A CRITERION-REFERENCED TEST BATTERY TO ASSESS OVERALL ABILITY IN TEAM HANDBALL

Shoukeh, Nart, A, PhD

Associate Professor, Deptartment of PE, Y.U. Irbid, Jordan

Mismar, Bassam, A, PhD

Professor, Department Of superv.& Inst. University of Jordan, Amman, Jordan

Hisham Ali Al Dmour, PhD

Dean of AL Karak College University Al Balqa Applied University, Jordan

Abstract

The main purpose of the study was to construct the team handball skill test battery, based on criterion measurement in order to evaluate the overall ability of pertinent skills for rudimentary level in the game. Ss were (N=60) students from the faculty of physical education at Yarmuk University in the northern part of the Hashemite Kingdom of Jordan. Those Ss were chosen purposively while registering in 2 team handball classes, each of which had (n=30) students during the academic year 2011/2012.

The researchers utilized descriptive methodology by designing the proposed test battery, which comprises of (overhead pass, ball catching and shooting with forward jump) and represents the overall ability of individuals in the game. In order to achieve the objective of the study, selected procedures and techniques have been used to check for scientific coefficient of the tool based on criterion referenced tests and measurement, with considering some indicators such as validity, reliability, and objectivity.

Data were collected by examining the Ss of the study throughout using the proposed tests battery at the end of the course. The results revealed that the skill tests battery was valid, reliable, objective, and could be a practical tool to obtain certain and similar objectives. The study recommends adopting this tests battery with the criterion –referenced based measurements to assess overall ability in team handball by assigning students to mastery and none-mastery states.

Keywords: Criterion-Referenced Measurement, Team Handball Skill Tests, Test Battery

Introduction

The idea of Criterion referenced measurement and its relative framework comes from utilizing quality behaviors (criteria) that students can accomplish in order to make the mission of evaluation in psycho- motor domain as easy and objective as possible. This could be true when making interpretations of individuals' skill performance by comparing it with motor skill criteria "relevant and quality standard". This gives meaning to the pertinent whole assessment process by describing one's ability in certain motor skill. Popham (1978) stated that CR tests should be called objective referenced tests, as a result of designing test items according to behavioral objectives of skill being taught. Therefore, there such needs to determine the behavioral objectives and relevant out comes specially when talking about acquisition of skills by allocating students to mastery states (Safrit, 1981), as well as using success rate of performing the motor task by utilizing systematic observation and criteria sheets in order to check for learning status(Mosston and Ashworth, 1986). These conditions make the scope and the sequence of behavioral domain to be observed, measured in given tasks, and capable of writing progress reports of the whole class according to the accomplishment of specific performance objectives(Siedentop, et al, 1986).

Comparing with Norm-referenced tests and measurements, there has been a lack of convincing evidence regarding the utilization of CR framework model in order to assess the overall ability of individuals especially in team sport events. As a result of individual variations in abilities and the state of individual differences, NR does not seem to make good fit with the mentioned circumstances. That is simply because it compares individual's performance with his/her classmates or finding the relative performance of one student comparing with his/her classmates. In order to yield scores pertaining to its normative group who took the same test, ones know who got the best or worst score. Safrit (1980) showed some concern related to this line of research by stating that there is such a popular usage of this academic area. In reality, her work concentrated on: how to use certain criteria to judge the acceptability of tests, especially by addressing imperative terms such as validity and reliability as well as objectivity of the tests. She addressed the future of this academic research in the next coming decades. In CR, it could be possible to refer one performance to domain scores depending upon which the test is used (Safrit, 1981). This process is called criterion-referenced measurement which leads to criterion-referenced evaluation. According to Shiffelett & Schuman (1982) CR test could be defined as a tool constructed to yield measurement that can be directly interpreted in terms of specified and well described performance.

This area of discipline has been covered by some outstanding and well known scientists in measurements and evaluation such as Safrit, Baumgartner, Jackson, Mood, etc. Those figures indicate the significance of such an area in teaching physical education curriculum by making a good contribution to the field by directing and controlling individual's efforts. Regardless of that, an invitation to researchers shall be made in order to consider working with this line of research and directing more attentions to conduct more research.

Significance and Problem of the study

One of the classic psychometric styles in measurement and evaluation in the area of education is what has been called "criterion –referenced tests". This style has been introduced as a result of the criticism that has been addressed about norm-referenced tests. That was one reason which made the researchers think about the following question: whether or not an individual possesses particular skills related to team handball? To answer such a question, the researchers utilized another type of measurement which deals directly with the quality of performance and well-described techniques related to the skills of the game.

