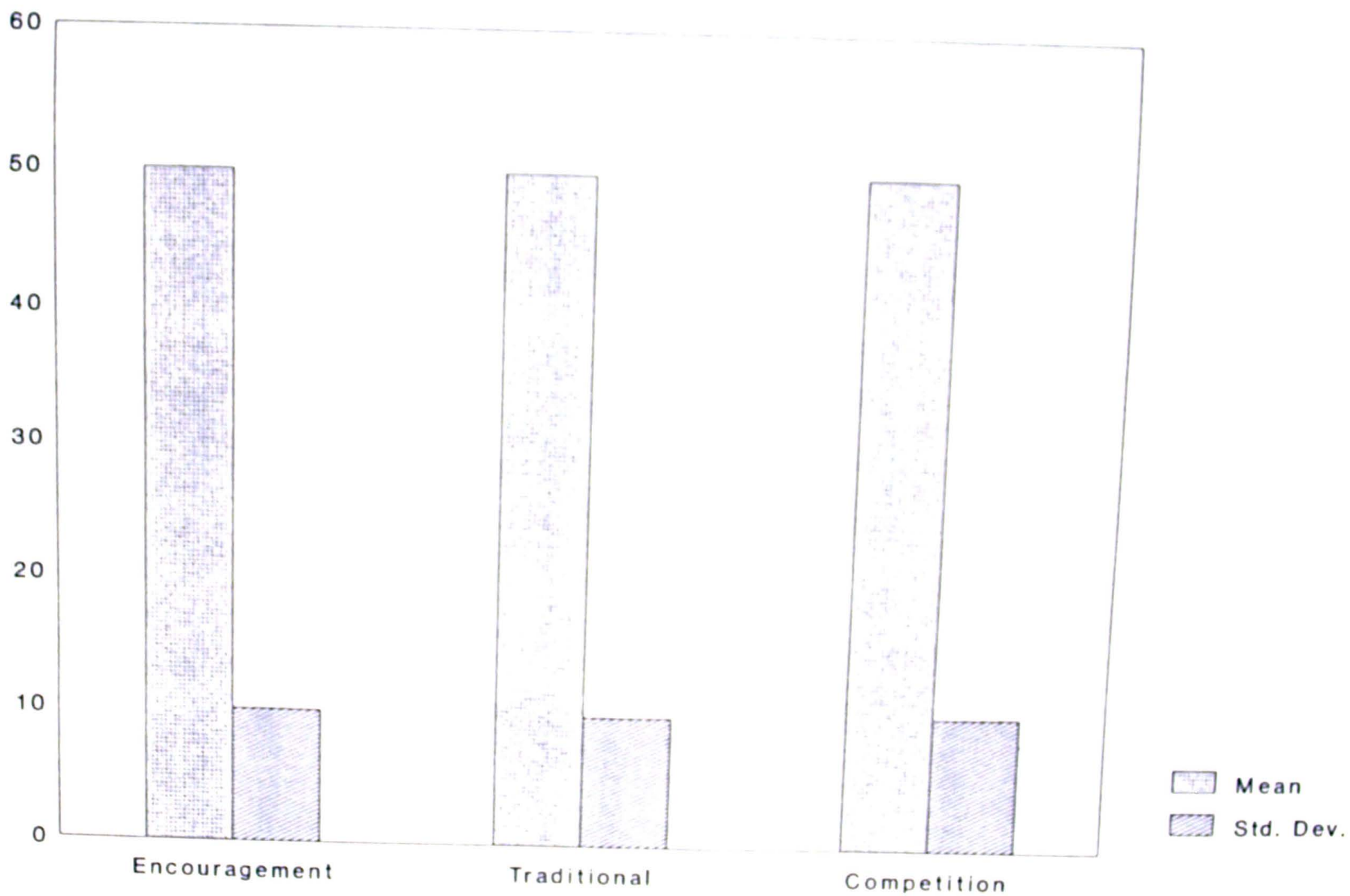


Figure 1: Mean Comparisons on the Groups

DISCUSSION

This study examined the effect that many variables can have on the performance of students being evaluated by a physical fitness test. The results of these tests are measured against national standardized norms. The tests administered in the field are taken under a wide variety of conditions using an equally wide variety of methods. It was reasonable to suggest that it is unfair to compare scores against norms that are established under very different conditions. However, this suggestion was not supported by the data. In actuality, the results showed no significant difference among the traditional, competition and encouragement methods. As might have been expected, the scores of the encouragement group were not significantly higher than the other two groups. In fact, the results showed little to no difference among the three test groups. This result was unexpected because the previous study (Kraft, 1991) obtained a difference.

