VARIABILITY IN MAXIMUM STRENGTH TESTING OF UNIVERSITY WOMEN

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS IN PHYSICAL EDUCATION IN THE GRADUATE SCHOOL OF THE TEXAS WOMAN'S UNIVERSITY

COLLEGE OF

HEALTH, PHYSICAL EDUCATION, AND RECREATION

BY

SHARON LOUISE DEWEY, B. S.

DENTON, TEXAS

SPRING, 1971

Texas Woman's University

Denton, Texas

May 3 19 71

We hereby recommend that the Thesis prepared under our supervision by Sharon Louise Dewey entitled VARIABILITY IN MAXIMUM STRENGTH TESTING OF UNIVERSITY WOMEN

be accepted as fulfilling this part of the requirements for the Degree of

Committee:

Chairman

Accepted:

Dean of Graduate Studies

ACKNOWLEDGMENTS

The author wishes to express her sincere appreciation to Doctor Joel Rosentswieg for his dedicated efforts in directing this study. Grateful acknowledgment is extended to Doctor Claudine Sherrill, Doctor Marilyn Hinson, and Doctor Bert Lyle whose helpful suggestions and recommendations were incorporated into this study.

A note of sincere thanks is extended to Susan Allen for her criticism, suggestions, and optimistic encouragement which she offered throughout the conduction of the investigation.

The author is particularly indebted to her family and close friends whose supporting influence and expressed confidence during the past year facilitated the completion of the study. Finally, special recognition and gratitude is offered to the students who gave of their time during the experimental session.

TABLE OF CONTENTS

Chapter

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Measures of all human phenomena vary, and the limits of such variation are not known for strength testing for women. The literature indicates that maximal strength efforts for women seem to be unreliable because women are believed to give up before a maximum effort is made.[1] Motivation is considered an important factor for strength performance[2] and the lack of motivation for women may stem from a cultural stereotype.[3, 4] If this is true, the reliability of maximum strength performance for women may be limited and therefore should not be attempted in any situation but carefully controlled research. The present investigation is concerned with variability in maximum.

---

[1]Herbert A. de Vries, Physiology of Exercise (Dubuque, Iowa: Wm. C. Brown Company Publishers, 1966), p. 315.

[2]Barry L. Johnson and Jack K. Nelson, "Effect of Different Motivational Techniques During Training and in Testing upon Strength Performance," Research Quarterly, XXXVIII (December, 1967), p. 630.

[3]Walter Kroll, "Test Reliability and Errors of Measurement at Several Levels of Absolute Isometric Strength," Research Quarterly, XLI (May, 1970), p. 155.

[4]Perry B. Johnson, et al., Physical Education: A Problem-Solving Approach to Health and Fitness (New York: Holt, Rinehart and Winston, 1966), p. 348.

strength performance in university women in the hope that such knowledge may have practical application to strength testing for women.

Statement of the Problem

The present investigation entailed a study of 100 women enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971, to determine the variability in maximum strength performance of the arm flexors and the muscles in the shoulder girdle. Maximum strength was measured on the Universal Gym Machine by use of the bench press lift every other week for a period of ten weeks. The subjects were unaware of their performance. A conclusion was drawn concerning the inter-individual and group variability of maximum strength performance in university women.

Definitions and/or Explanations of Terms

For the purpose of clarification, the following definitions and explanations have been established for use throughout the study:

A. Maximum Strength: The investigator accepts the definition of Mathews who states that maximum strength is "the force that a muscle or group of muscles can press against a resistance in one maximum effort."[1]

---

[1]Donald K. Mathews, Measurement in Physical Education (Philadelphia: W. B. Saunders Company, 1968), p. 5.

B. Bench Press: The bench press lift as used throughout the study may be described as lying on a bench with the head next to the Universal Gym Machine, the bend of the handles just above the chest, and the feet on the floor. The action is to press the weight up and exhale sharply, then return weight down with control.[1]

Purpose of the Study

The purpose of the study was to determine the extent of variability with women in repeated measures of maximum strength. Specifically, the following null hypotheses were tested:

A. There is no significant difference in the inter-individual variability of maximum strength performance as measured by the bench press lift on the Universal Gym Machine.

B. There is no significant difference in the group variability of maximum strength performance as measured by the bench press lift of the Universal Gym Machine.

Delimitations of the Study

The present study was subject to the following delimitations:

---

[1] Chuck Coker, Coaches Training Manual (Fresno, California: Universal Athletic Sales Company, 1970), p. 20.

A. The selection of 100 students enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971.

B. The participation of the subjects in the experimental session once every other week.

C. The extent to which the selected instrument reliably and validly measures the desired characteristic in each subject. Empirical validity is accepted for the Universal Gym Machine.

D. The procedure providing for random selection of weight variation on the selected instrument.

E. The assumption that strength should not be changed by any activity engaged in by the subjects during the testing period.

F. The possibility of learning effects from the repeated testing.

Summary

Measures of all human phenomena vary, and the limits of such variation are not known for strength testing for women. It has been hypothesized that maximal strength effort of women seems to be unreliable because many women are believed to give up before a maximum effort is made. The present investigation is concerned with determining the variability in maximum strength performance in 100 women for one exercise, the bench press. Research in this area may result in knowledge which has practical application to strength testing for women.