The significance of the study exists through presenting the skill tests battery in team handball by emphasizing the quality of skill movements as major variables. As a result of the paucity of research in the area of teaching physical education, related to the fore-mentioned question which seems to be quite hard to answer via the utilization of norm tests as a result of using percentages or rank orders to differentiate between class mates. These norm-reference based tests are available and commonly used in testing sports and psychomotor skill in physical education (Kirkendall, et. al, 1982). As observed by researchers and pertaining to their experiences, this topic may be considered as a well-chosen one because it seems worthwhile and useful in offering scientific based research with an objective framework to assess skill techniques of the students in team handball, by considering skills acquisition as an indicator to check for overall ability in the game by utilizing the most common and frequently used skills both in the game and during the match. Also, this study may provide PE teachers with guides to help them construct such tests for sports events in order to assign grades to students in such a way by being as objective as possible during the evaluation of lessons outcomes by the end of teaching units.

The Objective of the study

The objective of the study was to:

Construct the test battery related to selected skill tests in team handball based on criterionreferenced measurement in order to assess players overall ability in the game.

The questions of the study

- What was the possibility to construct a scientific tool based on CR type measurements to evaluate the overall skills ability of rudimentary level students in team handball?

The following minor questions were raised:

- Are the skill tests battery considered to have scientific indicators such as validity, reliability, and objectivity?
- How could one assess the overall ability in team handball by using the proposed criterion test battery?

Operational definition of terms

A criterion- referenced test: could be defined as a type of the test which concerns about quality of performance and movement, written or stated by pinpointing some standard determined by assessors.

A test Battery: Several tests items that represent the skills to be measured and evaluated by applying it on real situation.

Review of CR research:

Because of the lack of research in this line of endeavors, the researchers exert extensive effort to make a wide review in the field of specialty within the area of tests and measurement and more particular in team handball skills. This effort may encounter some hardship. The following related studies are reviewed in these instances:

Shifflett and Schuman(1985) published their study in which it aimed at constructing and evaluating a CR test for archery for beginners. SS were required to shoot certain arrows towards a target from the prescribed distance. The researcher used a specific technique for estimating validity and reliability of the test, and predetermine cut off score was implied. The results of the study revealed that the process of misclassifications were minimized as a result of the useful method presented by Berks methods in (1976).

Hudson (1985) conducted a study aimed at examining the utilization of selected biomechanical variables in order to predict basketball skills by using a CR measurement. Ss were U.S. college women B.B players who were divided into three groups (elite, good and poor) according to skill levels. Such discriminant analysis (DA) was employed to predict the ability of Ss. These DA came from accuracy, stability as well as the height of release. Results revealed that using discriminant analysis along with biomechanical variables could be a useful technique in the prediction of BB ability.

Mosher and Schutz (1982) conducted their study in order to examine and conclude the findings of children of British Columbia status quo pertaining to their fitness level by using CR evaluation. Ss were (3000) students tested on physical fitness components, by classifying

them to different categories, ranged from very satisfactory to weak .The results revealed that fitness levels of the Ss were rated between weak and marginally satisfactory, in which the criteria set by provincial evaluative panels.

Methodology

The methodology of this study was descriptive in nature due to its relevancy to the objective of the study. The population of the study were all physical education students who enrolled in the faculty of physical education at Yarmuk University during the academic year (2011/2012). The total number was (125) students. Subjects were selected purposively (N=60) students who were registered in the course of team handball in two mutually exclusive groups.

According to the level of expertise of the researchers as well as reviewing the pertinent literature in team hand ball, curriculum guides and instructional manuals, tests, measurement, and evaluation of psychomotor domain, the researchers have adopted and selected this study with relevant variables and skill tests in order to construct the battery of team handball skill test. This battery has been chosen to represent the most important skills, which are commonly used during the game and could reflect the overall ability for beginner players.

The tool of the Study

The researchers relied upon such sources as documents analysis as well as experts opinions in the field of team handball to assist in formulating and constructing the proposed battery with utilizations of scientific coefficient processing in the next stage such as validity, reliability, and objectivity. Actually, three tests have been determined in order to formulate the battery, these skill tests were: overhead pass, ball catching, and shooting with forward jump. Each skill was analyzed by using task analysis technique. Each skill test contains 3 parts (appendix 1), and each part consists of related certain behavioral objectives (criteria) and designated scores. The total score for each skill test is (10) and the total score of the whole tests battery is (30).

