Measurement of Reliability on Volleyball Serve Assessment through Cohen's Cappa Coefficient Analysis

Fauzi[1]
[1]*Faculty of Sport Science, Universitas Negeri Yogyakarta, Jl Kolombo No 1, Yogyakarta , Indonesia*

Keywords: Measurement, Reliability, Serve, Analysis

Abstract: Research objective is to find out the reliability of volleyball serve instrument, including 1) opening or introductory, 2) warming up, 3) preparation move, 4) main move, 5) final move, and 6) closing. The subjects of research consisted of two parties, namely 1) 72 male athletes aged between 16-20 years old who were members of the Pengda PBVSI (Regional Volleyball Association) in Yogyakarta Special Region club or association, and 2) raters, i.e. six volleyball coaches. The subjects was taken by random sampling, using following criteria, namely 1) athletes who have trained for more than three years, 2) athletes who have participated in competitions, and 3) clubs that developed continuous training, while for the reviewers / trainers, they should be coaches who actively train in volleyball clubs in Yogyakarta Special Region. Instrument reliability was measured Cohen's Cappa analysis that allowed six (6) raters / assessors when the subjects played pre-game. The research results showed 1) the opening or delivery instrument is 0.85; 2) warming up instrument 0.85; 3) preparation move instrument 0.83; 4) main move instrument 0.87; 5) final move instrument 0.84; and 6) closing instrument 0.84.

1 INTRODUCTION

Volleyball is a team game that requires athletes to master playing skills. The process of volleyball training in clubs teaches technique skills for volleyball athletes and teaches how to apply them in competitive situations. The volleyball training process is focused on the athlete in developing technical skills to play volleyball, namely the skill in playing the ball.

An instrument can fulfil the requirements as good measuring tool if it can describe the true state of the object to be measured and can provide data that are related correctly and precisely. Good instrument is said to be good if it can meet the requirements of validity, reliability, objectivity and practice.

2 MANUSCRIPT PREPARATION

This research tried to measure the reliability of volleyball serve instrument items through observation by raters / assessors on male volleyball athletes aged 16-20 years who were members of PBVSI Pengda (Regional Volleyball Association Administrator) clubs or associations in Yogyakarta Special Region. The method of observation was done, so that the measured ability really expressed the real ability displayed when players play volleyball.

To determine the performance of volleyball athlete, assessment is needed. So far, there are many volleyball assessments in the form of sport skill tests. Sport skill tests have weaknesses including only one or two techniques in one skill test and require various facilities (infrastructure, high walls). Therefore, it is important and necessary to conduct research on assessment of the performance of volleyball skills that meet the reliability criteria of the instrument. Assessment of performance can be done in two ways; 1) process performance and 2) product performance. Assessment of the performance of volleyball skills in the process performance by which athletes train volleyball techniques covering all aspects of the technique with simulations or demonstrations as shown when playing or competing.

Based on the description above, it is necessary to develop the assessment of the performance of the process (before competing or pre-game). With the

Fauzi, .
Measurement of Reliability on Volleyball Serve Assessment through Cohen's Cappa Coefficient Analysis.
DOI: 10.5220/0009796505470553
In *Proceedings of the 3rd Yogyakarta International Seminar on Health, Physical Education, and Sport Science in conjunction with the 2nd Conference on Interdisciplinary Approach in Sports (YISHPESS and CoIS 2019)*, pages 547-553
ISBN: 978-989-758-457-2

assessment of the performance of the process, it will be easier for the trainer to see the progress of the athletes to achieve optimal volleyball skills. The assessment that was developed in this research was the assessment of volleyball skills on the performance of the process (before playing or pre-game) of volleyball serve technique.

2.1 Assessment of Training- Coaching Volleyball

Development and implementation of training program performed by coaches influences the assessment applied to athletes at club. Anderson (2003: xi) states that assessment is the process of gathering information to make decisions. Lund & Linn (1990: 5) claim that assessment is a systematic process in collecting, analysing, and interpreting information that includes work performance, knowledge, attitudes, skills, work tasks that they know and can do their jobs.

Assessment includes all the methods used to collect data about individuals (Mardapi, 2012: 12). The assessment focuses on the individual, so the decision is also on the individual. Huba & Freed (2000: 8) define assessment as a process of gathering and testing information to improve the clarity of understanding of what has been learned by learners or students from their experiences. Anderson (2003: 15) explains assessment is an action that is very closely related to decision making. The more the number of incidents of decision making from the assessment of training, the more serious the consequences and implications in the long term.