The purpose of this study was to determine the extent of variability with women in repeated measures of maximum strength. The present investigator hypothesized that there is no significant difference in the inter-individual variability of maximum strength performance as measured by the bench press lift on the Universal Gym Machine; and there is no significant difference in the group variability of maximum strength performance as measured by the bench press lift of the Universal Gym Machine.

In Chapter II a review of selected literature will be presented.

CHAPTER II

REVIEW OF RELATED LITERATURE

The present investigation entailed a study of 100 subjects enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971, to determine the variability in maximum strength performance of the arm flexors and the muscles in the shoulder girdle. Maximum strength was measured on the Universal Gym Machine by use of the bench press lift every other week for a period of ten weeks. All subjects were unaware of their scores. A conclusion was drawn concerning the inter-individual and group variability of maximum testing in university women. A review of selected studies which were believed pertinent to the present study follows.

Kroll[1] conducted a study of twenty male subjects to determine the reliability of right wrist flexor strength in test-retest situations. Each subject was tested three days in succession, and five trials, spaced a minute apart, were given. The subjects were tested during a similar time of day, and exertion for any trial was limited to five

---

[1]Walter Kroll, "Reliability Variations of Strength in Test-Retest Situations," Research Quarterly, XXXIV (March, 1963), p. 50-55.

seconds. The subjects were retested three weeks later and again three months later. The apparatus used isolated action to the right wrist flexor muscle group.

Analysis of variance technique was used for the purpose of determining the experimental error components of variance and securing reliability estimates for each of the test conditions. The intraclass correlation technique was used also for each of the three test conditions.

> Significant differences for means of daily tensiometer strength scores between test condition one and both retest conditions affected test-retest reliability and suggested intrusion of non-pertinent research factors.[1]

It was suggested that these factors might be caused by a physiological response to initial measurement procedures, a learning effect, and/or a combination of these factors.

Kroll,[2] in 1962, completed a study similar to the previous one, to assess the reliability of a selected measure of human strength of fifty male subjects. The subjects were tested three days in succession at similar times of day. The apparatus used was designed to isolate action of the right wrist flexor muscle group. Standardized instructions were read at each testing period and five trials spaced a minute apart were given. Exertion on each trial was limited to five seconds.

---

[1] Ibid., p. 50.

[2] Walter Kroll, "Reliability of a Selected Measure of Human Strength," Research Quarterly, XXXIII (October, 1962), p. 410-417.

The intraclass correlation technique was used to assess the reliability of measurement with a resultant coefficent of 0.93. The investigator concluded that other factors intruded and tentatively identified these factors as fatigue, a day effect resulting from strength development and/or learning effect. Kroll suggested that these factors be taken into consideration in measurement programs of right wrist flexor strength.

In a third study of reliability utilizing the right elbow flexor muscles, Carlson and Kroll[1] conducted a study of thirty-six male college students at the University of Texas, Austin, Texas, who volunteered and were paid. They determined the reliability of isometric strength assessment by means of an intraclass correlation technique and estimated the reliability one might obtain under similar conditions with different combinations of days and trials. The subjects reported to the laboratory three times a week until sixteen testing sessions were completed. One half of these testing sessions were devoted to isometric contractions, while the remaining eight sessions contained isotonic contractions. So that no subject had the same order of testing, the method of exercise during the sixteen sessions was randomly assigned to each subject.

An analysis of variance model was used to obtain the

---

[1]Robert Carlson and Walter Kroll, "The Use of Analysis of Variance in Estimating Reliability of Isometric Elbow Flexion Strength," Research Quarterly, XLI (May, 1970), p. 129-134.

variability components necessary for computation of an estimate of reliability. The data obtained in this study yielded a coefficient of reliability of .99.

In order to study reliability, Meyers and Piscopo[1] constructed a test of the push-up action. The subjects were tested once a day at one week intervals for three weeks.

Both the cable tension test and the manuometer push apparatus were used. The Pearson Product-Moment Method was used, indicating a high reliability coefficient for the cable tensiometer. The investigators concluded that for the cable tension test there was a high reliability coefficient and there was no improvement between the second and third trials. On the manuometer test, the comparative inconsistency and increased correlation between tests two and three indicate the influence of the learning factor and perhaps the fatigue effect.

In another study, conducted by Alderman and Banfield,[2] the reliability of isometric strength of eight different measures was determined. The investigators found that the reliability coefficients ranged from moderately high (.74) to (.98).

To determine the reliability and errors of measure-.

---

[1]Carlton R. Meyers and John Piscopo, "Reliability Study of Cable Tension Strength as Compared to Manuometer Push Apparatus," Research Quarterly, XXXV (May, 1964), p. 213-214.

[2]Richard B. Alderman and Terry J. Banfield, "Reliability Estimation in the Measurement of Strength," Research Quarterly, XL (October, 1969), p. 449-455.

ment at several levels of strength, Kroll[1] conducted a study of seventy-five college females and thirty male subjects. Each subject made two visits to the laboratory one week apart. Five maximal isometric wrist flexion strength trials of five seconds duration and spaced one minute apart were performed by each subject on each visit. The subjects were divided into high, middle, and low levels of absolute strength groups. Intraclass correlation techniques were used to partition error variance and true score variance.