Validity of the test battery

The content validity check was verified by (7) experts in the field of game sports as well as in team handball; they have been requested to verify the criteria and check the proposed score for each skill test (technique). They provided some remarks on the tool related to some (kpi) to be considered such as: what it is intended to measure? Are the items of the test battery appropriate and could be observed considering to be free of biasness? Are the directions of the battery cleared and convenient to examinees? All notes and remarks were taken into considerations while preparing the last revision of the tool of the study.

Reliability of the test battery

After determining the content validity of the test battery, the researchers utilized a sample of (10) students from the study population, outside the sample of the study. As far as reliability depends upon the utilization of tests themselves, researchers decide to apply them in two different occasions by using Pearson correlation coefficient in order to check for the application in the two occasions(test-re-test) with a period of one week between them (tables 2,3,and 4) show the values of r. This gave indications relative to the extent of suitability in producing tests with desirable test scores as well as the consistency of tests items.

Objectivity of the test battery

The researchers verified the objectivity coefficient in terms of the clarity of skill test items and how the two examiners got close images in perceiving the items of test battery. A sample of (10) subjects has been chosen to check for objectivity coefficient by using two different examiners who got some training about the test items. This gives indications that the tests scores have an objective base. Tables (4, 5, and 6) show more details.

Testing protocol

The following procedures have been of great concerns in this research:

All testing procedures were taking place in the outdoor play ground area of the faculty of physical education at Yarmuk University. Pilot testing process was taking place before data collection in order for the assessors to be ready in assigning grades. Each Ss engaged in worm-up session in terms of getting prepared physically and mentally and taking some skill trials with the attendance of instructors. The skills were taught during the semester before testing take place at the final exam week. The total score for the test battery is (30), see (appendix 1). Details instructions were given to both examiners and Ss. The location of each Ss was predetermined. Two examiners recorded the skills performance of the skill test battery by utilizing the criterion sheets. Each Ss has performed (3) trials in each test, where the best trial among them was counted. Data were gathered and analyzed by the appropriate statistical treatments by using (spss).

The results and discussions

The results of the study were organized in accordance to research questions. In this instance, the first question to be answered: What was the possibility to construct a scientific tool based on CR type measurements to evaluate the overall skills ability of rudimentary level students in team handball?

To answer this question, the researchers believe that there is such a strong possibility to construct the criterion-referenced tests and measurements in psychomotor domains for all aged groups related to sport events. This assumption may be true if ones follow prescribed procedures accurately and develop the study protocol in order to construct a testing tool. Also, answering the following minor question could be of great value in the process of tests construction.

"Are the skill tests battery considered to have scientific indicators such as validity, reliability, and objectivity?"

To answer this question, the researchers used a statistical technique to double check the validity of this battery by applying the collected data for all Ss. Table (1) exemplifies the loading factors and Eigen values of each test and its parts among the battery as an indicator of validity.

Skill	Skill Part	Loading Factors	Eigen values	% of Variance
	Preparatory Part 1	0.68	3.75	26.80
Overhead Pass	Preparatory Part 2	0.74		
	Preparatory Part 3	0.82		
	Main Part	0.81		
	Closing Part	0.75		
Ball Catching	Preparatory Part 1	0.69	3.09	22.06
	Preparatory Part 2	0.77		
	Preparatory Part 3	0.55		
	Main Part	0.74		
	Closing Part	0.53		
Shooing with	Preparatory Part 1	0.80	2.93	21.30
Forward Jump	Main Part 1	0.84		
	Main Part 2	0.87]	
	Closing Part	0.75		

Table 1. : Factorial Analysis results (using Varimax) of the three skills representing the test battery

Table (1) indicates that the results of the factorial analysis by using principles compared method and factor rotations with varimax. These values were fair enough for the three skill tests representing the test battery. The loading factors values were high enough for the current purpose of the study. The values were ranging between 0.87 for main part 2 of shooting with forward jump skill to 0.53 for closing part of ball catching skill, where all saturated values were above the minimum absolute value (0.30) according to Guilford criterion. Also, table (1) showed the Eigen values of 3.75, 3.09. and 2.93 for the three tests respectively. These values were greater than (1.00) which considered as a cut off score with respect to Kizer criterion. In addition, the explained variance percentage values for each factor was also high as noticed in the table with values of 26.80, 22.06, 21.30 respectively ,and with a total of 73.6 the whole test battery. This shared value of the 3 components is considered good enough to give an interpretation of almost ³/₄ from the total explained variance. When applying such test battery, the results indicate a clear instance the appropriate

statistical validity of the test battery that measures the overall ability of team handball players at rudimentary level.