The training- coaching volleyball in clubs has various objectives, including move techniques, strategy, tactics, cooperation, and fairness skills. Other supporting factors are used as illustration by the coach to assess athletes, while training- coaching also emphasize on aspects of knowledge and behaviour. The results of training- coaching are certain competencies or abilities both cognitive, affective, and psychomotor that are achieved or mastered by athletes after following the training-coaching process.

A coach has assessed the training- coaching results involving measurements using measuring instruments such as athlete's move observation and the existing volleyball test (sports skill test) that has been adopted. The assessment is carried out by the coach by using the volleyball skill instrument on the sports skill test without realizing there was a weakness, namely the existence of culture bias. The weakness of culture bias is that for tests that are adaptive in nature, there are several norms that are not necessarily suitable to be applied or used in other cultures. The use of athlete's move observations was not accompanied by the preparation of norms as a reference for the success of athletes in the process of practicing.

The assessment process involves gathering evidence about the achievement of training results of athletes which in the implementation involves a measurement process. Measurement is the determination of numbers in a systematic way to state the individual's condition or object (Mardapi, 2012: 5). Systematic way means that measurement activities are carried out through certain procedures or certain rules, so that the results of measuring the state of the same object, even though different places and times will give the same measurement results.

Some of the main principles that must be considered by coaches in assessment process for athletes, according to Popham (1995: 16) are validity, reliability and absence-of-bias. According to Nitko and Brookhart (2007: 38), validity is the accuracy of the interpretation and usefulness of the results of the assessment. Therefore to validate the interpretation and usefulness of the assessment, it must combine evidence from other sources which indicate that the interpretation and usefulness of the assessment results are appropriate. Validity is a result of judgment made after considering various evidences from various relevant sources.

Reliability is coefficient that shows the level of consistency of the results of the test (Mardapi, 2012: 51). Consistent measurement results are using the same measuring instrument for different people or at different times but the same conditions. The degree of reliability of the assessment results determines the level of confidence of the results achieved. The reliability of an assessment result does not guarantee the validity of the assessment results. It is just that reliability increases trust in determining decisions related to assessment results.

Thus, it can be concluded that the training-coaching assessment is all the methods used by coach to collect data on the results of training athletes through systematic measurement.

2.2 Psychomotor Assessment

Assessment of psychomotor aspect or skill is the success of playing the ball with various volleyball techniques. Reid (2010: 1) states that in order to determine the development and improvement of athletes and motivate them, the coach must use the opportunity to measure the success of the exercise by making an assessment at the end of the year. Assessment of volleyball skills basically rests on the activities of athletes with many using performance assessments. With the performance assessment it is

expected that athletes can demonstrate activities in the form of volleyball techniques. According to Stiggins (1997: 34), performance assessments call upon the examinee to demonstrate specific skills and competencies, that is, to apply skills and knowledge they have mastered.

Performance assessment of volleyball skills is basically a process of gathering evidence of achievement that relies on the performance of athletes. That is, the performance assessment in its implementation involves athletes in an activity that demands to demonstrate its capabilities in the form of processes and products. This assessment is expected that athletes can demonstrate that they can do certain tasks, such as doing volleyball playing techniques. In performance assessment the important thing is how the coach observes and scores the athletes' performance abilities. Therefore, coaches as raters must pay attention to the validity and reliability factors of the measuring instrument so that the expected results of the assessment are not subjective. In addition, in carrying out the assessment of volleyball skills the trainer needs assessment guidelines that aim to facilitate the assessor or rater in assessing so as to minimize measurement errors.

2.3 Performance Assessment

Performance assessment or often referred to as authentic assessment (authentic assessment) (Metzler, 2005: 180). Authentic assessment in other books (except Wiggins) is equated with the name of alternative assessment (performance assessment) or performance assessment (Herman & Winters, 1992: 146). According to Zainul (2005: 3), performance Assessment is a variety of tasks and situations where test participants are asked to demonstrate understanding and application of in-depth knowledge, as well as skills in various contexts. This means that performance assessment or performance assessment is an assessment that asks test takers or athletes to perform performance or demonstrate their knowledge in various contexts according to the desired criteria.

