THE DEVELOPMENT OF A SINGLE-ITEM TEST AS A MEASURE OF SOCCER SKILL

by

JOSEPH ROBERT JOHNSON

B.P.E. University of British Columbia, 1962.

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF PHYSICAL EDUCATION

We accept this thesis as conforming to the
required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April, 1963.

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Physical Education

The University of British Columbia,
Vancouver 8, Canada.

Date 6th May, 1963

ABSTRACT

The purpose of this study was to develop and establish a wall-volley type test as a method of measuring soccer skill. The subjects were students in attendance at The University of British Columbia. The test was administered to 75 students who represented five distinct soccer ability groups: The Thunderbirds (Varsity first team), the Chiefs (Varsity second team), the Braves (Varsity third team), a Physical Education major class, and a Service Programme class. Each group consisted of 15 subjects.

Subjects were rank-ordered by the experimenter according to ability. The test required each subject to perform the wall volley test of three 30-second trials. The subject's aggregate score was correlated against the experimenter's rank ordering of players.

This test differed from previous wall-volley type tests of soccer ability in the dimensions of the target area; the distance of the restraining line; the use of a moving ball at the commencement of the test, and in the method of scoring.

Test results proved satisfactory, and the test suggests itself as a speedy, economic means of evaluating the soccer skill of players by coaches and physical educators. It was noted from repeated testing that subjects perform better after at least two practice sessions.

TABLE OF CONTENTS

CHAPTER 1

STATEMENT OF PROBLEM

Physical educators have long realized the value of skill tests as a means of more objective measurement, and they have also recognized the desirability of employing scientific and sound procedures to obtain better results in their work. However, although the area of tests and measurement is given lip service as being essential in the evaluation of physical education programmes, many physical educators resist the use of tests on the grounds of unreliability or non-validity. Many teachers also feel that the administration of tests is too time consuming and requires too much supervisory control.

In 1933, one of North America's leading physical educators, J. B. Nash, (1) wrote "the game of soccer is rapidly gaining acceptance across North America and has already replaced football in many of our schools." This statement reveals the popularity of soccer thirty years ago. Today the game enjoys considerable acceptance in all forms of educational institutions: the elementary school, junior and senior high schools, and colleges and universities.

In Canada, as a result of the Duke of Edinburgh's address to the Canadian Medical Association Convention in 1959, and more recently, the introduction by the Federal Government of Bill C-131 (2) directed towards the promotion of National Fitness and Amateur Sport, greater interest has been aroused in programmes of physical education and subsequently in the objective evaluation of such programmes.

The problem here was to provide a quick, reliable and economic method of testing the soccer skill of University students. There was also the necessity of providing a test which would serve the layman soccer coach as well as the physical education teacher.

It was hypothesized then, that a wall volley-type test can be developed which will effectively measure soccer skill, and provide an objective means of classifying players.

Statement of Problem: The problem is to develop a test of general ball control ability using as nearly as possible actual game skills. The purpose of this test is to measure the soccer skill of the individual and to serve as a means for classifying players.

Sub-Problems Were:

(a) To decide the dimensions of the target area.

(b) To decide the distance of the restraining line from the target area.

(c) To determine the method of administering the test.

(d) To develop an interim scoring scale as a measure of the soccer skill of players.

(e) To determine the effects of practice or learning by repeated re-testing.

REFERENCES

1. Caswell, John, E., "Preface", Soccer for High Schools, A. S. Barnes & Co., New York, 1933.

2. Bill C-131, An Act to encourage Fitness and Amateur Sport. As passed by the House of Commons, 25th September, 1961. Queen's Printers, Ottawa, 1961.

CHAPTER II

JUSTIFICATION OF PROBLEM

Relatively few tests have been devised to measure ability in soccer. Most of the existing tests are in battery form, with their concomitant disadvantages of requiring too much space, time, equipment and supervision. While some battery-type tests have provided a reasonable basis for predicting or measuring soccer ability, they have enjoyed little acceptance among teachers, coaches, and physical educators. Voltmer and Esslinger (1) state, "A vast majority of physical educators today do not share this enthusiasm for testing and are prone to resist attempts to set up such programmes." Their chief criticisms are that many of the available tests are neither reliable nor valid and that a testing programme to meet their needs would involve too much time and, in many cases, too much expense.

However, teachers and physical educators do prefer to objectively evaluate their work when possible, and it is with this in mind that this test has been developed.

REFERENCES

1. Voltmer, E. F. and Esslinger, A.A., The Organization and Administration of Physical Education, 3rd Ed. (Appleton-Century-Crofts, Inc., New York, 1958, p.505.

CHAPTER III

REVIEW OF LITERATURE

Evelyn F. Schaufele (1) studied the value of objective tests for girls in determining soccer ability. Her tests were given to a group of eighty-four girls in the ninth and tenth grades in the schools of Fairview Village, Ohio, in the fall of 1938. Criteria were secured by having each girl rated subjectively by three instructors and two senior students, the observations being carried out during two class periods. A conclusion was that the best single test for measuring soccer ability of high school girls was that of wall-volleying, which had a correlation with the criteria of .57. Schaufele also correlated her wall-volleying test with the sum of several other tests which had been T-scored. Her figures showed that the test ranked third in correlation with subjective criteria and first with the combined tests criteria having a relationship of .77 by rectilinear correlation.

Schaufele's wall-volley consisted of rebounding a ball against a wall 15' wide by 10' high. She used an initial restraining line of 15 feet, but once the test had begun, the restraining line did not enter into the test; however, she did set limits on the forecourt area which was 30 feet square. Each subject had two trials of one minute duration and only the best score was counted.

Marjorie L. Heath and Elizabeth J. Rogers (2) made a study on the use of knowledge and skill tests in soccer, using as subjects for the experiment unskilled children in the fifth and sixth grade.

The skill tests were:

1. Soccer dribble test.
2. Throw-in for accuracy.
3. Kicking a dead ball for distance.
4. Scoring goals, kicking a dead ball from the penalty spot.
5. Knowledge tests.

Results were:

1. In grade V the correlation between the composite T-score of the soccer skill test and the judgement rating was .602.
2. In grade VI the relationship between the composite T-score and judged playing ability was .624.

Bontz (3) developed a dribble and shoot test for fifth and sixth grade children. The test was given to 142 girls. In a review of the literature and from questionnaire returns she established that kicking constituted the basic skill of soccer and the skill most often tested and taught. Her test proved rather unwieldy as she required four successful runs for: dribbling, passing, and scoring. The child used first the left foot and then the right foot before the test was completed. (See diagram No. 1). The test suffered because of the necessity of successful completion, and involved too much time. The validity of the test was reported as .58 by correlating player's times with player's ratings. The reliability coefficients were calculated by using the odd and even performances of the player's left and right feet. These were given as .85 and .91 respectively.

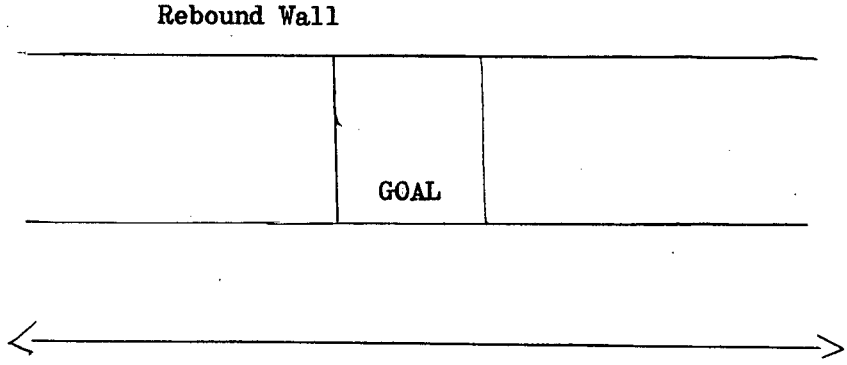


Diagram I

Crawford (4) revised the Schaufele wall-volley test for use with women majoring in Physical Education at the University of Oregon. Crawford did not bother with a restraining line, but supplied a retriever service to retrieve the ball in the event of miskicks, poor kicks and misses. Crawford found a validity coefficient of only .252 by correlating the rating of three judges with the initial test results. Crawford had more success with a rebound and trap test which she devised. The ball had to be rebounded from a wall and trapped behind an 8 foot restraining line. She reported a validity coefficient of .450 between the best score of three trials and the judges' ratings, and a validity coefficient of .537 between the best score of the three trials and the total test criteria. The test reliability was .704 using the second and third trial scores.

In 1950 Konstantinov (5) experimented at Springfield College with the purpose of developing and evaluating a battery of soccer skills as an index of ability in the game of soccer. His tests, developed by means of expert opinion and composite score criteria, were to be used for both classification and diagnosis. Data were

obtained from the results of testing of seventy-four varsity, freshmen varsity, and soccer class members. He used factor analysis to reveal that as far as the tests used in his study were concerned, there were three fundamental factors related to soccer - skill, power, and speed.

Vanderhoof (6) devised a battery test consisting of ten items: dribbling, trapping, throwing, tackling, place-kicking, volleying, corner-kicking and goalkeeping. She did not give any figures of student performance but suggested that the tests be considered as measures, or be evaluative, of players performance in each of these areas.