In term of reliability coefficient, tables (2, 3, and 4) show the values of Pearson Product Moment Correlation between application and reapplication of the test battery:

Skill Parts	Reliability Coefficients	Significance Levels
Preparatory Parts:	0.916	0.000
Preparatory Part 1	0.910	0.000
Preparatory Part 2	0.870	0.001
Preparatory Part 3	0.942	0.000
Main Part	0.783	0.007
Closing Part	0.883	0.001
Total	0.904	0.000

Table 2: Pearson correlation coefficient results of Test-Retest of the Overhead Pass Skill (n = 10)

Table (2) indicates that the values of correlation coefficient between test-re-test were high enough with statistical significance where alpha level was equals or less than(.05). The highest correlation was between the practices related to part three of the preparatory period (r=.942), where as the total for the whole reliability value was 0.904, which gives indication about the status of consistency between the two practices of the first test in the battery which was the overhead pass skill. Table (2) also showed the highest and lowest reliability coefficients of the tests. These values of (r) are considered high enough for the purpose of the study.

Skill Parts	Reliability Coefficients	Significance Levels
Preparatory Parts:	0.971	0.000
Preparatory Part 1	0.968	0.000
Preparatory Part 2	0.980	0.000
Main Parts:	0.943	0.000
Main Part 1	0.856	0.002
Main Part 2	0.886	0.001
Closing Part	0.983	0.000
Total	0.986	0.000

Table 3: Pearson correlation coefficient of Test-Retest of the Ball Catching Skill (n = 10)

Table (3) indicates that the values of Pearson correlation coefficient between test-retest were also high enough with statistical significance less than (.05). The highest correlation was in the closing part of ball catching(r=.983), where as the total for the whole reliability value(0.986), which gives indication about the status of consistency between the two practices related to the skill test. In terms of ball catching which was the second skill test in the battery, the value s of (r) are considered high enough for the study purpose.

Table (4) showed the values of Pearson Product Moment Correlation Coefficient as seen below:

Skill Parts	Reliability Coefficients	Significance Levels			
Preparatory Part	0.937	0.000			
Main Part	0.988	0.000			
Main Part 1	0.722	0.018			
Main Part 2	0.931	0.000			
Closing Part	0.840	0.002			
Total	0.931	0.000			

Table4: Pearson correlation coefficient of Test-Retest of the Shooting with Forward Jump Skill (n = 10)

The values of correlation coefficient between test and re-test which were also high enough with statistical significance were less than (.05). The highest correlation was in the main part where (r=.988), where as the total for the whole reliability value was (0.931), which gives indication about the status of consistency between the two practices related to shooting with forward jump test. The values of (r) are high enough for the purpose of the study.

Objectivity coefficient

In order to check for objectivity coefficient, the researchers used two team handball trained instructors who got some training in the skill tests battery and its related parts, tables (5, 6, and 7) exemplify this matter:

Table (5) showed descriptive statistics as well as t-test between the means of the two assessors:

Skill Parts	Assessor	Means	Standard	Calculated t-	Significance
	110000001		Deviations	values	Levels
Preparatory Part	First	0.68	0.11	0.01	0.996
1	Second	0.68	0.10		
Preparatory Part	First	1.56	0.16	0.01	0.997
2	Second	1.56	0.13		
Preparatory Part	First	0.70	0.12	1.50	0.168
3	Second	0.68	0.11	-	
Preparatory Part	First	2.94	0.30	0.51	0.619
	Second	2.92	0.26	-	
Main Part	First	2.34	0.28	0.80	0.440
-	Second	2.39	0.14	-	
Closing Part	First	2.32	0.36	0.89	0.396
-	Second	2.38	0.21		
Total	First	7.60	0.89	0.56	0.587
	Second	7.69	0.52		

Table 5. t Test	Decults of both	Assessors on the	Overhood Dece	Skill (n = 10)
	NESULS OF DOLL	Assessors on the	: Overneau i ass	3KIII (II - IV)

Table(5) shows that the value of calculated t which revealed no significant differences between the two assessors while they evaluate the first test, over hand pass, where (P) value was greater than (0.05) level of significance. This gives meaning that the assessments of the two Assessors were quite close. Calculated t-values ranged from (1.50) to (0.01) as appeared in the table.