The smallest error variance estimates for trials and days on both wrists were demonstrated by the low strength females. The middle and high strength females were two to three times larger in error variance than the low strength females. The largest error variance was exhibited by the male subjects in the right wrist, some two to four times larger than the female subjects. Since the magnitude of the strength score seemed to be mildly but positively related to the error variance, the derived reliability coefficient can be expected to be an overestimate of the true reliability. The investigator also discussed the possibility that motivation in reliability assessment of maximal isometric strength may be exaggerated. This is surmised because the low strength females demonstrated lower error variance estimates than the male subjects or high strength females who exhibited larger error variance.

---

[1] Kroll, 1970, op. cit., p. 155-163.

A study on the effects of knowledge of performance on isometric strength test scores of fifteen paid medical students was conducted by Pierson and Rasch.[1] Each subject participated twice a day, five days a week, for two weeks. Once a day the subject was aware of his test results and once he was not aware of his results. Between each of the daily trials, a rest of not less than five minutes was taken. Isometric strength was tested by an apparatus consisting of a bar connected by a cable and adjustable link chain to a load cell which activated a dial indicator in view of the subject.

A split-plot Latin square experimental design was used in which priority of administration and the effects of day of week were balanced. With no knowledge of results, the mean strength score was 97.8 pounds with a standard deviation of 17.0 pounds. With knowledge of results, the mean strength score was 100.8 pounds with a standard deviation of 15.6 pounds. The correlation between the two test scores ($r=0.98$) was significant. An analysis of the effects of the day of week upon isometric strength indicated that knowledge of results or that performance without the knowledge of results did not make a significant difference. Pierson and Rasch concluded that when the subject has knowledge of his performance, isometric strength scores are greater, but

---

[1]William R. Pierson and Philip J. Rasch, "Effect of Knowledge Results on Isometric Strength Scores," Research Quarterly, XXXV (October, 1964), p. 313-315.

not to a significant degree, than when the subjects does not have the knowledge of his results.

Johnson and Nelson[1] conducted a study of 120 male subjects to investigate the effect of applying different motivational techniques during training and testing upon strength performance. The investigators found that motivational techniques during training promote significant strength gains. Hansen,[2] however, reported no significant differences for five selected motive-incentive conditions upon the effectiveness of a six-week isometric training program for the development of strength in the elbow flexor muscle group. The investigator stated that at the .05 level of confidence, no significant differences were found between the means of the scores.

Malina[3] states that test-retest sessions using both the product-moment correlation and analysis of variance techniques yielded different reliability estimates. The former essentially provides an estimate of performance stability over time, whereas the latter provides an estimate of the components underlying reliability and thus permits a more accurate estimate of test reliability.

---

[1]Johnson and Nelson, op. cit.

[2]Gary F. Hansen, "Effect of Selected Motive-Incentive Conditions upon Development of Strength Through an Isometric Training Program," Research Quarterly, XXXVIII (December, 1967), p. 585-592.

[3]Robert M. Malina, "Reliability of Different Methods of Scoring Throwing Accuracy," Research Quarterly, XXXIX (March, 1968), p. 149-160.

In a report by Henry[1] the test-retest was also discussed. Henry reported that the reliability coefficient is equal to the true score variance divided by the total variance. The former is inter-individual variance, the latter is the sum of inter-individual variance, intra-individual variance, and error of measurement. Measurement error is a characteristic of the test and may or may not be large enough to reduce the reliability coefficient appreciably. Variations between and within individuals characterize behavior, regardless of measurement error which may or may not be reliable.

In another study by Kroll[2] the relative effects of different measurement schedules upon strength development were reported. The investigator concluded that the varying number of trials on the first day had no significant effect upon the strength measures. Kroll,[3] in a later study, discussed the problem of selecting a criterion measure when several trials are available for reliability analysis. If no trend in trials was present, based upon reliability theory, the correct criterion measure was the mean of all available trials. When a trend was present, the recommended procedure

---

[1]F. M. Henry, "Reliability, Measurement Error, and Intra-Individual Difference," Research Quarterly, XXX (March, 1959), p. 21-24.

[2]Walter Kroll, "Reliable Method of Assessing Isometric Strength," Research Quarterly, XXXIV (October, 1963), p. 350-355.

[3]Walter Kroll, "Reliability Theory and Research Decision in Selection of a Criterion Score," Research Quarterly, XXXVII (October, 1967), p. 412-419.

was to have a measurement schedule free of systematic measurement error variance.

Summary

The present investigation entailed a study of 100 subjects enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971, to determine the variability in maximum strength performance of the arm flexors and shoulder girdle muscle groups. Maximum strength was measured on the Universal Gym Machine by use of the bench press lift every other week for a period of ten weeks for the experimental group. The control group was measured at the beginning and end of the study. The subjects were unaware of their scores. In this chapter the literature pertinent to this study was summarized.

Studies determining reliability of selected measures of human strength were discussed. The analysis of variance along with the intra-class correlation technique were the most widely used statistical treatments in assessing reliability estimates.