Winterbottom (7) tried to test three skills: place-kicking using a moving ball, accuracy in kicking using a moving ball, and controlled heading ability. He tested sixty of the top professional soccer players in England, but the average results given by him for five kicks with the left foot in each category (2 out of 5) would indicate that either the test itself was too difficult, or the players were not motivated to perform to the best of their ability. This was also true of his two heading tests which produced low average results of one in five.

MacDonald (8) experimented with wall-volley testing by reducing the restraining line distance from an initial trial distance of 30 feet to one of 9 feet. MacDonald's target area was 30 feet wide and $11\frac{1}{2}$ feet high. The players started the test with a stationary ball from behind the restraining line and two spare balls were placed 9 feet behind the restraining line in the centre of the testing area.

Use of the spare balls allowed the subject to make a recovery in the event of his losing the original ball. MacDonald allowed the subjects to control the rebounding ball behind the restraining line in any manner possible including use of the hands. In the event of the subject retrieving a poorly hit ball or selecting a spare ball, he was permitted to use his hands to return the ball to the starting position behind the restraining line and continue the test. Each subject was allowed four trials of 30 seconds each. The score was the highest score of any three trials.

MacDonald tested three groups of varsity players engaged in soccer. He used the subjective ratings of the three coaches of the three groups against their performance scores in computing the validity of the test. He obtained the following coefficients of correlation.

Number	Group	Correlation with Subjective Rating
17	Varsity Team	.94
18	Junior Varsity Players	.63
18	Freshmen Players	.76
53	Combined Groups	.85

Mitchell (9) used a revision of the MacDonald wall-volley test to determine the suitability of wall-volley testing as a technique for evaluating the soccer ability of grade five and six elementary school boys in West Vancouver, British Columbia. A restraining line of 6 feet and a target area of eight feet long by four feet high was used. Three trials of 20 seconds, and use of the spare balls technique

and a retriever service were employed. The subjects started the test by kicking a stationary ball from behind the restraining line. No use of the hands was permitted in recovering a badly played ball. Mitchell correlated a coaches' rating with test and re-test performance scores. The test and re-test were administered to six groups on the same day. Mitchell used combinations of trial scores and correlated these with the coaches' ratings. The following validity coefficients of correlation using two methods were obtained.

Method	Groups						Mean
	1	2	3	4	5	6	
Rank Difference	.864	.859	.821	.846	.825	.841	.84
Product Moment	.768	.808	.699	.813	.748	.717	.76

The wall-volley testing technique has also been used to measure ability in other sports such as tennis, badminton, volleyball and handball.

Miller (10) made use of the wall-volley technique to devise a badminton test as a measure of badminton ability. A restraining line at 10 feet and a rebounding area extending above a line drawn on the wall at a height of $7\frac{1}{2}$ feet. Three 30-second volleys or rallies were allowed and the score consisted of the sum of the three trials. The test reliability was determined by the test-retest method and found to be .94, while the test validity of .83 was determined by correlating the test scores with the results of a round-robin tournament.

Dyer (11) used a similar technique in setting up a test to measure tennis ability and also as a means of classifying subjects. A restraining line of 5 feet from the base of the backboard which was 10 feet high by 15 feet wide, with a line drawn at a height of 3 feet to represent the net. A box of extra balls was provided at the side of the restraining line. The total score of three 30-second trials was taken.

The test reliability was found by the test-retest method to be .86, while the validity of the test was determined by correlating the test scores with subjective judgements of three experts, and also by correlating the best scores with standings obtained by the subjects in several round-robin tournaments. The first of these methods revealed a validity coefficient of correlation of .85, while the second method produced coefficients ranging from .85 to .92.

Brady (12) made use of the wall-volley as a means of classifying and grading college men in volleyball. Brady experimented with several test items but found the wall-volley test to be most valid. No restraining line was used, but a smooth rebounding area was necessary. A line 5 feet long and $11\frac{1}{2}$ feet high was drawn on the wall, and vertical lines extended towards the ceiling at the ends of the horizontal line. The subject had to volley the ball for one minute. The test was begun by throwing the ball against the rebounding area. Only legal volleys counted. The test-retest method revealed a reliability of .925 (for 282 subjects), while the test validity of .86 was determined by correlating the test scores against the subjective judgements of 4 judges.

Russell and Lange (13) also made use of a wall-volley test to measure volleyball ability in junior high school girls. Their test was really a modification of an earlier test developed by French and Cooper (14). They used a restraining line of 3 feet, and a rebounding area 10 feet wide and $7\frac{1}{2}$ feet high. The score was the total number of legal hits in three 30-second trials. The test reliability, using the test-retest method, was found to be .87, while the test validity of .80 was obtained by correlating the test scores of the subjects with the subjective ratings of seven judges. French and Cooper used the same technique, except that they allowed ten 15-second trials. The score taken was the sum of the five best trials.

Cornish (15) devised a series of tests to measure handball ability including the 30-second wall-volley. The ball was served from the service zone. The total number of rebounds across the service line was counted. In the event of a ball getting out of control, a judge handed the contestant another ball. Combining the 30-second volley with the Service Placement test provided the best coefficient (.667) when correlated with the criterion (the subject's total points score in relation to his opponents after 23 games).

REFERENCES

1. Schaufele, Evelyn, F., "The Establishment of Objective Tests for Girls of the Ninth and Tenth Grades to Determine Soccer Ability". (Unpublished Master's Thesis Abstract, Springfield College, 1951).

2. Heath, Marjorie, L. and Rogers, Elizabeth, G., "A Study in the Use of Knowledge and Skill Tests in Soccer", Research Quarterly, vol. 33, no. 3, December, 1932.

3. Bontz, Jean, "An Experiment in the Construction of a Test for Measuring Ability in Some of the Fundamental Skills Used by Fifth and Sixth Grade Children in Soccer". (Unpublished Master's Thesis, State University of Iowa, 1942).

4. Crawford, Elinor, A., "The Development of Skill Test Batteries for Evaluating the Ability of Women Physical Education Major Students in Soccer and Speedball". (Unpublished Doctoral Thesis, University of Oregon, 1958).

5. Konstantinov, J., (Unpublished Master's Thesis, Springfield College, 1950).

6. Vanderhoof, Mildred, "Soccer Skill Tests", Journal of Health and Physical Education, vol. 42, no. 3, October, 1932.

7. Winterbottom, Walter, Soccer Coaching, revised ed., Naldrett Press Ltd., London, 1960, pp. 243-244.

8. MacDonald, Lloyd, G., "The Construction of a Kicking Skill Test as an Index of General Soccer Ability". (Unpublished Master's Thesis Abstract, Springfield College, 1951).

9. Mitchell, Reid, "A Wall-Volley Test for Measuring Soccer Ability in Fifth and Sixth Grade Boys". (Unpublished Master's Thesis, University of Oregon, 1963).

10. MacDonald, op.cit.

11. Dyer, Joanna, T., "Revision of Backboard Test of Tennis Ability", Research Quarterly, vol. 9, no. 1, March, 1938.

12. Brady, George, F., "Preliminary Investigations of Volleyball Playing Ability", Research Quarterly, vol. 16, no. 1, March, 1945.

13. Russell, Naomi, and Elizabeth Lange, "Achievement Tests in Volleyball for Junior High School Girls", Research Quarterly, vol. 2, no. 4, December, 1940.

14. French, Esther, L., and Cooper, Bernice, "Achievement Tests in Volleyball for High School Girls", Research Quarterly, vol. 8, no. 2, May, 1937.

15. Cornish, Clifford, "A Study of Measurement of Ability in Handball", Research Quarterly, vol. 20, no. 2, May, 1949.

CHAPTER IV

METHODS AND PROCEDURES

An appropriate test of soccer skill should approximate as nearly as possible the elements present in the actual soccer game: shooting, passing, dribbling, tackling, trapping and heading. A game-like situation requires that the aforementioned elements be performed whilst the ball is in motion and an "under pressure" situation in existence. Many tests fail to meet either of these criteria. Finney (1), Meisl (2), and Czaknady (3), who played representative soccer for their respective countries in international competition, each list the aforementioned skills as prerequisites of the accomplished soccer player. A poll of the opinions of the top four clubs in the Vancouver, B.C. Pacific Coast Soccer League (Canadians, Columbus, Firefighters and Victoria United) revealed that they rated the six foregoing skills above all others in determining soccer ability.

We may visualize each of these elements entering into a wall-volley test situation: shooting to make the ball strike the target area sooner; passing to make sure that the ball strikes the target area; trapping by controlling an awkwardly bouncing ball; tackling by intercepting and controlling a difficult return; dribbling by returning a rebounding ball quickly to the restraining line, and, of course, heading a high return.

In a review of previous tests which attempted to measure soccer skill, the wall-volley-type test came closest to approximating

"under pressure" conditions of the game situation. The changes introduced in this study were:

(a) The size of the target area - For logical reasons, the target area decided upon was the same as the regulation goal measurement (24 feet by 8 feet). This size was chosen as one that would be familiar to all players, and an area towards which all players would be accustomed to playing a ball.

(b) The distance of the restraining line - The restraining line distance of 15 feet was decided upon after experimentation with restraining lines of 30 feet and 24 feet.

(c) Starting the test by kicking a moving ball - Previous wall-volley tests had used a dead ball start behind the restraining line. Other battery tests had used a ball thrown or rolled by the researcher. It was felt that the ball should not be dead, nor should it be subject to variations in velocity, height, bounce or angle of arrival, so it was decided that the player should hold the ball in his hands behind the restraining line. On the command "go" he should put it into play as quickly as possible, thus, leaving the control of the ball entirely to the individual.