Morrow (2005: 131) states that assessment is a real life setting, one that is less contrived and artificial than traditional forms of testing. This means that an assessment takes the form of a task designed in accordance with real life. According to Mueller (2006), authentic assessment is a form of assessment in which students are asked to present assignments in real situations by demonstrating the application of essential skills and knowledge (direct assessment).

Based on the description above, it can be concluded that performance assessment is an assessment that asks test takers to perform performance or demonstrate their knowledge and skills into various contexts according to the desired criteria.

Performance assessment is characterized as assessment that emphasizes the form of assignments so that the results of the work are in the form of athletes' work called products, and printing rubrics or guidelines developed and designed according to the content of the task and used to assess athletes' products. There are two important components in performance assessment, namely task performance and performance rubrics.

Task is assignment that will be performed to make performance assessment, while the rubric consists of a list of criteria that are realized with work dimensions, process aspects or concepts that will be assessed and quality gradations ranging from the most perfect level to a bad level (Zainul, 2005: 13). Important properties in performance measurement require reliability. Johnson, *et al* (2009: 22) state that reliability means the consistency of scores across such factors as occasions, tasks, and raters. in others words, reliability addresses whether examining scores would be the same if she was to take the exam on a different occasion, complete different tasks, or scored by different raters. In the performance assessment using the inter-rater reliability method, it requires the assessor to assess or measure some aspects of the behaviour of students or athletes in the assignment. To achieve high level of reliability among raters, it is necessary to design and apply the rubric. The development of the rubric is very necessary in the performance assessment, which is used as the basis for measurement.

Rubrics are guidelines for the product of assessment. Rubrics or criteria are guidelines for giving clear scores and agreed upon by teachers and students (Zainul, 2005: 9). This guide explains to teachers and students about the standards that exist in performance (Herman & Winters, 1996: 47). Lund and Mary (2002: 43) argue that rubric indicates the criteria for a person scoring performance-based assessment should use when doing evaluations. Furthermore, the criteria for performance evaluation are often called scoring criteria, scoring guidelines, rubrics, and rubric scoring but have the same meaning (Lund & Mary, 2002: 44). The design of the assessment rubric requires specifications of the criteria to assess the quality of performance and choice of assessment procedures (Brenan, 2006: 394). The criteria in the rubric describe the important elements of performance and become the source of the criteria to be assessed.

Criteria that are clearly the essence of performance assessment so students can be assessed consistently (Arter, 1996: vi). The rubric criteria used by teachers are product criteria (Guskey 1996b: 4). The product criterion in this case is about the results that are clearly displayed by athletes in the volleyball club.

In the process of training- teaching sports in the club, the product criteria are the performance of the athlete in the accuracy of the ball towards the opposing field, the accuracy of the ball being directed and passed to team mates in the team and to block the ball from the opponent's smash attack. Assessment of the performance of volleyball skills where athletes are asked to display the process performance of performing sports techniques and display the results of the accuracy of the ball.

In performance assessment, it involves two kinds of relevance, namely in accordance with competence and meaningful in real life. That is, a performance assessment criteria or indicators must be in accordance with the competencies or abilities trained by athletes and meaningful with real life.

2.4 Assessment of Performance-based Volleyball Skills

In assessing the achievement of the competencies of athletes in volleyball clubs, the coach needs to apply comprehensively in assessing the success of practicing volleyball during training by paying attention to the indicators of the athlete's abilities / performance in volleyball which are described from the training program indicators.

Assessment is an important and inseparable part of the practice of practicing volleyball, so the coach will make continuous assessments as part of the training-coaching situation. Assessment can apply various ways and use of various assessment tools to obtain information on the extent to which learners' learning outcomes or the achievement of competencies (sets of abilities) of students (Carol, 1994: 5).

Training- coaching volleyball technique involves many various skill abilities. Skill skills can be assessed in various ways, observation is one of the best judgments available. Coaches and athletes can observe skills and provide assessments designed to improve the process of practicing volleyball training in clubs. Observation is an assessment performed by observing the athlete during the training process and / or outside the training activities. Observations are made to collect qualitative and quantitative data in accordance with the competencies assessed, and can be done both formally and informally.

In the process of training, it mostly involves athlete behaviour, the use of coach observation is one of them which is currently the most practical. The coach gives an assignment to display, then sees how well the athlete's movements behave. Observations on performance are done to collect data, so that it can be seen how far the athlete has mastered a skill based on the performance displayed during, after, and or after several times practicing training at the club.