The studies reviewed for this thesis did not agree on the importance of motivation in testing for maximum strength. Kroll and Hansen suggested that motivation in reliability assessment of maximal isometric strength may be exaggerated. Johnson and Nelson, however, found that motivational techniques induced during testing for maximal strength is important with respect to performance.

A study on the effects of knowledge of performance

on isometric strength test scores was conducted by Pierson and Rasch. The investigators concluded that when the subject has knowledge of his performance, isometric strength scores are greater, but not to a significant degree, than when the subject does not have the knowledge of his results.

The test-retest situation was used commonly in the studies reviewed. The time interval and the number of trials in the studies differed. The time interval ranged from three months (Kroll, 1963) to three days (Kroll, 1970) to twenty (Pierson and Rasch). The present study used a time interval of ten weeks and five trials.

In Chapter III, the investigator will present the procedures followed in the development of the study.

CHAPTER III

PROCEDURES FOLLOWED IN THE DEVELOPMENT OF THE STUDY

The present investigation entailed a study of the variability in maximum strength performance of 100 undergraduate students enrolled in the Texas Woman's University in Denton, Texas, during the academic year of 1970-1971. The procedures will be reported under the following headings: preliminary procedures, sources of data, selection of subjects, selection of instrument, procedures followed in the collection of data, organization and treatment of data collected, and procedures followed in writing the final report. The chapter concludes with a brief summary.

Preliminary Procedures

Prior to the actual collection of data, a series of preliminary procedures were necessary. These procedures included surveying, studying, and assimilating all literature pertinent to the study; securing permission from the Dean of the College of Health, Physical Education and Recreation at the Texas Woman's University to conduct the study during the academic year of 1970-1971; developing and presenting a tentative outline of the study at a Graduate Seminar of the College of Health, Physical Education and Recreation at the

Texas Woman's University in Denton, Texas; revising the outline in accordance with the suggestions offered by members of the thesis committee, and filing a prospectus of the approved study in the Office of the Dean of Graduate Studies.

Sources of Data

Both human and documentary sources were utilized in the development of the present study. The human sources included 100 undergraduate students enrolled in the required physical education program at the Texas Woman's University in Denton, Texas, during the academic year of 1970-1971. Other human sources enlisted were members of the faculty at the Texas Woman's University who served upon the thesis committee and qualified persons who served as resource persons in specific aspects of the proposed study. The documentary sources consisted of books, pamphlets, periodicals, bulletins, research studies, microcards, theses, dissertations, and other unpublished materials.

Selection of Subjects

One criterion was established for use in the selection of subjects: each subject should be enrolled in the required physical education program at the Texas Woman's University in Denton, Texas, during the academic year of 1970-1971. One hundred subjects participated in the study from four service classes selected in an arbitrary manner prior to the registration period. The control group were women from golf

and archery classes, while the experimental group were women from track and field and body mechanics classes. The respective classes were selected by the subjects by personal choice during the registration period prior to each semester. No control of this selection was attempted.

Selection of Instrument

Prior to the selection of the instrument to be used to determine maximum strength, criteria were established. The criteria set for the present study were: availability of the instrument, administrative feasibility, objectivity, validity, and reliability. The Universal Gym Machine met each of the stated criterion and therefore was selected.

Procedures Followed in the Collection of Data

Prior to the administration of the test, data were obtained concerning the age, height, weight, and class or year classification. The information was collected in case there was a change during the testing with the control group. Menstrual data were collected also but were not included in this study because previous studies indicate that menstruation does not influence participation in physical activities.[1, 2]

---

[1] Frances A. Hellebrandt and Margaret Meyer, "Physiological Data Significant to Participation by Women in Physical Activities," Research Quarterly, X (March, 1939), p. 10-21.

[2] M. A. Garlic and E. M. Bernauer, "Exercise During the Menstrual Cycle: Variations in Physiological Baselines," Research Quarterly, XXXIX (October, 1968), p. 533-542.

A schedule was established for the administration of the tests. The control group was tested both at the beginning and the end of the experimental period. The experimental group was tested once every other week for ten weeks. Testing once every other week was chosen because it was believed that strength development would not take place from the test itself during this period. All data were collected between the period from September, 1970, through November, 1970.

Standard procedures were used in testing for maximum strength. On the initial test each subject was tested for maximum strength using increments of five pounds. The beginning weight was one half the subject's total weight. During the remaining four visits of the experimental group, the initial amount of weight (maximum strength) was varied randomly from minus twenty pounds to plus twenty pounds. Weight was added or subtracted until maximum strength test scores were achieved. The maximum score was assumed to be at the point that was last lifted when an additional five pounds could not be lifted. Between trials, subjects rested a minimum of three minutes.

Standardized instructions which requested full co-operation were read to every subject but the investigator made no further efforts to motivate the performance. The exact instructions used may be found in the appendices of the study. The subjects were not informed of their strength scores. The subjects provided their own motivation for each trial to achieve maximal strength scores. Each subject was tested for maximum strength each time approximately at the same time of day.