(d) The use of spare balls - Subjects were not to be penalized for using spare balls.

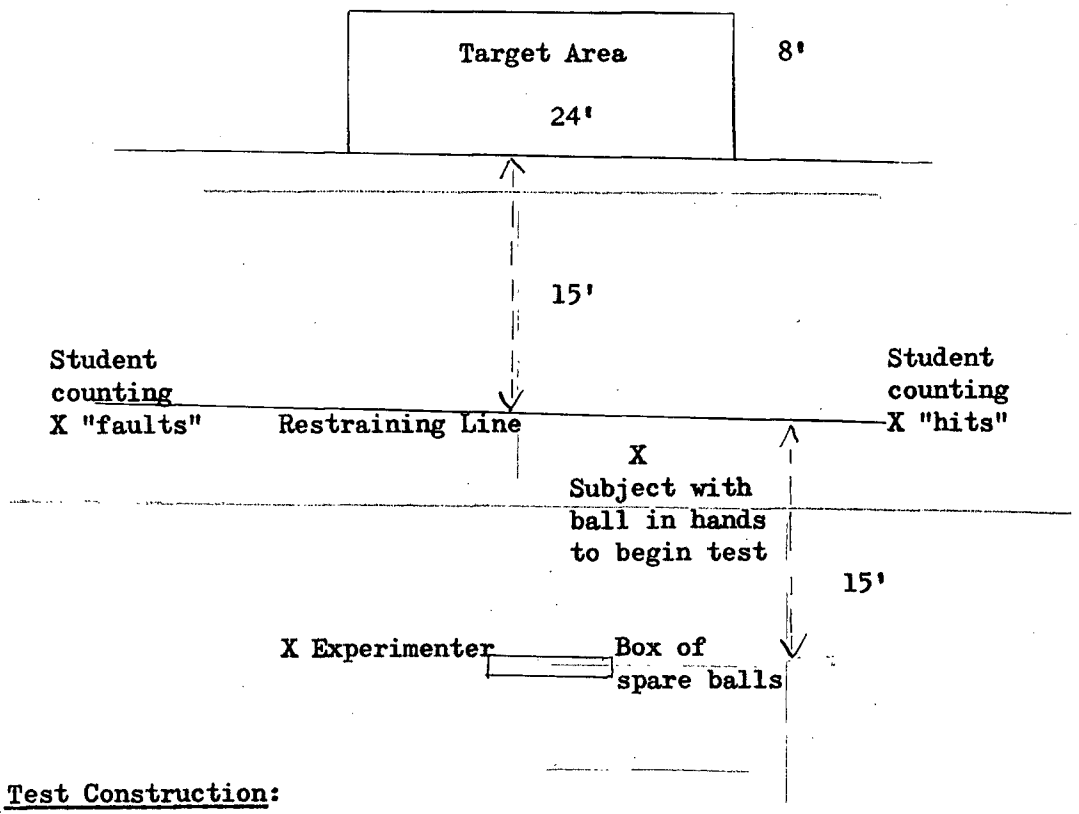
(e) The method of scoring - The aggregate score to count rather than the best trial or best two trials.

A rank ordering of player technique was adopted and rated

by the investigator. Weiss and Scott (4) state "Judges ratings can be a satisfactory criterion if the judges are competent and well trained, and if they have an adequate chance to observe before rating." In an attempt to overcome any element of subjectivity which may have been occasioned by the system of rank ordering, the players were rated after observing them play twelve soccer games with their respective teams. In assessing the rating of players whom the experimenter subjectively tie-ranked, a "checklist" technique was used. Players who could shoot hard and accurately with both feet were rated higher than those who could only shoot well with one foot. Players who could pass accurately were preferred to those who gave careless or incompleted passes. Players who made interceptions or successful tackles were rated above those who were poor at intercepting or tackling. Players who completed successful dribbles when forced to do so were preferred to those who were not successful. Players who revealed good trapping control were rated above those who did not, and players who were successful in heading a ball when challenged by an opponent were rated superior to those who were not.

The test was constructed so that it would demand the effective combined use of the six basic skills of shooting, passing, trapping, tackling, heading and dribbling, all performed "under pressure".

DIAGRAM 2



Test Construction:

(a) The target area was set at regulation goal size (24' x 8' high).

(b) The restraining line was fixed at 15' and parallel to the base of the target area.

(c) The box of spare balls was kept at a distance of 15' behind the centre of the restraining line.

(d) The decision to use a spare ball was left to the player (subject).

(e) The subject to start the test stood behind the centre of the restraining line facing the target area, holding a ball in both hands at waist height.

(f) The aggregate score of three 30-second trials counted as the subject's score.

(g) The experimenter gave the commands "go" and "stop" and took the time on a calibrated stop watch.

(h) A student counted the total number of hits to cross the restraining line during each trial. ("Hits": any ball played correctly by the subject from behind the restraining line, striking the target area and rebounding across the restraining line).

(i) A student counted the number of "faults" made by the subject during each trial, and these were deducted from the total number of rebounds noted by the first student. ("Fault": any infringement of a "hit" made by a subject within the trial limit).

(j) A third student provided a retriever service by returning a ball abandoned by the subject to the spare ball box.

(k) Only regulation size rubber soccer balls, inflated to 12 lbs. pressure, were used.

The wall-volley test was developed by the experimenter to serve as a means of measuring soccer skill, and also as a method of grading or classifying large groups of players. The test illustrated that five distinct categories of soccer ability may be revealed: superior, good, average, below average, and poor. While in this study the test was administered to university students, it is believed that its administration will separate subjects of all ages into these five categories.

Administration of the Test:

1. The test consists of three 30-second trials of rebounding a

soccer ball from behind a restraining line at a distance of 15' against a target area of 24' long and 8' high. The forecourt surface is tarmacadam and dry.

2. To start the test the subject stands behind the restraining line facing the target area and holds the ball in both hands at waist height.

3. A box of spare balls is provided 15 feet behind the centre of the restraining line.

4. On the command "go" the subject drops the ball from his hands and commences to rebound the ball against the target area as often as possible within the 30-second trial period.

5. The ball may be directed by the foot, leg, knee, or other part of the body except the hands or arms, to the area marked on the target area.

6. A ball that does not rebound across the restraining line can be retrieved by the subject by dribbling it across the restraining line, or by rebounding it from the target area across the restraining line. In either case the incompleted rebound would not count.

7. The use of the hands at any time to steady or to retrieve the ball is not allowed.

8. On the command "stop" at the end of 30-seconds, the subject's score is the total number of correctly completed rebounds.

9. A retrial is to be given the subject in the event that the retriever (student) interferes in any way with the subject's performance.

10. The three trials are to be carried out with a minimum of delay between trials.

Instructions to Subjects:

1. On the command "go", start the test immediately. Drop the ball; it need not bounce before you play it against the target area. Continue to play the ball to the target area until the command "stop" at the end of 30-seconds.
2. You may use any skill or combination of skills. You must play all balls from behind this restraining line (indicate the line clearly).
3. You may cross the line to retrieve the ball, but any "hits" made in such a position do not count. You may use any number of balls. If for any reason you lose close contact with the ball in play, do not try to retrieve it. Take another ball from this box (indicate clearly) and put it in play as you did at the start.
4. Each ball striking the wall in the marked area and returning over the restraining line before the word "stop" counts as a hit and scores one point.
5. You will each be given three trials today. The final score on the test is the sum of the scores on the three trials.

The following points are to be demonstrated:

1. One ball in hands.
2. Start test by dropping ball, then play it.
3. Demonstrate a few times, showing various skills: side-foot, instep, knee, thigh and head volley.

4. Cross restraining line to retrieve a ball, make a low hit to keep it in play, and retreat for next shot.
5. Make a wild shot to show how taking another ball saves time. Put this new ball in play as at the start.

Read the following paragraph to make certain that each person understands the test procedure and his duties.

No. 1 takes the test. At the signal "ready" he stands in the centre behind the restraining line facing the target area with a ball in his hands, prepared to start the test at the word "go". No. 2 counts the number of balls which strike the wall in the marked area and recross the restraining line before the word "stop", and enters them on the score card opposite the approximate trial number. If any infringements are reported by No. 3, these are deducted before the score for the trial is recorded. A ball recrossing the restraining line coincident with the word "stop" counts. No. 3 watches the player in relation to the restraining line. He reports to the scorer at the end of the trial the number of hits, if any, made while the player was standing closer to the wall than the restraining line. No. 4 collects the balls before the start of a trial and puts them in the box. During the trial he retrieves and returns to the box any balls going out of play.

Method of Scoring

Dr. McCloy (5) suggested in the revision of the backboard test of tennis ability that the aggregate total should be each player's score rather than simply taking the best of three trials.

He suggested that the time lost in putting another ball in play might cause sufficient reduction in score without any additional penalties, such as subtracting the number of extra balls used from the number of hits scored. This method was adopted in developing the soccer test described in this study.