Table 6: t-Test Results of both Assessors on the Ball Catching Skill (n = 10)					
Skill Parts	Assessor	Mean	Standard	Calculated t-	Significance
			Deviation	value	Level
Preparatory Part	First	2.34	0.20	0.56	0.587
1	Second	2.30	0.26	1	
Preparatory Part	First	0.66	0.16	1.50	0.168
2	Second	0.65	0.16	1	
Preparatory Part	First	3.00	0.29	1.00	0.343
	Second	2.95	0.33	1	
Main Part 1	First	2.22	0.26	0.98	0.351
	Second	2.27	0.25	1	
Main Part 2	First	0.65	0.10	1.16	0.273
	Second	0.62	0.10	1	
Main Part	First	2.87	0.34	1.96	0.081
	Second	2.89	0.33	1	
Closing Part	First	1.49	0.14	0.55	0.591
	Second	1.50	0.12	1	
Total	First	7.36	0.62	1.00	0.343
	Second	7.34	0.61	1	

Table (6) shows that the value of calculated t which revealed no significant differences between the two assessors while they evaluate the second test, ball catching, where (P) value was greater than 0.05 level of significance. This gives meaning that the assessment of the two judges was somehow close to each others, t - test values ranged

Tublett	Test Results of b	off Hissessons on th	ie bliooting with I	of ward sump BK	· · ·
Skill Parts	Assessor	Mean	Standard	Calculated t-	Significance
			Deviation	value	Level
Preparatory Part	First	2.30	0.20	0.61	0.555
	Second	2.34	0.19		
Main Part 1	First	2.32	0.18	1.80	0.104
	Second	2.27	0.15		
Main Part 2	First	1.52	0.13	1.24	0.244
	Second	1.55	0.12		
Main Parts	First	3.84	0.30	1.96	0.081
	Second	3.84	0.30		
Closing Part	First	1.51	0.14	0.80	0.443
	Second	1.53	0.14		
Total	First	7.65	0.61	0.53	0.606
	Second	7.69	0.49]	

Table7: t-Test Results of both Assessors on the Shooting with Forward Jump Skill (n = 10)

between(1.96)and (0.55).

Table (7) shows that the values of calculated t test which revealed no significant differences between the two assessors while they evaluate the third test, shooting with forward jump, where(P) value was greater than (.05) level of significance. This gives meaning that the assessment of the two judges was somehow close to each others. Calculated t test was ranged between (1.96) and (0.53).

The second minor question to be answered was:

"How could ones assess the overall ability in team hand ball by using the proposed criterion test battery?"

To answer this question, the researchers utilized all calculated data for SS (N=60) of this study, and treat them statically. The test battery has been applied on real setting by examining the Ss by the end of spring semester and assigning grades to them. Tables (8, 9 &10) show the pertinent descriptive statistics for each skill tests as well as the mean composite scores for the sample of the study.

Tuble of means and brandard Dematoris of the overhead russ shift (n = ov)				
Skill Parts	Means	Standard Deviations		
Preparatory Parts:	2.87	0.31		
Preparatory Part 1	0.68	0.10		
Preparatory Part 2	1.49	0.16		
Preparatory Part 3	0.70	0.11		
Main Part	2.29	0.23		
Closing Part	2.23	0.26		
Total	7.40	0.74		

Table 8: Means and Standard Deviations of the Overhead Pass Skill (n = 60)

Table(8) displayed descriptive statistics such as mean scores as well as standard deviations of the two groups, where the main part of the first skill got highest mean scores with 2.29 out of 3 comparing to lowest mean scores of 2.87 out of 4 for preparatory part of the same skill test. The composite mean score for the skill test (overhead pass) was 7.40 out of 10 with standard deviation of (.74). These results indicate that the process of teaching and instruction of team handball, particularly in teaching passing skills (overhead) went very well. The level of expertise of the instructors seem to play an effective role in determining this improvements especially by utilizing behavioral objective strategy as well using teaching cues(TC) as well as (KR)and (KP) throughout teaching learning process.