Organization and Treatment of Data Collected

Score cards were used for recording the data. The investigator recorded the raw scores yielded by the test chosen to measure maximum strength after each test. A copy of the score card may be found in the appendices of the study.

The investigator selected the two way analysis of variance for repeated measures model statistical techniques to determine test reliability. The mean, standard deviation, and standard error of the mean were also computed. The data were organized and presented in appropriate tables. The investigator then analyzed, interpreted, summarized and stated a conclusion to the study from the findings of the investigation.

Procedures Followed in Writing the Final Report

The investigator prepared and submitted the written report to members of the thesis committee for corrections, suggestions, and final approval. Revisions of the report were made in accordance with the recommendations made by members of the thesis committee and the final written report was prepared of the study as a whole, including the conclusion to the study, implications of the findings for physical education, recommendations for further studies, a classified bibliography, and appendices.

Summary

The procedures followed in the development of the study were presented in this chapter. Preliminary procedures involved the selection of the instrument and the selection of subjects.

Subjects for the study were 100 undergraduate students enrolled in the required physical education program at the Texas Woman's University in Denton, Texas, during the academic year of 1970-1971. The instrument selected was the Universal Gym Machine. The basic testing procedures used in testing for maximum strength were described prior to the initial test. Each subject was tested for maximum strength using increments of five pounds. During the remaining four visits, the initial amounts of weight were varied randomly from minus twenty pounds to plus twenty pounds. Weight was added or substracted until maximum strength test scores were achieved.

Procedures for analyzing the data consisted of the analysis of variance to determine test reliability. The final procedures included those related to determining a conclusion to the study, the implications of the findings for physical education, and writing the final report.

In Chapter IV the investigator will present and discuss the findings resulting from the analysis of the data.

CHAPTER IV

PRESENTATION AND INTERPRETATION OF THE DATA

Introduction

The present investigation entailed a study of 100 women enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971, to determine the variability in maximum strength testing of the arm flexors and shoulder girdle muscle groups. Maximum strength was measured on the Universal Gym Machine by use of the bench press lift. The experimental group was tested every other week for a period of ten weeks. The control group was tested at the beginning of the test period and again ten weeks later. The subjects were unaware of their performance. The data were tabulated, organized into tables, analyzed statistically through analysis of variance techniques, and interpreted for presentation in this chapter of the thesis. All raw data may be found in the appendices.

Findings of the Study

The findings of the study are presented in the following paragraphs. In Table 1 the ranges, means, standard deviations and standard errors of the means of the maximum strength scores are presented. A study of Table 1 reveals the range for the experimental group was 65 pounds and for

the control group was 45 pounds. The means were 70.06 pounds and 66.25 for the experimental group and control group respectively. The standard deviations were 12.81 and 9.14 for the experimental group and control group respectively. The standard error of the means was 1.28 for the experimental group and .578 for the control group. The greater variation for the experimental group was anticipated because they were measured five times in comparison to the control group that was only measured twice. It should be noted that five pound weight decrements were used when testing for maximum strength.

TABLE 1

RANGES, MEANS, STANDARD DEVIATIONS, AND STANDARD ERRORS OF THE MEANS FOR ALL MAXIMUM STRENGTH SCORES. (N=100)

Group	Range Lbs.	Mean Lbs.	SD Lbs.	SE_m
Experimental	65 (50-115)	70.06	12.81	1.28
Control	45 (45-90)	66.25	9.14	.578

In Table 2, the summary table for the analysis of variance of the experimental group is presented. A study of Table 2 reveals that there was no significant difference between the five trials in the experimental group as shown by an F ratio of 2.313 at the assigned confidence level (p=.05). This finding may be interpreted to mean that the subjects were putting forth a consistent effort on the test for maximum strength in all five trials and that the consistent effort was of maximum magnitude. The between trials

variance approached significance with an obtained value of 2.313. The F ratio would have to be 2.41 to be significant at the .05 level of confidence. Any variation that is not a chance occurrence, most likely could be attributed to learning effects, strength development, a combination of the two, and/or the social-cultural role of women phenomenon. A significant difference appeared between the subjects in the experimental group as shown by an F ratio of 58.76. This difference was anticipated and is considered normal because individuals do differ in maximum strength abilities.

TABLE 2

SUMMARY TABLE FOR ANALYSIS OF VARIANCE OF THE EXPERIMENTAL GROUP

Source	SS	df	ms	F	p
Between Trials	1226	4	306.5	2.313	ns
Between Subjects	381541	49	7786.551	58.76	.05
Subjects x Trials	25974	196	132.520		
Total	408741	249			

F (4, 196) (.05) = 2.41
F (49, 196) (.05) = 1.42

The control group was measured in their bench pressing ability twice, once at the beginning of the experimental period and again at the end of the period. A study of Table 3, summary table for the analysis of variance for the control group, reveals that there was no significant difference between the two trials in the control group as shown by an F ratio of 1.25 at the assigned significance

level of .05. This was interpreted as being indicative that the subjects were putting forth maximum effort on the test for maximum strength in both trials because of the consistency. There was a significant difference between the subjects as shown by an F ratio of 9.4188 at the assigned significance level of .05. This difference was anticipated and is considered normal because individuals differ in maximum strength abilities.