Gathering the Data

Five groups of students playing soccer at the University at the time of this study were selected as the experimental groups. The five groups were the three university representative soccer teams: The Thunderbirds, the Chiefs and the Braves; a Physical Education Major soccer class, and a Required Programme soccer class. The experimenter, through a programme of visitation during class periods, observed these groups playing soccer on no fewer than twelve occasions. Following this period of observation, a rank order was determined for the members of each group. Players of each group were ranked in order of ability, from one to fifteen by the experimenter. Goalkeepers were excluded from rank orders. Following the administration of the test the players were rank-ordered again from one to fifteen on the basis of their scores. The groups were tested within a one-week period between 1:00 and 2:00 P.M. on separate dry days, thus standardizing ground and atmospheric conditions.

All players wore running shoes. The test was conducted on a blacktop area against a smooth cement wall 8 feet high and 24 feet wide. This area (same dimensions as a goal area) was clearly marked. A restraining line was drawn 15 feet from, and parallel to, the base

of the wall, and was clearly visible.

Treatment of Data

The test scores of each group were rank-ordered from one to fifteen, or from highest to lowest. Using the rank difference method of correlation, the test scores of each group were compared with the experimenter's pre-test rating of the players in each group. The degree of relationship between these two rank orders was determined by the coefficient of correlation obtained by using the Spearman Brown Rank Order formula:

$$r = 1 - \frac{6 \times d^2}{N(N-1)}$$

Thus, five rank difference coefficients of correlation were obtained; one for each group.

A total group validity coefficient of correlation was obtained by using the formula for quintiserial correlation as outlined by Jaspen (6):

$$\text{quint } r = \frac{ZaYa(Zb-Za)\ Yb+(Zc-Zb)\ Yc+(Zd-Zc)\ Yd-ZdYe}{6y\ \frac{Za^2}{a}+\frac{(Zb-Za)^2}{b}+\frac{(Zc-Zb)^2}{c}+\frac{(Zd-Zc)^2}{d}+\frac{Zd^2}{e}}$$

The reliability of the test was found by comparing the scores obtained on each of the three trials. Thus, three reliability correlation coefficients were obtained, i.e. by comparing the scores of the first trial with the scores obtained in the third trial, by comparing the scores obtained on the second trial with those obtained on the third trial, and lastly, by comparing the scores

obtained on the first trial with those obtained on the second trial.

To obtain these coefficients of correlation, the following two formulas were used:

(a) The Pearson Product Moment Formula, and

(b) The Spearman Brown Prophecy Formula.

Garrett (7) states "Increasing the length of a test or averaging the scores obtained from several applications of the test or from parallel forms will increase reliability. Fortunately a good estimate of the effect of lengthening or repeating a test can be obtained by use of the Spearman Brown Prophecy formula."

The standard deviation, the mean, the median and the aggregate scores were also calculated for each group and also for the total group of 75 subjects.

REFERENCES

1. Finney, Thomas, Instructions to Young Footballers, Museum Press Ltd., London, 1955, pp. 16-34.

2. Meisl, W., Soccer Revolution, Phoenix Sports Club, London, 1955, pp. 20-34.

3. Czaknady, Jeno, Learn to Play the Hungarian Way, Hungarian Sport Publishing House, Budapest, 1954, pp. 17-71.

4. Weiss, Raymond, A., and Scott, Gladys, M., "Construction of Tests". Research Methods, 2nd ed., American Association for Health, Physical Education and Recreation, Washington, D.C., 1959, pp. 239-240.

5. Dyer, Joanna, T., "Revision of the Backboard Test of Tennis Ability", Research Quarterly, vol. 9, no. 1, March 1938, p. 27.

6. Jaspen, Nathan, "Serial Correlation", Psychometrika, March, 1946, pp. 23-30.

7. Garrett, Henry, E., Statistics in Psychology and Education, 5th ed., Longmans, Green, and Co., New York, 1958, p. 343.

CHAPTER V

PRESENTATION OF DATA

The five groups - the Thunderbirds, the Chiefs, the Braves, the Physical Education Major Soccer Class, and a Required Programme Soccer Class were all given the wall-volley test as described in Chapter 3. A total of 75 subjects were tested.

There were 15 subjects in each group and each group was rank-ordered by the experimenter from 1-15. The scores of each subject and of each group are indicated in Tables 1 to 5 of the appendix.

The range of scores of each group was:

Thunderbirds	Chiefs	Braves	P.E. Majors	Required Programme
58 - 40	50 - 27	47 - 25	34 - 21	36 - 16
(18)	(23)	(22)	(13)	(21)

TABLE 1

The mean, median and standard deviation scores of each group were:

Group	Mean Score	Median Score	Standard Deviation
Thunderbirds	48.00	47.00	4.99
Chiefs	36.93	38.00	6.26
Braves	32.60	31.00	6.34
P.E. Majors	25.66	24.50	4.59
Req'd. Programme	24.20	23.50	5.76

Coefficients of correlation for the test validity were obtained for the five groups by using the rank order difference

method of computation. The following validity coefficient of correlations were obtained for each group.

Thunderbirds	Chiefs	Braves	P.E. Majors	Required Programme
.577	.840	.812	.944	.975

A validity coefficient of correlation for the total group tested was obtained by using the quintiserial technique (Jaspen) (1) of obtaining a coefficient of correlation. The quintiserial method of computing a correlation coefficient was obtained by comparing the five group categories with the five group test scores. Using this technique, the total group test validity was found to be .856.

The reliability coefficients of correlation were found by comparing the scores obtained by the subjects in the three trials. Thus, three reliability coefficients of correlation were found: one between the first and the third trial; one between the second and third trials and one between the first and second trials. These reliability correlation coefficients were obtained by means of the Pearson Product Moment Formula and the Spearman Brown Prophecy Formula.

The following reliability correlation coefficients were obtained:

1st & 3rd Trials	2nd & 3rd Trials	1st & 2nd Trials
.900	.924	.921

The reliability of the test had already been established in previous studies of wall-volley tests. Garrett (2) states " a highly valid test cannot be unreliable since its correlation with a criterion is limited by its own index of reliability."

Standard deviations for the five test groups are included in Table 1. The standard deviation for the whole group was found to be 10.33.

A sixth group of subjects composed of 16 members of another Physical Education major soccer class were given the test on successive weeks to determine:

(a) the effects of practice

(b) when peak performance was reached.

This sixth group was not rank-ordered, but merely given the test with the same test instructions, and under the same conditions. Over a five-week period their aggregate scores were:

Subject	1st Trial	2nd Trial	3rd Trial	4th Trial	5th Trial
A	26	28	37	43	30
B	24	18	27	40	30
C	22	33	32	26	39
D	34	36	36	45	32
E	28	29	24	34	38
F	31	41	49	36	—
G	22	25	38	34	51
H	36	40	48	44	36
I	36	37	42	48	34
J	25	29	26	28	27
K	26	35	28	32	40
L	34	39	32	42	32
M	17	25	28	34	—
N	10	8	14	21	25
O	28	28	38	33	33
P	24	25	—	—	—

Table 2 shows the mean, the range, and the standard deviation scores for the sixth group over five trials.

TABLE 2

Trial Number	Mean Scores	Range of Scores	Standard Deviation
1	27.06	10 - 36	6.85
2	29.75	8 - 41	7.64
3	33.27	14 - 49	8.99
4	36.00	21 - 48	7.39
5	31.93	25 - 41	6.46

TABLE 3

Smoothing Frequency Distribution Curve

Mid-Point of Interval	Frequency	Correction
13.5	0	.66
16.5	2	1.33
19.5	2	4.66
22.5	10	6.66
25.5	8	9.33
28.5	10	8.66
31.5	8	8.00
34.5	6	6.00
37.5	4	5.66
40.5	7	5.00
43.5	4	5.66
46.5	6	4.66
49.5	4	3.66
52.5	1	2.33
55.5	2	1.33
50.5	1	1.00
61.5	0	.33

"To find an adjusted or "smoothed" f, we add the f on the given interval and the f's on the two adjacent intervals and divide the sum by 3." (3) For example, the smoothed f for mid-point 22.5 is

$$\frac{10 + 2 + 8}{3} = \frac{20}{3} = 6.66$$

TABLE 3

Step Interval	Frequency	Cumulative Frequency	Percentile Score
57 - 59.9	1	75	P_{100} = 60
54 - 56.9	2	74	P_{95} = 51.75
51 - 53.9	1	72	P_{90} = 48.17
48 - 50.9	4	71	P_{85} = 46.37
45 - 47.9	6	67	P_{80} = 43.8
42 - 44.9	4	61	P_{75} = 41.68
39 - 41.9	7	57	P_{70} = 40.25
36 - 38.9	4	50	P_{65} = 38.06
33 - 35.9	6	46	P_{60} = 35.14
30 - 32.9	8	40	P_{55} = 33.62
27 - 29.9	10	32	P_{50} = 32.75
24 - 26.9	8	22	P_{45} = 30.66
21 - 23.9	10	14	P_{40} = 29.4
18 - 20.9	2	4	P_{35} = 28.27
15 - 17.9	2	2	P_{30} = 27.15
			P_{25} = 25.78
	75		P_{20} = 24.35
			P_{15} = 23.90
			P_{10} = 22.05
			P_{5} = 20.62
			P_{0} = 14.90

Presentation of Data

The curve of the frequency distribution of the total group was smoothed by using the technique as explained by Guilford (4). As the distribution did not reproduce a normal curve, it was decided to show percentile scores rather than transform scores into standard scores.

REFERENCES

1. Jaspen, Nathan, "Serial Correlation", Psychometrika, March, 1946, pp. 23-30.

2. Garrett, Henry, E., Statistics in Psychology and Education, 5th ed., Longmans, Green and Co., New York, 1958, pp. 471.