Skill Parts	Mean	Standard Deviation
Preparatory Parts	2.90	0.29
Preparatory Part 1	2.25	0.20
Preparatory Part 2	0.66	0.14
Main Parts	2.93	0.25
Main part 1	2.28	0.19
Main Part 2	0.65	0.09
Closing Part	1.47	0.20
Total	7.30	0.61

Table 9: Means and Standard Deviations of the Ball Catching Skill (n = 60)

Table(9) displayed the descriptive statistics such as mean scores as well as standard deviations of the two groups, where main parts and closing part got the highest mean score with 2.93 out of 4 and 1.47 out of 2 comparing to lowest mean score of 2.90 out of 4 for preparatory part of the skill test. The total mean score was 7.30 with .61 of standard deviation. This result indicates that the teaching and learning environment of team handball, relative to ball catching went smoothly. It seems that the level of expertise of the instructors

play positive role in determining this improvements especially by utilizing some strategies such behavioral objective, teaching cues (TC) as well as knowledge of results and performance(KR)and (KP) throughout teaching learning process.

$\mathbf{r} = \mathbf{r} + $				
Skill Part	Mean	Standard Deviation		
Preparatory Part	2.31	0.22		
Main Part	3.78	0.35		
Main Part 1	2.27	0.22		
Main Part 2	1.50	0.15		
Closing Part	0.65	0.11		
Total	6.74	0.58		

Table10: Means and Standard Deviations of the Shooting with Forward Jump Skill (n = 60)

Table (10) displayed the descriptive statistics such as mean scores as well as standard deviations of the two groups, where the preparatory part got highest mean score with 2.31 out of 3 comparing to lowest mean scores of 0.65 out of 2 for closing part of the same skill test The total mean score for the performance of the sample of the study in this skill test (shooting with forward jump skill) was 6.74 with 0.58 as standard deviation. These results indicate that the process of instruction related to team handball, particularly in teaching shooting with forward jump skill along with other skills of the game by using variety of drills, went very well. Regardless the lessen value of the mean scores of the instructors played a good role in determining this improvements. This may happen by utilizing behavioral objective strategy as well using teaching cues(TC) along with modern style of teaching such as (E)the inclusion(Mosston and Ashworth,1986).

The composite mean score for the whole test battery was(21.44) out of (30), when

calculating this value to percentage, that means that (71.4) of student scores which formulate

the average scores of the Ss. The performance seems improved after training and retraining

throughout lectures during the semester by using in(closed skill) as well as open skill

training and through scrimmages along with the previously mentioned teaching strategies.

Conclusion

The researchers conclude the followings:

- The study proved that the test battery and its parts related to team handball skills were valid, reliable, and objective.

- This test battery seems applicable as a criterion –referenced tests and measurements to assess overall ability in team handball by assigning students to mastery and none-mastery levels.

Recommendations the researchers recommend the followings:

- The steps and procedure followed in constructing this battery are recommended to develop more test batteries in this game or other sports games.

- Further endeavors in this line of research shall be of great value to improve motor skill tests, especially in the area of sports science and physical education.

References:

Hudson.(1985)Prediction of Basketball skills using biomechanical variables. Research Quarterly for Exercise and Sport, 56, 2,115-121.

Kirkendall, D.Gruber, J. and Johnson.(1982).Measurement and evaluation for physical educators(2nd ed.) Champaign, IL, Human Kinetics, P.308.

Mosher, R.and Schutz, C.(1982). Physical fitness of students of British Columbia: A criterion-Referenced evaluation. Can. Jour. Of Applied Sport Sciences. 4.249-257.

Mosston, M. and Sara, Ashworth (1986). Teaching physical education (3ed, Ed.) Columbus, OH. Merrill

Popham, W. (1978). Criterion _ Referenced Measurement. N.J Englewood Cliffs

Safrit, M.J. (1977). Criterion-Referenced measurement: Application to motor skills. Theory into practice, 2, 1, 21-35.

Safrit, M.J. & C.L. Stam (1980).Reliability estimate for CR measurement in psychomotor domain. Research Quarterly for exercise and sport, 51,359-368.

Safrit, M.J. (1981). Evaluation in physical education (2nd, ed). Englewood Cliffs, N.J.p,20

Siedentop, D. Mand, C.&Taggart, A. (1986). Physical education: Teaching and curriculum strategies for grades 5-12.California, Mayfield.

Shifflett,B. &Schuman.(1982).A Criterion- Referenced Test for Archery. Research Quarterly for Exercise and Sport,53,4,330-335.