TABLE 3

SUMMARY TABLE FOR ANALYSIS OF VARIANCE OF THE CONTROL GROUP

Source	SS	df	ms	F	p
Between Trials	20.25	1	20.25	1.25	ns
Between Subjects	7456.25	49	152.168	9.419	.05
Subjects x Trials	792.25	49	16.16		
Total	8268.75	99			

F (1, 49) (.05) = 4.035

The maximum strength scores were submitted to further statistical treatment through analysis of variance to compare the total difference of maximum strength scores between the experimental and control subjects upon the first and last trials. In Table 4, the summary table for the analysis of variance between the experimental and the control group is presented. A study of Table 4 reveals that there was a significant difference between the experimental and the control group (F ratio 9.72) in the ability to lift weights in the bench press lift. The two groups were arbitrarly

determined, with those students selecting the archery and golf classes placed in the control group and those students enrolled in track and field class or body mechanics placed in the experimental group. This was done to facilitate testing procedure and to avoid having the subjects miss classes. It would appear that the subjects selecting the archery or golf classes to fulfill a university requirement, are not as strong as the students who selected the track and field or body mechanics classes. This difference does not effect the design or results of the study because variability was the measure considered rather than group differences.

There appears to be no interaction between the two groups. In other words, the relationship between the means tended to increase or decrease. There was no significant difference between the trials of subjects or the trials of the two groups as determined by evaluating F ratios of .0100 and .203 respectively. This indicates consistency of measurement. This finding may be interpreted as suggesting that the subjects in both groups were putting forth maximum effort on the test for maximum strength.

TABLE 4

SUMMARY TABLE FOR ANALYSIS OF VARIANCE BETWEEN THE EXPERIMENTAL GROUP AND THE CONTROL GROUP

Source	SS	df	ms	F	p
Between Subjects	4987.5	99	-	-	
Between Groups	450.	1	450.	9.72	.05
Error	4537.5	98	46.3	-	
Within Subjects	19600.	100	-	-	
Between Trials	2.	1	2.	.0100	ns
Trials x Groups	40.5	1	40.5	.203	ns
Error	19557.5	98	199.566		
Total	24587.5	199			

F (1, 98) (.05) =3.94

The estimates of the average reliability coefficients[1] for the experimental group and the control group of all trials are .99 and .89 respectively. This is a highly dependable relationship between the trials over the ten week experimental period. The investigator assumes than that there was no learning effect. The formula for this computation may be found in the appendices.

Tests of Hypotheses

Upon the basis of the results of the analysis of data through the application of the appropriate statistical test---analysis of variance, the hypotheses stated in the first chapter were examined. The results of the applied tests are presented on the next page.

---

[1] B. J. Winer, Statistical Principles in Experimental Design, (New York: McGraw-Hill Book Company, Inc., 1962), p. 127.

Hypothesis I

> There is no significant difference in the inter-individual variability of maximum strength performance as measured by the bench press lift on the Universal Gym Machine.

The data collected for this study failed to provide sufficient information to rejects the hypothesis.

Hypothesis II

> There is no significant difference in the group variability of maximum strength performance as measured by the bench press lift of the Universal Gym Machine.

The data collected for this study provided information that caused the investigator to reject the hypothesis. Variability was the parameter under study and the data previously presented indicated that the two groups were internally consistent. The difference that was found between the groups was most likely a result of the initial selection and placement of the students and is not an unusual or particularly important statistic.

Summary

The present investigation entailed a study of 100 women enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971, to determine the variability in maximum strength testing of the arm flexors and shoulder girdle muscle groups. Maximum strength for the experimental group was measured on the Universal Gym Machine by use of the bench press lift every other week for a period of ten weeks. The control group was tested twice, ten weeks apart. The subjects were unaware of their performance. The

data were tabulated, organized into tables, analyzed statistically, presented and interpreted.

Based upon the findings there was no significant difference in the inter-individual or the group variability of maximum strength testing as measured by the bench press lift of the Universal Gym Machine. This was believed to indicate that the subjects were putting forth maximum effort on the test for strength in both groups and in all the trials. There was a significant difference between the subjects of both the control and experimental groups. This was expected. Individuals differ in maximum strength. There was a greater variation in the experimental group which may be explained by the more trials that were given to the experimental group than to the control group.

Based upon the data collected for this study, the null hypothesis concerned with the inter-group variability of maximum strength performance was rejected. The group comparison found a significant difference which may be explained as a result of the initial arbitrary placement of the subjects into groups.

In Chapter V, the summary, conclusion to the study, and recommendations for further studies will be presented.

CHAPTER V

SUMMARY, CONCLUSION AND RECOMMENDATIONS FOR FURTHER STUDIES

Summary of the Investigation

The investigation entailed the collection, organization, and analysis of data to determine the variability of maximum strength of the arm flexors and shoulder girdle muscle groups, of 100 women enrolled at the Texas Woman's University, Denton, Texas, during the academic year of 1970-1971. Maximum strength was measured on the Universal Gym Machine by use of the bench press lift. The experimental group was measured every other week for a period of ten weeks. The control group was tested twice, once at the beginning and again at the end of the ten week period. The subjects were unaware of their performance. Maximal strength efforts for women seem to be unreliable because women are believed to give up before a maximum effort is made.