3. Ibid., P. 13.

4. Guilford, J. P., Fundamental Statistics in Psychology and Education, 2nd ed., McGraw Hill Book Co., New York, 1950, pp. 51-54.

CHAPTER VI

ANALYSIS OF DATA

The test results show that the experimenter correctly categorized the five groups as being representative of five different levels of soccer skill displayed by students at the University of British Columbia. Substitution of the five group names by the five classes of superior, good, average, below average, and poor would seem to be quite justified. On the basis of the numbers tested, and of the results obtained in this study, a table of norms is suggested for use in classification.

Superior	42 and over
Good	37 - 41
Average	31 - 35
Below Average	25 - 30
Poor	24 and below

The five group validity coefficients of correlation show that greatest difficulty in rank-ordering was experienced with subjects who participated in the higher classes of soccer (superior and good).

Mathews (1) states that validity coefficients from .80 - .85 may be interpreted as very good, and above .85 as excellent; those falling within the range .70 - .79 may be considered acceptable, especially where a subjective judgement is involved.

Garrett (2) states that "the validity of a test depends upon the fidelity with which it measures what it purports to measure."

The validity of this test has been determined experimentally by finding the correlation between the test and the external criterion (the experimenter's rating).

Group validity coefficients of correlation which fall in the range between .737 and .975 are quite acceptable. The only validity correlation coefficient which falls below acceptable standards is that obtained for the Superior or Thunderbird Groups, .577. This is understandable considering the difficulties that present themselves in ranking subjects who play in the superior category, and whose test scores revealed that two-thirds of the group fell between 5 scores above and 5 scores below the mean of 48.

It is obvious that the Superior or Thunderbird Group is quite distinctly differentiated in performance from all other groups. The groups representing the "below average" and "poor" categories are much less different from each other than they are different from the average, good and superior groups. It was assumed that the Physical Education majors group would be superior to the Required Programme group because they were older, and as Physical Education students might be expected to reveal greater interest, coordination and sports skill ability. This difference, it was felt, might be more pronounced if the test was administered to greater numbers of students in those respective groups.

The validity coefficient of .856, obtained by using the quintiserial correlation method, is considered excellent according to the criteria described by Mathews (3). This validity coefficient

is higher than that of .57, obtained by Schaufele (4), and the .76 obtained by Mitchell (5) and very similar to MacDonald's (6) .85. It should also be noted that each of these researchers obtained their test validity coefficient figures by taking a mean of the group validity coefficients.

The difference of 4.33 between the mean scores of the Chiefs (good) and the Braves (average) groups, while significant, might reasonably have been expected to be greater. The season (1962-63) of Varsity soccer has produced an unbeaten Thunderbird team, an extremely weak Chiefs team, and a rather strong Freshmen, or Braves representative team. One would normally have anticipated more equality in the intervals between the means of the first three groups. These peculiarities of the 1962-63 season saw the Thunderbird (superior) group win league and cup tournaments, while the Chiefs (good) group finished at the bottom of their division. The Braves (average) group won promotion to a higher division which probably contributed quite considerably to the inequality of the intervals between the mean scores of the groups. However, this unevenness in playing skill might occur by chance in any given year in Varsity representative soccer teams. In view of the difficulties described, it is fortunate that the groups did reveal distinct differences in mean performance score, in aggregate performance score, and in group validity correlations.

The validity correlations of the five groups were produced by means of the rank difference correlation method and were all acceptable except that of .577 found for the Thunderbird group.

These correlations express the degree of relationship between the experimenter's rank ordering of the subjects in terms of soccer ability, and the rank order of the subjects according to their total performance score on the test.

The experimenter was a former professional soccer player with the Glasgow Rangers for seven years; a Canadian All-Star and current player coach of the University of British Columbia Thunderbird team. This suggests that the background of experience and participation qualifies him as an expert capable of assessing the ability of players after having had the opportunity to observe them at play on no less than 12 occasions. The Required Programme group rank order correlation of .975 was almost a perfect correlation because the experimenter was able to detect quite distinct differences in the individual abilities of these subjects. This was possible because for many of these subjects this was their first experience in soccer, and the various differences in skill and ability were more clearly defined in this group than in any other. These variations in skill and ability, while quite clearly evident to the trained observer in the Physical Education Major and Required Programme groups, became progressively less differentiated in the representative teams. Perhaps a possible source of error in rank-ordering the Thunderbird group was that the experimenter was thoroughly familiar with the player's previous soccer experience, and was also a fellow player with them. Thus, most of the observations of this group were done by the experimenter as a participant, while those of the other groups were done while the experimenter was in the role of a spectator.

The fact that the Thunderbirds (superior) group was a homogeneous group of high ability, made rank-ordering difficult, while the Physical Education Majors (below average) group was a homogeneous group of low ability, yet easily rank-ordered. This would suggest that group homogeneity alone does not preclude rank-ordering, but rather homogeneity plus high skill performance makes rank-ordering an unrealistic technique. The Chiefs (good), the Braves (average) and the Required Programme (poor) groups were heterogeneous groups and readily lent themselves to a rank-ordering technique. The superior group reveals a small range (18); the score distributions form a normal curve about the mean and the mean ± 1 standard deviation encompasses two-thirds of the scores.

Another significant feature of the rank ordering by the experimenter is that in all cases the first ranked occupied the same position in the performance rankings. Thus, choosing the best player did not pose difficulties. The greatest area of difficulty in ranking, in groups other than the Thunderbirds group, presented itself in the eleventh to fourteenth positions. While it was possible to predict the last ranked, or lowest performer, almost as accurately as the first ranked, it was rather difficult to distinguish between the subjects beyond the tenth rank order. It was, therefore, not surprising to find many tie-rankings in the performance scores between the eleventh and fifteenth rank-ordered positions.

If the mean of 34 for the whole group is accepted as the "pass" performance score, then only 33 subjects obtained a "pass" while 42 subjects failed. However, this average has been boosted

by the high performance scores of the Thunderbird group. The medium score for the total group lies between 30 and 29, and this would appear to be a more suitable "pass" mark in this instance.

The test reliability was determined by correlating the scores obtained on the three trials. The correlations of .742 between the first and second trials, .801 between the second and third trials, and .796 between the first and third trials, are quite acceptable in this form of test, especially when one considers that only 75 subjects were used in the test and that the trials were performed one after the other without any noticeable rest period in between.

If the test were administered on three separate occasions, if three parallel forms of the test were administered, then "the reliability of the averaged scores will be the same as the reliability obtained by tripling the length of the test." (7) Thus, these reliabilities could be increased by tripling the length of the test to .900, .924, and .921 by using the formula,

$$r_{nn} = \frac{nr11}{1 + (n-1)r_{11}}$$

where r_{nn} = the correlation between n forms of the test and n alternate forms, r_{11} = the reliability of Trial 1. Mathews (8) states "most tests in physical education should show reliability within the range .90 - .99." He further adds, "tests objective in nature should give highly consistent results when being measured. Therefore, when evaluating tests in terms of reliability containing

such objective measurements, one should expect the coefficients of correlation to fall within this range in order to be acceptable."

The analysis of the test results would appear to support a conclusion that the test is valid and reliable, and therefore, is a useful means of obtaining an effective measure of the soccer ability of university students.

The test validity and reliability correlations are superior to any produced, by acceptable statistical methods, in other wall-volley-type tests, and are also higher than any obtained in battery-type tests. Thus, the test appears to be a useful, economical and a quick means of evaluating soccer skill. The test lends itself to use by the coach or physical educator who wishes a means of grading and classifying large groups of subjects.

The results of the sixth group tested reveal that players do improve their score in repeated trials and the amount of improvement decreases after several tests. There appears to be a tendency for subjects to reach a peak performance score.

The sixth group scores reveal a reduced range after five trials, but there is also a reduced mean score. Thus, repeated trials may produce a general increase in performance score, but after a top performance is achieved there appears to be a tendency for the subject to regress in performance. This may be due to reduced interest in the test after top performance.

In the first trial eleven subjects score in the "below average" category, while after four trials only three subjects are

scoring "below average." This reflects considerable general improvement. It would therefore be more appropriate to allow the subjects practice equivalent to two or three performances of the test to ensure that each subject would do his best.

REFERENCES

1. Mathews, Donald, K., "Measurement in Physical Education", W.B. Saunders Co., London, 1961, p.24.

2. Garrett, Henry, E., Statistics in Psychology and Education, 5th ed., Longmans, Green and Co., New York, 1959, p. 354.

3. Mathews, op.cit., p.

4. Schaufele, Evelyn, F., "The Establishment of Objective Tests for Girls of the Ninth and Tenth Grades to Determine Soccer Ability". (Unpublished Master's Thesis Abstract, Springfield College, 1951).

5. Mitchell, Reid, (Unpublished Master's Thesis, University of Oregon, 1963).

6. MacDonald, Lloyd, G., "The Construction of a Kicking Skill Test as an Index of General Soccer Ability". (Unpublished Master's Thesis Abstract, Springfield College, 1951).

7. Garrett, op.cit., p. 343.

8. Mathews, op.cit.,

CHAPTER VII

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This study was an attempt to develop and establish the usefulness of a wall-volley-type test as a means of measuring the soccer skill of players participating in soccer at The University of British Columbia.