3 RESEARCH METHODS

The type of data in this research was quantitative. Quantitative data were obtained through the results of measuring the performance of the technique of volleyball skills. Pre-match technical process performance (pre-game), where athletes simulated or demonstrated each technique of volleyball sports with indicators: 1) opening or introductory, 2) warming up, 3) preparatory move, 4) main move, 5) final move, and 6) closing. Indicators are performed on serve techniques, bottom fitting, top fitting, hitting, and blocking mastered by athletes. Data from the measurement of the performance skills of the volleyball serve assessment process was obtained using pre-match assessment instruments (pre-game).

The instrument of this research used a volleyball grid to measure the level of achievement of volleyball serve skills with a volleyball athlete as the respondent. The data collection technique used the athlete's performance tasks, namely the athlete's job performance in the form of pre-match simulation (pre-game) which was observed using observation guidelines equipped with a rubric sheet.

The analysis used to find out the reliability coefficient of the instrument was inter-rater reliability analysis (inter-rater reliability). The assessor or rater in the research used 6 (six) raters and the one who became the rater was the volleyball coaches who had national coach certification. The reliability coefficient of the volleyball sports skills assessment instrument used SPSS assistance with the Cohen Cappa K. program The Cohen Cappa K program aims to find out the understanding of the assessor, the understanding of the assessor when observing athletes in performing the pre-game process with serve techniques that include introduction, warming up, preparatory move, main move, final move, and closing. The amount of the Cohen Cappa coefficient produced compared to the minimum permissible criteria was 0.70 (Mardapi, 2012). The results of volleyball skill assessment on volleyball serve technique was analyzed by using

descriptive analysis, which described the assessment done by athletes during performing performance tests on serve technique.

The subject of this research consisted of two elements: 1) athletes totalling 72 male athletes aged 16-20 years who were members of Pengda PBVSI clubs in Yogyakarta and 2) volleyball coaches (raters) who were (6) coaches. Determination of research subjects was taken by random sampling, with criteria, namely: 1) athletes who had practiced for more than three years, 2) athletes who had participated in the competition, and 3) clubs that developed continuous training, while the reviewers / trainers were trainers who actively trained in volleyball clubs in Yogyakarta Special Region. Measurement of instrument reliability used Cohen's Cappa coefficient analysis involving six (6) raters / assessors when the subject played pre game. Instruments can be said to be good if the reliability of the instrument shows that the reliability coefficient is greater than or equal to 0.60, meaning that the instrument can provide level of understanding among the raters.

The type of data in this research is quantitative. Quantitative data is obtained through the results of measuring the performance of the technique of volleyball skills. Pre-match technical process performance (pre-game), where athletes simulate or demonstrate each technique of volleyball sports with indicators: 1) opening or introductory, 2) warming up, 3) preparatory move, 4) main move, 5) final move, and 6) closing. Indicators are performed on serve techniques, bottom fitting, top fitting, hitting, and dam controlled by athletes. Data from the measurement of the performance skills of the boa volley serve assessment process was obtained using pre-match assessment instruments (pre-game).

The instrument of this research used a volleyball grid to measure the level of achievement of volleyball serve skills with a volleyball athlete as the respondent. The data collection technique used the athlete's performance tasks, namely the athlete's job performance in the form of pre-match simulation (pre-game) which was observed using observation guidelines equipped with a rubric sheet.

The analysis used to find out the reliability coefficient of the instrument was inter-rater reliability analysis (inter-rater reliability). The assessor or rater in the research used 6 (six) raters and the one who became the rater was the volleyball coaches who had national coach certification. The reliability coefficient of the volleyball sports skills assessment instrument uses SPSS assistance with the Cohen Cappa K. program The Cohen Cappa K program aims to find out the understanding of the assessor, the understanding of the assessor when observing athletes in performing the pre-game process with serve techniques that include introduction , warming up, preparation move, main move, final move, and closing. The amount of the Cappa K Cohen coefficient produced compared to the minimum permissible criteria is 0.70 (Mardapi, 2012). The results of volleyball skills assessment on volleyball serve techniques were analysed using descriptive analysis, which describes the assessment carried out by athletes during performing performance tests on serve techniques.