The maximum strength scores obtained from the administration of the bench press lift on the Universal Gym Machine were tabulated, treated statistically through analyses of variance technique, and analyzed. A conclusion concerning the variability in maximum strength testing of the arm flexors and shoulder girdle muscle groups of 100 women was determined.

The summary of the findings of the study, the conclusion, and recommendations for further studies are presented in this chapter.

Findings of the Study

The hypotheses that guided the present investigation stated that (1) there would be no significant differences in the inter-individual variability of maximum strength performance of university women, and (2) there would be no significant differences in the group variability of maximum strength testing of university women. The investigator failed to reject the first hypothesis. No significant difference in variability was found for either group.

Two-way analysis of variance for repeated measures indicated that both the control group which was tested only twice, once at the beginning and once at the end of a ten week period, and the experimental which was tested every other week for ten weeks did vary significantly. The rationale for this finding may be due to the arbitrary selection of the subjects. Students enrolled in golf and archery classes were placed in the control group while students enrolled in track and field or body mechanics classes became the experimental group. The subjects selecting golf and archery may have selected these classes due to their sedentary nature as opposed to track and field or body mechanics, thus presaging the group difference.

Conclusion of the Study

From the findings determined by this study, the following conclusion seems justified: confidence can be placed on maximum

strength testing scores of university women. The 100 university women at the Texas Woman's University who participated in this study did not vary significantly in their bench press ability and thus it is believed that the subjects did not give up before a maximum effort was made. The possibility of the subjects remembering any level of effect other than maximum is remote. It would seem that motivation is not a major factor for university women when testing for stability of maximum strength on the Universal Gym Machine.

Limitation of the Study

A limitation to the study is believed necessary to explain the fact that the control group and the experimental group were not equated in strength ability nor were randomization techniques applied.

Recommendations for Further Studies

The following recommendations for further studies are suggested:

1. A study to determine the variability in maximum strength testing of muscle groups other than the arm flexors and shoulder girdle of university women.
2. A study comparing physical education majors and non-majors with respect to variability in maximum strength testing of university women.
3. A study comparing the variability in maximum strength testing of university women before, during and after participation in selected recreational activities.

4. A study comparing the variability in maximum strength testing of university women and selected motivation techniques.
5. A study comparing the variability in maximum strength testing of university women with men.
6. A study comparing the variability in maximum strength testing and different age groups.
7. A study comparing the variability in maximum strength testing of different levels (high, middle, and low) of strength of university women.

Instructions

This is a test to determine the variability involved with strength testing. You are to press as much weight as you possibly can. There will be at least a two minute rest between trials to find the maximum amount of weight you can press.

Please have no fear of developing bulging muscles while participating in this experiment. Women's muscles are not made to develop as men's are. You would need to work out for several hours, several days a week to develop an increase in your muscle size.

Motivation to do your best has got to come from you. I will not be able to encourage you at anytime. So that you will not be motivated by the amount of weight you press, you will not know how much weight you do press.

RAW SCORES FOR THE EXPERIMENTAL GROUP

Subjects	Trials				
	1	2	3	4	5
S_1	65	70	75	65	70
S_2	50	50	55	55	55
S_3	90	105	100	105	110
S_4	65	60	60	65	65
S_5	60	60	65	65	60
S_6	90	90	90	95	85
S_7	115	105	115	110	115
S_8	80	75	80	80	75
S_9	75	85	80	85	85
S_{10}	60	60	55	60	60
S_{11}	60	60	60	65	60
S_{12}	75	75	80	75	75
S_{13}	80	85	80	80	80
S_{14}	60	65	65	70	65
S_{15}	60	55	60	60	55
S_{16}	70	65	70	70	70
S_{17}	60	65	60	65	60
S_{18}	65	70	65	65	70
S_{19}	60	55	60	65	55
S_{20}	65	60	60	60	60
S_{21}	60	60	65	65	65
S_{22}	70	70	75	80	70
S_{23}	80	80	70	80	75
S_{24}	70	65	70	65	65
S_{25}	80	85	75	85	75
S_{26}	55	55	55	55	50
S_{27}	60	60	60	60	60
S_{28}	50	55	55	55	55
S_{29}	65	60	70	60	55
S_{30}	60	70	65	70	65

RAW SCORES FOR THE EXPERIMENTAL GROUP --- CONTINUED

Subjects	Trials				
	1	2	3	4	5
S_{31}	75	75	75	75	75
S_{32}	85	85	85	80	80
S_{33}	85	85	85	80	80
S_{34}	80	95	100	90	90
S_{35}	50	50	50	50	50
S_{36}	75	70	75	70	70
S_{37}	75	75	70	70	70
S_{38}	70	60	60	60	65
S_{39}	75	75	80	70	60
S_{40}	65	65	70	65	65
S_{41}	80	80	75	80	75
S_{42}	65	65	60	60	60
S_{43}	75	80	80	70	75
S_{44}	55	60	60	55	55
S_{45}	75	75	70	75	70
S_{46}	65	70	55	70	65
S_{47}	55	60	65	60	60
S_{48}	75	80	70	70	75
S_{49}	60	60	60	60	60
S_{50}	85	85	90	90	80