The total group tested was 75 subjects. The total number was composed of 5 teams or groups of 15 subjects. Each group was chosen as being representative of the various levels of soccer played at the University.

The performance scores were correlated against a rank order scale. The rank order scale was the external criterion used in the test to determine test validity.

The final test form was developed as a result of earlier experiments with other subjects to determine the distance of the restraining line; the number of balls to be used, and the method of scoring. The test consisted of three trials of 30-seconds each of volleying a regulation rubber soccer ball from behind a 15 foot restraining line against a target area 8 feet high by 24 feet wide. The ball was put in play by dropping it from the hands at waist height.

The subjects of each group used in the test were rank-ordered by the experimenter from one to fifteen, in order of ability. The subjects were observed playing over a twelve-game

period before the rank orders were finished.

The performance scores of the subjects were correlated with the experimenter's rank-ordering of subjects to determine test validity. The experimenter also used a quintiserial correlation technique to provide a total test validity coefficient.

The total test validity score was .856 using the quintiserial method, while the five-group validity coefficients were:

Thunderbirds	Chiefs	Braves	P.E. Majors	Required Programme
.577	.840	.812	.944	.975

The experimenter felt that the low validity correlation found for the Thunderbird group was due mainly to conditions conforming to the old saying "the spectator sees most of the game." (The experimenter as a playing member of the Thunderbird team made his observations from a participant standpoint). The test results show that it has the highest validity coefficient of any soccer wall-volley-type test currently mentioned in the literature. The results also show that rank-ordering, although a subjective technique, is quite satisfactory when the observer is qualified.

The reliability of the test merely confirms Garrett's (1) statement that "a highly valid test cannot be unreliable."

The groups tested appeared homogeneous as regards high or low ability within groups and heterogeneous between groups. This conclusion was based on the range of scores and the size of group standard deviations. Thus, for example, the Thunderbirds were a

homogeneous group of high ability. The Physical Education Majors were a homogeneous group of low ability.

The test would appear to serve the purpose of testing soccer skill of the University students, and provides reliable, economic and time saving means of grading and classifying large numbers of university students. The test lends itself to the development of norms as a basis for classifying students in five distinct categories of soccer ability: superior, good, average, below average, and poor.

The results of the tests of the 16 members of the 6th group would indicate that the players do improve their scores with regular practice, and that such improvement decreases as the player appears to approach or reach a peak performance score.

CONCLUSIONS

1. On the basis of present evidence, this wall-volley test appears to be a valid and reliable instrument for measuring soccer skill.
2. Repetition of the test would permit the development of more refined norms.
3. The test reliability and validity indicate that it is a more appropriate measure of soccer ability than any existing battery or single item type tests illustrated in the literature.
4. The test is most useful for teachers and coaches who require an economic and time saving means of grading and classifying large groups of players.
5. The test may be used to categorize students as superior, good, average, below average and poor.

6. Repeated practice of the test does cause improvement in performance.
7. Continued practice of the test is a good means of improving soccer ability.
8. Use of the wall-volley technique is a useful device for introducing and developing soccer interest and ability in students.

RECOMMENDATIONS

1. It is suggested that test efficiency be further tested by repeating it with further groups and greater numbers of subjects.
2. That the test scores be used as a means of grading students in one of five categories: superior, good, average, below average and poor, rather than the teacher or coach subjectively assess ability.
3. That the number of trials remain at three and the aggregate score be taken.
4. That the test be conducted indoors to determine usefulness.
5. That trials be conducted with appropriate groups to determine suitable dimensions and test conditions.
6. That tables of norms be developed for varsity, high school and elementary school students.

REFERENCES

1. Garrett, Henry, E., Statistics in Psychology and Education, 5th Ed., Longmans, Green, and Co., Inc. New York, 18, 1960, p. 361.

APPENDIX

APPENDIX

TABLE 1

Thunderbird Scores

Rank Order	Trials 1	2	3	Aggregate	Performance Order	Group Data
1	18	19	21	58	1	
2	16	17	15	48	7	
3	17	18	21	56	2	
4	17	19	18	54	3	
5	16	18	13	47	8.5	Mean 48.00
6	16	15	14	45	11	Standard
7	14	16	14	44	12.5	Deviation
8	13	16	12	41	14	4.99
9	15	14	17	46	10	Rank. Diff.
10	18	16	17	51	4	Correlation
11	16	13	21	50	5	.576
12	16	14	17	47	8.5	
13	16	16	17	49	6	
14	15	13	16	44	12.5	
15	13	14	13	40	15	
Totals	236	238	246	720		

TABLE 2

Chief Scores

Rank Order	Trials 1	2	3	Aggregate	Performance Order	Group Data
1	16	15	19	50	1	
2	14	12	17	43	3.5	
3	15	12	12	39	6.5	
4	13	14	16	43	3.5	
5	13	11	11	35	9	Mean 36.93
6	14	13	12	39	6.5	
7	15	8	15	38	8	S.D. = 6.72
8	14	17	15	46	2	
9	15	11	15	41	5	
10	10	14	10	34	10	
11	11	9	10	30	12	Rank Diff.
12	13	9	9	31	11	Correlation
13	10	8	11	29	13.5	= .840
14	9	8	10	27	15	
15	6	12	11	29	13.5	
Totals	188	173	193	554		

TABLE 3

Braves

Rank Order	Trials 1	2	3	Aggregate	Performance Order	Group Data
1	16	17	14	47	1	
2	14	9	15	38	4	
3	16	8	15	39	3	Mean 32.60
4	10	12	13	35	6	
5	13	13	11	37	5	S.D.= 6.26
6	11	9	12	32	7	
7	5	9	16	30	8.5	
8	11	5	13	29	10.5	
9	9	11	9	29	10.5	
10	9	9	12	30	8.5	
11	15	12	13	40	2	Rank Diff.
12	12	8	6	26	13	Correlation
13	9	6	12	27	12	.812
14	10	7	8	25	14.5	
15	9	8	8	25	14.5	
Totals	169	143	177	489		

TABLE 4

P.E. Majors

Rank Order	Trials 1	2	3	Aggregate	Performance Order	Group Data
1	13	9	12	34	1	
2	9	11	12	32	3	
3	10	9	11	30	4	
4	13	11	9	33	2	
5	12	8	8	28	5	Mean 25.66
6	10	8	9	27	6	
7	8	8	8	24	8.5	S.D.= 4.59
8	7	8	10	25	7	
9	9	9	7	24	8.5	
10	8	7	6	21	13	
11	7	8	6	21	13	Rank Diff.
12	8	6	9	23	10	Correlation
13	8	6	7	21	13	.944
14	7	8	6	21	13	
15	9	5	7	21	13	
Totals	138	120	127	385		

TABLE 5

Required Programme

Rank Order	Trials 1	2	3	Aggregate	Performance Order	Group Data
1	12	12	12	36	1	
2	13	11	9	33	2	Mean 24.20
3	10	9	11	30	3	
4	10	9	8	27	5	S.D.= 5.76
5	7	11	10	28	4	
6	8	8	8	24	7	
7	7	8	10	25	6	
8	9	8	6	23	9	
9	8	6	7	21	11	
10	9	7	7	23	9	
11	8	6	9	23	9	Rank Diff.
12	8	5	7	20	12	Correlation
13	5	9	4	18	13	.975
14	5	7	4	15	14	
15	5	5	5	15	15	
Totals	124	121	117	362		

TABLE 6

Total Group Aggregate Scores

Rank Order	Thunderbirds	Chiefs	Braves	P.E. Majors	Reqd. Prog.	Group Data
1	58	50	47	34	36	
2	48	43	38	32	33	N = 75
3	56	39	39	30	30	
4	54	43	35	33	27	Aggregate
5	47	35	37	28	28	score
6	45	39	32	27	24	= 2510
7	44	38	30	24	25	
8	41	46	29	25	23	Mean
9	46	41	29	24	21	= 33.47
10	51	34	30	21	23	
11	50	30	40	21	23	Median
12	47	31	26	23	20	= 29
13	49	29	27	21	18	S.D.
14	44	27	25	21	16	= 10.33
15	40	29	25	21	15	
Totals	720	554	489	385	362	
Mean	48.00	36.93	32.60	25.66	24.20	
Rank Difference Correlation	.576	.840	.812	.944	.975	

The Rank Difference Correlations for each group were found by using the formula:

$$r = 1 - \frac{6 \times D^2}{N(N-1)}$$

where r = coefficient of correlation form rank differences.

ΣD^2 = the sum of the squares of the differences in rank.

N = the number of subjects or paired rankings.