It is impossible to ensure that all variables are controlled although there are some things that can be

improved. Consequently, qualified professionals, adequate facilities and testing conditions are needed (Rees, 1994). Classroom teachers are wonderful in the classroom, but they lack the specialized training for administering fitness tests. Yet, these teachers can and do plan quality daily periods of physical activity. Thus, it is important that the curriculum teachers follow be changed to reflect the modern reason for testing.

To maximize fitness test results, it is best that conditioning activities be included in the physical education program (Deal, 1993). In the past, physical fitness tests were based more on athletic ability than fitness (Corbin, 1987). Hence, students who have a natural talent do well while the other students struggle. Fortunately, today's physical fitness tests are more health related, and they are a better measure of fitness than previous tests.

References

Blazer, S., Leeds, M., McSwegin, P., & Petray, C.

(1989). Journal of Physical Education Recreation and Dance, 60(1), 42-45.

Cicero, L., Gilliam, G., & Ray, T. (1988). A physical education dilemma: Team sports or physical fitness. Journal of Physical Education, Recreation and Dance, 45, 69-72.

Cohen, M. W. (1986). Research on motivation: New content for the teacher preparation curriculum. Journal of Teacher Education, May-June, 1986.

Corbin, C., Longjoy, P., Whitehead, J. (1988). Youth physical fitness awards. Quest, 40, 200-218.

Corbin, C. (1987). Youth fitness, exercise and health: There is much to be done. Research Quarterly, 58, 308-314.

Deal, T., Gallahue, D., & Updike, W. (1993). Curriculum effectiveness in promoting physical fitness in fifth and seventh grade students. Physical Educator, 50(3), Fall, 1993.

Fox, K. (1991). Motivating children for physical activity: Towards a healthier future. Journal of Physical Education, Recreation and Dance, Sept. 1991.

Hill, G., Mullen, S., & Randle, K. (1992). What children think, feel and know about physical fitness testing. Journal of Teaching Physical Education, 14, 408-417.

Kraft, R. (1991). A comparison of fitness administration measures. (ERIC ED 330669).

McSwegin, P., Pemberton, C., & Petry, C. (1989). An education plan. Journal of Physical Education, Recreation and Dance, 60(1), 32-34.

Melms, B. (1990). They need more exercise. Journal of Pediatric Health Care, 4(4), 167-168.

Nelson, J. (1982). Reducing administration times while improving reliability and validity of fitness tests. Journal of Physical Education, Recreation and Dance, 53, 63-65.

Pangrazzi, R. (1993). Physical fitness: Questions teachers ask. Journal of Physical Education, Recreation and Dance, 64(7), 14 Sept. 1993.

Petry, C. (1989). Designing the fitness testing environment. Journal of Physical Education, Recreation and Dance. 60(1), 35-38.

Plowman, S. (1983). A workshop model that works: Implementation strategies for health related fitness testing. Journal of Physical Education, Recreation and Dance. 54(2), 46-47, Feb. 1983.

Rees, S. (1994). How fit are you? American Health. 13(2), 68, March, 1994.

Seidenberg, S. (1989). The effects of daily physical education on the fitness levels of elementary school students. Austin Peay State University (ED 318719).

Veal, M., & Compazmone, N. (1995). How sixth graders perceive effort and skill. Journal of Teaching in Physical Education. 14(4), July 1995.

West, B. H. (1987). Color coding children's fitness performance. Physical Educator. 44(4), 400-405.

Department of the Army, Headquarters. Physical Fitness Training. FM 21-20.