RAW SCORES FOR THE CONTROL GROUP

Subjects:	Trial 1	Trial 2
S_1	65	65
S_2	70	75
S_3	65	55
S_4	70	60
S_5	70	75
S_6	65	65
S_7	70	75
S_8	45	50
S_9	60	60
S_{10}	70	65
S_{11}	60	60
S_{12}	65	60
S_{13}	65	70
S_{14}	70	65
S_{15}	75	75
S_{16}	55	60
S_{17}	55	45
S_{18}	70	65
S_{19}	55	55
S_{20}	60	60
S_{21}	65	65
S_{22}	65	70
S_{23}	65	65
S_{24}	80	75
S_{25}	55	55
S_{26}	60	50
S_{27}	70	75
S_{28}	65	60
S_{29}	65	70
S_{30}	60	60

RAW SCORES FOR THE CONTROL GROUP --- CONTINUED

Subjects	Trial 1	Trial 2
S_{31}	70	70
S_{32}	65	65
S_{33}	55	60
S_{34}	70	90
S_{35}	55	60
S_{36}	45	45
S_{37}	80	85
S_{38}	70	70
S_{39}	65	75
S_{40}	65	70
S_{41}	70	70
S_{42}	65	70
S_{43}	80	85
S_{44}	65	75
S_{45}	70	75
S_{46}	80	75
S_{47}	80	85
S_{48}	60	60
S_{49}	65	70
S_{50}	80	80

Individual Data Sheet

NAME: _____ AGE: _____

ACTIVITY CLASS & TIME: _____

CIRCLE ONE: FRESHMAN, SOPHOMORE, JUNIOR, SENIOR, GRADUATE

	Test 1	Test 2	Test 3	Test 4	Test 5
HEIGHT:					
WEIGHT:					
MENSTRUATION:					
PRE-PERIOD:					
DURING PERIOD:					
POST-PERIOD:					

Do not write below this line.

Test Raw Scores

Weight Pressed

Test 1: _____

Test 2: _____

Test 3: _____

Test 4: _____

Test 5: _____

FORMULA[1]

$$r = 1 - \frac{\text{MS within subjects}}{\text{MS between subjects}}$$

[1]Winer, op. cit., 124.

BIBLIOGRAPHY

Books

Coker, Chuck. Coaches Training Manual. Fresno, California: Universal Athletic Sales Company, 1970.

de Vries, Herbert. Physiology of Exercise. Dubuque, Iowa: Wm. C. Brown Company Publisher, 1966.

Johnson, Perry B.; Updrke, Wynn F.; Stolberg, Donald C.; and Schaefer, Maryellen. Physical Education: A Problem-Solving Approach to Health and Fitness. New York: Holt, Rinehart and Winston, 1966.

Mathews, Donald K. Measurement in Physical Education. Philadelphia: W. B. Saunders Company, 1968.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, Inc., 1962.

Periodicals

Alderman, Richard B. and Banfield, Terry J. "Reliability Estimation in the Measurement of Strength." Research Quarterly, XL (October, 1969), 449-455.

Carlson, Robert and Kroll, Walter. "The Use of Analysis of Variance in Estimating Reliability of Isometric Elbow Flexion Strength." Research Quarterly, XLI (May, 1970), 129-134.

Hansen, Gary F. "Effect of Selected Motive-Incentive Conditions upon Development of Strength Through an Isometric Training Program." Research Quarterly, XXXVIII (December, 1967), 585-592.

Henry, F. M. "Reliability, Measurement Error, and Intra-Individual Differences." Research Quarterly, XXX (March, 1959), 21-24.

Johnson, Barry L. and Nelson, Jack K. "Effect of Different Motivational Techniques During Training and in Testing upon Strength Performance." Research Quarterly, XXXVIII (December, 1967), 630.

Kroll, Walter. "Reliability of a Selected Measure of Human Strength." Research Quarterly, XXXIII (October, 1962), 410-417.

_____. "Reliability Variations of Strength in Test-Retest Situations." Research Quarterly, XXXIV (March, 1963), 50-55.

_____. "Reliable Method of Assessing Isometric Strength." Research Quarterly, XXXIV (October, 1963), 350-355.

_____. "Reliability Theory and Research Decision in Selection of a Criterion Score." Research Quarterly, XXXVII (October, 1967), 412-419.

_____. "Test Reliability and Errors of Measurement at Several Levels of Absolute Isometric Strength." Research Quarterly, XLI (May, 1970), 155.

Malina, Robert M. "Reliability of Different Methods of Scoring Throwing Accuracy." Research Quarterly, XXXIX (March, 1968), 149-160.

Meyers, Carlton R. and Piscopo, John. "Reliability Study of Cable Tension Strength as Compared to Manuometer Push Apparatus." Research Quarterly, XXXV (May, 1964), 213-214.

Pierson, William R. and Rasch, Philip J. "Effect of Knowledge Results on Isometric Strength Scores." Research Quarterly, XXXV (October, 1964), 313-315.