Thus, for the Thunderbirds the formula is replaced by:

$$r = 1 - \frac{6 \times 237}{15 \times 224}$$

$$= .576$$

Substituting in the formula for the Chiefs we have:

$$r = 1 - \frac{6 \times 89.50}{15 \times 224}$$

$$= .840$$

Substituting in the formula for the Braves we have:

$$r = 1 - \frac{6 \times 105.50}{15 \times 224}$$

$$= .812$$

Substituting in the formula for the P.E. Majors we have:

$$r = 1 - \frac{6 \times 31.50}{15 \times 224}$$

$$= .944$$

And substituting in the formula for the Required Programme Group we have:

$$r = 1 - \frac{6 \times 14}{15 \times 224}$$

$$= .975$$

The formula for quintiserial correlation (a correlation coefficient for the total group) is expressed as:

$$r_{quint} = \frac{ZaYa(Zb-Za)\ Yb\ (Zc-Zb)\ Yc\ (Zd-Zc)\ Yd-Zd\ Ye}{Oy \sqrt{\frac{Za^2}{a} \quad \frac{(Zb-Za)^2}{b} \quad \frac{(Zc-Zb)^2}{c} \quad \frac{(Zd-Zc)^2}{d} \quad \frac{Zd^2}{e}}}$$

$$= \frac{.2800x48(.3863-.2800)36.93\ (.3863)32.60\ (.2800-3863)25.66-2800x24}{10.34 \sqrt{\frac{.2800^2}{.20} \quad \frac{(.3863-2800)^2}{.20} \quad \frac{(.3863-.3863)^2}{.20} \quad \frac{(.2800-.3863)^2}{.20} \quad \frac{.2800^2}{.20}}}$$

$$= \frac{7.96}{9.26}$$

$$= .856$$

"The effect of serial correlation is to normalize the segmented distribution at the time that the correlation coefficient is obtained. If the number of segments is large, and if the segmented variable is already normally distributed, the resulting correlation will be the same as a Pearson Product Moment correlation.

Symbolism

The following symbolism will be adopted:

Let y be a continuous variable, x be a continuous segmented variable, normally distributed, and r be the coefficient of correlation (linear) between x and y.

Let a = the proportion of cases in the top right-most segment of the x distribution

b = the proportion of cases in the second highest segment,

c = the proportion of cases in the third highest segment, etc., and

f = the proportion of cases in the f-th segment of the distribution.

Then a b c ... = 1.

Let qa = a,

qb = a b,

qc = a b c, etc.,

qf = a b ... f = the area above the left boundary of f-th segment, and

qf-1 = a b ... up to but not including f = the area above the right boundary of the f-th segment.

Let Za = the ordinate of the normal curve, assuming a unit normal distribution at qa,

Zb = the ordinate of the unit normal curve at qb, etc.,

Zf = the ordinate at qf, and

Zf-1 = the ordinate at qf-1.

Let Ya = the mean of the y's in the top (right most) segment of the x distribution.

Yb = the mean of the y's in the second highest segment, etc., and

Yf = the mean of the y's in the f-th segment.

Let Xa = the mean of the x's in the top segment of the x distribution, etc., and

Xf = the mean of the x's in the f-th segment.

Except for a few modifications, conventional symbolism has been adhered to. The symbol q is sometimes taken in the literature to represent the proportion of cases in one of the segments (i.e. $q = q^1 - q^2$, a decumulated frequency), as well as the cumulation of proportions or frequencies from a given line of truncation to the end of the curve. In tables of the normal probability integral oriented in terms of q, and therefore, in this paper q always represents the area from the given line of truncation to the end of the curve.

Since the normal curve is symmetrical, the ordinates z are equivalent for complementary q's. Consequently most normal tables oriented in terms of q carry the argument only from zero to .500, and it is there necessary to consult the complement of q for values of q higher than .500. This, of course, does not disturb the meaning of q.

In this paper the upper segments of the x distribution (the desirable pole of the trait or measure in question) are placed to the right of the lower (or less desirable) segments of the normal curve." (1)

The reliability coefficients of correlation were obtained using the product-moment formula,

$$r = \frac{N\ xy - (x)(y)}{\sqrt{\left[N\ x^2 - (x)^2\right]\left[N\ y^2 - (y)^2\right]}}$$

where r = the coefficient of correlation of reliability between two trials

x = the sum of the first trial scores.

y = the sum of the second trial scores.

x^2 = the sum of the squares of the first scores.

y^2 = the sum of the squares of the second scores.

xy = the sum of the first score times the second scores.

N = the number of subjects

Thus, substituting in trials 1 and 3 we found,

$$r = \frac{75 \times 822 - (45 \times 12)}{\sqrt{\left[75 \times 959 - (45)^2\right]\left[75 \times 1126 - (12)^2\right]}}$$

$$= .796$$

Substituting for trials 2 and 3 we find,

$$r = \frac{75 \times 831 - (31 \times 12)}{\sqrt{\left[75 \times 959 - (31)^2\right]\left[75 \times 1726 - (12)^2\right]}}$$

$$= .801$$

Substituting in the formula for trials 1 and 2 we find,

$$r = \frac{75 \times 751 - (45 \times 31)}{\sqrt{\left[75 \times 959 - (-45)^2\right]\left[75 \times 934 - (31)^2\right]}}$$

$$= .742$$

Applying the Spearman Brown Prophecy formula for extending the correlation between n forms of a test and n comparable forms, we have,

$$r_{nn} = \frac{N r_{11}}{1+(n-1) r_{11}}$$

where r_{nn} = the correlation between n forms of the test and n alternate forms

r_{11} = the reliability coefficient of Trial 1.

Thus, by tripling the test, the Spearman Brown Prophecy formula would alter the correlations of reliability as follows,

Trials 1 and 3

$$r_{3111} = \frac{3 \times .796}{1 + (2 \times .796)}$$

$$= .900$$

for trials 2 and 3

$$r_{3111} = \frac{3 \times .801}{1 + (2 \times .801)}$$

$$= .924$$

for trials 1 and 2

$$r_{3111} = \frac{3 \times .796}{1 + (2 \times .796)}$$

$$= .921$$

The standard deviation for the five groups and for the total group was found by using the following formula

$$\sqrt{\frac{x^2}{N} - M^2}$$

Substituting in Thunderbird Group:

$$S.D. = \sqrt{\frac{34934}{15} - (48)^2} = 4.99$$

Substituting in Chiefs Group:

$$S.D. = \sqrt{\frac{21134}{15} - (35.6)^2} = 6.26$$

Substituting in Braves Group:

$$S.D. = \sqrt{\frac{16529}{15} - (32.60)^2} = 6.34$$

Substituting in P.E. Major Group:

$$S.D. = \sqrt{\frac{10193}{15} - (25.66)^2} = 4.59$$

Substituting in Required Programme Group:

$$S.D. = \sqrt{\frac{9232}{15} - (24.13)^2} = 5.76$$

Substituting in total group:

$$S.D. = \sqrt{\frac{92022}{75} - (33.47)^2} = 10.33$$

Standard Deviation Scores for Sixth Group Trials

1st Trial $S.D. = \sqrt{\frac{Ex^2}{N} - (M)^2} = \sqrt{\frac{11919}{16} - (26.44)^2} = 6.85$

2nd Trial $S.D. = \sqrt{\frac{Ex^2}{N} - (M)^2} = \sqrt{\frac{15453}{16} - (29.94)^2} = 7.64$

3rd Trial $S.D. = \sqrt{\frac{Ex^2}{N} - (M)^2} = \sqrt{\frac{17815}{15} - (33.27)^2} = 8.99$

4th Trial $S.D. = \sqrt{\frac{Ex^2}{N} - (M)^2} = \sqrt{\frac{20256}{15} - (36)^2} = 7.39$

5th Trial $S.D. = \sqrt{\frac{Ex^2}{N} - (M)^2} = \sqrt{\frac{15909}{13} - (34.38)^2} = 6.46$

Smoothing Frequency Distribution Curve

Mid-Point of Interval	Frequency	Correction Calculation	Correction
13.5	0	$\frac{0 + 0 + 2}{3}$	.66
16.5	2	$\frac{0 + 2 + 2}{3}$	1.33
19.5	2	$\frac{2 + 2 + 10}{3}$	4.66
22.5	10	$\frac{2 + 10 + 8}{3}$	6.66
25.5	8	$\frac{10 + 8 + 10}{3}$	9.33
28.5	10	$\frac{8 + 10 + 8}{3}$	8.66
31.5	8	$\frac{10 + 8 + 6}{3}$	8.00
34.5	6	$\frac{8 + 6 + 4}{3}$	6.00
37.5	4	$\frac{6 + 4 + 7}{3}$	5.66
40.5	7	$\frac{4 + 7 + 4}{3}$	5.00
43.5	4	$\frac{7 + 4 + 6}{3}$	5.66
46.5	6	$\frac{4 + 6 + 4}{3}$	4.66
49.5	4	$\frac{6 + 4 + 1}{3}$	3.66
52.5	1	$\frac{4 + 1 + 2}{3}$	2.33
55.5	2	$\frac{1 + 2 + 1}{3}$	1.33
58.5	1	$\frac{2 + 1 + 0}{3}$	1.00
61.5	0	$\frac{1 + 0 + 0}{3}$	.33

Calculation of Percentiles

$$P_{100} = 57 + \frac{75-74}{1} \times 3 = 60$$

$$P_{95} = 51 + \frac{71.25-71}{1} \times 3 = 51.75$$

$$P_{90} = 48 + \frac{67.5-67}{4} \times 3 = 48.17$$

$$P_{85} = 45 + \frac{63.75-61}{6} \times 3 = 46.37$$

$$P_{80} = 42 + \frac{60-57}{5} \times 3 = 43.8$$

$$P_{75} = 39 + \frac{56.25-50}{7} \times 3 = 41.68$$

$$P_{70} = 39 + \frac{52.5-50}{6} \times 3 = 40.25$$

$$P_{65} = 36 + \frac{48.75-46}{4} \times 3 = 38.06$$

$$P_{60} = 33 + \frac{45-40}{7} \times 3 = 35.14$$

$$P_{55} = 33 + \frac{41.25-40}{6} \times 3 = 33.62$$

$$P_{50} = 30 + \frac{37.3-32}{8} \times 3 = 32.75$$

$$P_{45} = 30 + \frac{33.75-32}{8} \times 3 = 30.66$$

$$P_{40} = 27 + \frac{30.22}{10} \times 3 = 29.4$$

$$P_{35} = 27 + \frac{26.25-22}{} \times 3 = 28.27$$

$$P_{30} = 27 + \frac{22.5-22}{10} \times 3 = 27.13$$

$$P_{25} = 24 + \frac{18.75-14}{8} \times 3 = 25.78$$

$$P_{20} = 24 + \frac{15-14}{8} \times 3 = 24.35$$

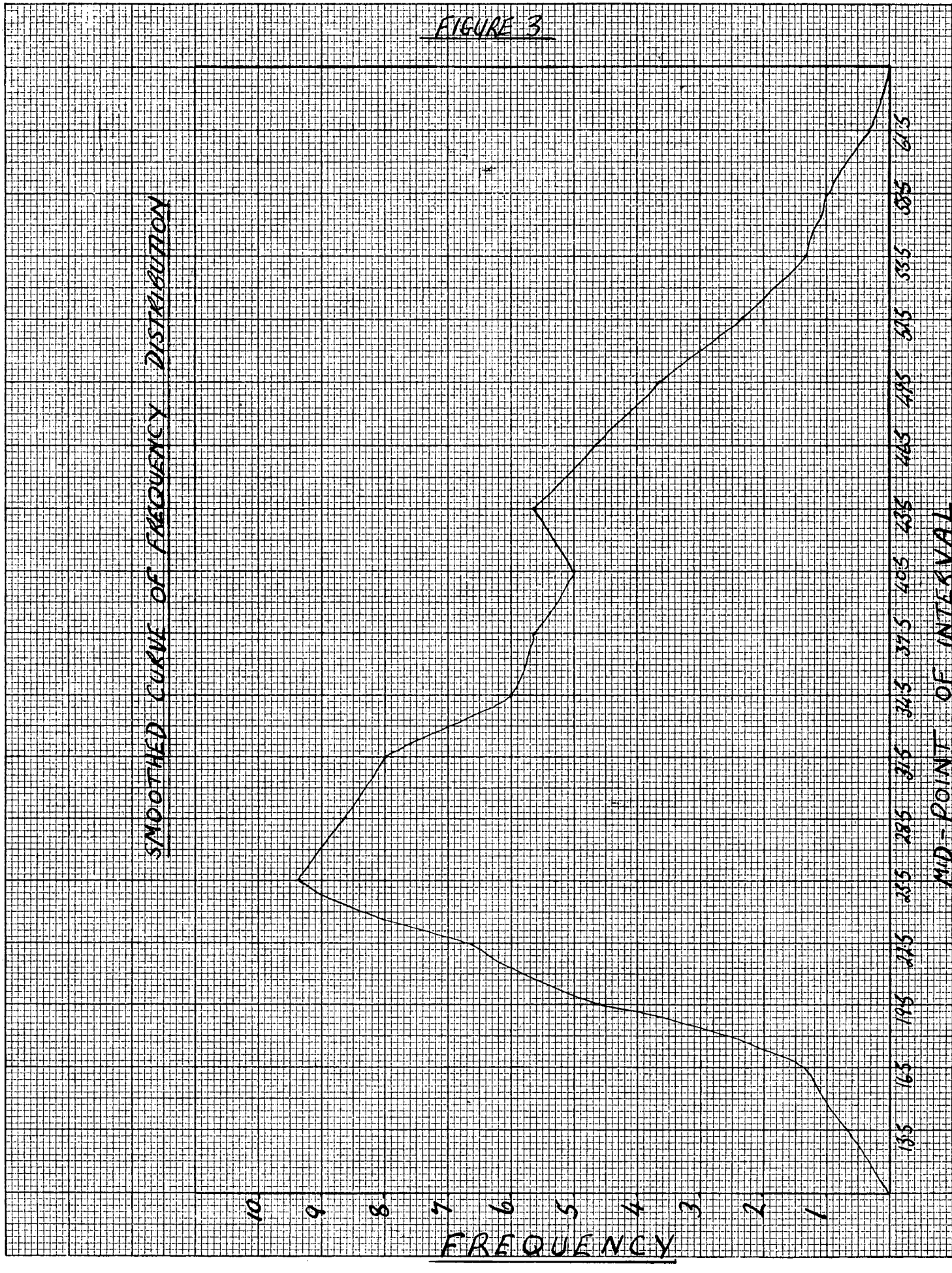
$$P_{15} = 21 + \frac{11.25-4}{10} \times 3 = 23.90$$

$$P_{10} = 21 + \frac{7.5-4}{10} \times 3 = 22.05$$

$$P_{5} = 18 + \frac{3.75-2}{2} \times 3 = 20.62$$

$$P_{0} = 14.9$$

FIGURE 3


SMOOTHED CURVE OF FREQUENCY DISTRIBUTION
FREQUENCY
10
9
8
7
6
5
4
3
2
1
135
165
195
225
255
285
315
345
375
405
435
465
495
525
555
585
615
MID-POINT OF INTERVAL

REFERENCES

1. Jaspen, Nathan, "Serial Correlation", Psychometrika, March, 1946, pp. 23-30.

BIBLIOGRAPHY

BOOKS

Caswell, John, E., Soccer for High Schools, A.S. Barnes and Co., New York, 1933.

Csaknady, Jeno, Learn to Play the Hungarian Way, Hungarian Sport Publishing House, Budapest, 1954, pp. 17-71.

Finney, Tom, Instructions to Young Footballers, Museum Press Limited, London, 1955, pp. 16-34.

Garrett, Henry, E., Statistics in Psychology and Education, 5th ed., Longmans, Green and Co., New York, 1958, pp. 471.

Guilford, J. P., Fundamental Statistics in Psychology and Education, 2nd ed., McGraw Hill Book Co., New York, 1950, pp. 51-54.

Mathews, Donald, K., Measurement in Physical Education, W. B. Saunders Co., Philadelphia, 1958, 349 pp.

Meisl, Willy, Soccer Revolution, Phoenix Sports Club, London, 1955, pp. 20-34.

Voltmer, Edward, F., and Esslinger, Arthur, A., The Organization and Administration of Physical Education, 3rd ed., Appleton-Century-Crofts, Inc., New York, 1958, 547 pp.

Winterbottom, Walter, Soccer Coaching, Naldrett Press Ltd., London, 1955.

PERIODICALS

Bill C-131, An Act to encourage Fitness and Amateur Sport. As passed by the House of Commons, 25th September, 1961. Queen's Printers, Ottawa, 1961.

Brady, George, F., "Preliminary Investigations of Volleyball Playing Ability". Research Quarterly, vol. 18, no. 1, March, 1945.

Cornish, Clifford, "A Study of Measurement of Ability in Handball". Research Quarterly, vol. 20, no. 2, May, 1949.

Dyer, Joanna, T., "Revision of Backboard Test of Tennis Ability". Research Quarterly, vol. 9, no. 1, March, 1938, p. 27.

French, Esther and Cooper, Bernice, "Achievement Tests in Volleyball for High School Girls". Research Quarterly, vol. 8, no. 2, May, 1937.

Heath, Marjorie, L. and Rogers, Elizabeth, G., "A Study in the Use of Knowledge and Skill Tests in Soccer", Research Quarterly, vol. 33, no. 3, December, 1932.

Jaspen, Nathan, "Serial Correlation", Psychometrika, March, 1946, pp. 23-30.

Miller, Frances, A., "A Badminton Wall Volley Test". Research Quarterly, vol. 22, no. 2, May, 1951.

Russell, Naomi, and Lange, Elizabeth, "Achievement Tests in Volleyball for Junior High School Girls", Research Quarterly, vol. 2, no. 4, December, 1940.

Vanderhoof, Mildred, "Soccer Skill Tests", Journal of Health and Physical Education, vol. 42, no. 3, October, 1932.

Weiss, Raymond, A., and Scott, Gladys, M., "Construction of Tests", Research Methods, 2nd ed., American Association for Health, Physical Education and Recreation, Washington, D.C., 1959, pp. 239-240.

THESES

Bontz, Jean, "An Experiment in the Construction of a Test for Measuring Ability in Some of the Fundamental Skills Used by Fifth and Sixth Grade Children in Soccer". (Unpublished Master's Thesis, State University of Iowa, 1942).

Crawford, Elinor, A., "The Development of Skill Test Batteries for Evaluating the Ability of Women Physical Education Major Students in Soccer and Speedball". (Unpublished Doctoral Thesis, University of Oregon, 1958).

Konstantinov, J., "Construction of a Soccer Kicking Test". (Unpublished Master's Thesis, Springfield College, 1950).

MacDonald, Lloyd, G., "The Construction of a Kicking Skill Test as an Index of General Soccer Ability". (Unpublished Master's Thesis, Springfield College, 1951).

Mitchell, Reid, "A Wall-Volley Test for Measuring Soccer Ability in Fifth and Sixth Grade Boys". (Unpublished Master's Thesis, University of Oregon, 1963).

Schaufele, Evelyn, F., "The Establishment of Objective Tests for Girls of the Ninth and Tenth Grades to Determine Soccer Ability". (Unpublished Master's Thesis, State University of Iowa, 1940).