The subject of this research consisted of two elements, namely: 1) athletes totalling 72 male athletes between the ages of 16-20 years who were members of the PBVSI Pengda club in Yogyakarta, and 2) a volleyball coach (rater) who there are six (6) trainers. Determination of research subjects was taken by random sampling, with criteria, namely: 1) athletes who have practiced for more than three years, 2) athletes who have participated in the competition, and 3) clubs that carry out continuous training, while the reviewers / trainers are trainers who actively train in volleyball clubs in the Special Region of Yogyakarta. Measurement of instrument reliability uses Cohen's Cappa coefficient analysis involving six (6) raters / assessors when the subject plays pre game. Instruments can be said to be good if the reliability of the instrument shows that the reliability coefficient is greater than or equal to 0.60, meaning that the instrument can provide a level of understanding among the raters.

4 RESULTS AND DISCUSSION

Data results can be presented by inter-rater coefficient analysis. The inter-rater coefficient is one of the means to see the level of consistency or consistency between the assessors / raters in giving ratings to the performance of athletes in demonstrating volleyball skills at the club. For the purposes of testing the consistency of inter-rater or appraiser, researchers used the coefficients of Cohen's Cappa.

Data collection of the research was done by 6 raters / assessors who gave rating on the assessment of volleyball serve technique skills. In this assessment there are 6 (six) items that are the objects of research, namely, opening / introductory, warming up, preparation move, main move, final move, and closing. The summary of the results of the calculation of consistency and agreement of the six raters on volleyball skills in serve technique is presented in Table 1.

Table 1: Results of the K (Cappa) Coefficient among Raters on Volleyball Serve Technique Skill

Item	Raters (Assessors)															Mean
	.2	1.3	1.4	1.5	1.6	2.3	2.4	2.5	2.6	3.4	3.5	3.6	4.5	4.6	5.6	(Average)
1	.89	.92	.79	.82	.81	.87	.78	.81	.81	.82	.84	.79	.92	.87	.95	.85
2	.92	.87	.82	.77	.77	.84	.85	.75	.80	.95	.90	.85	.90	.90	.90	.85
3	.82	.84	.76	.79	.83	.86	.79	.77	.84	.84	.91	.81	.84	.88	.86	.83
4	.89	.90	.85	.87	.90	.89	.84	.82	.85	.84	.87	.85	.92	.90	.93	.87
5	.92	.87	.75	.82	.80	.84	.82	.80	.82	.78	.85	.78	.87	.90	.93	.84
6	.88	.95	.76	.83	.81	.84	.83	.76	.83	.81	.88	.81	.83	.90	.83	.84
Mean	.89	.88	.79	.82	.82	.86	.83	.78	.83	.84	.88	.82	.87	.90	.89	.85
Overall Mean of Cappa Coefficient														**0.85**		

Table 1 shows that the results of the average K coefficient (Cappa) serve technique in the opening item 1 instrument (introduction) with an assessment of 6 raters are 0.85. The average results of the K (Cappa) coefficient of serve technique in the item 2 warming up instrument with assessment of 6 raters are obtained 0.85. The results of the mean K coefficient (Cappa) serve technique in the item 3 of the preparatory move with assessment of 6 raters were obtained 0.83. The average results of the K (Cappa) coefficient of serve technique in the item 4, move instrument with gesture move assessment of 6 raters were obtained 0.87. The results of the average of K coefficient (Cappa) on serve technique in the final instrument item 5 with 6 rater's ratings are obtained 0.84. The results of the mean K coefficient (Cappa) serve technique in the closing instrument item 6 with an assessment of 6 raters are obtained 0.84.

The level of consistency and agreement of the overall assessors in assessing volleyball serve technique skills at clubs can be determined by taking the average Cappa Coefficient by six raters 0.85. The value of 0.85 illustrates that the six assessors (raters) have perception and understanding of the valuation construct of 85%. The K (Cappa) coefficient value is greater than the minimum criteria used at 0.75, so the instrument meets the reliability coefficient requirements.

5 CONCLUSION

Based on the research results through Cohen's Cappa coefficient analysis, the reliability of serve technique instruments can be obtained as follows: 1) Reliability of opening or introductory instrument is 0.85; 2) Reliability of warming up instrument is 0.85; 3) Reliability of the preparatory move instrument is 0.83; 4) Reliability of the main move instrument is 0.87; 5) Reliability of the final move instrument is 0.84; and 6) Reliability of the closing instrument is 0.84.

From the conclusion above, that the use of the Cohen's Cappa coefficient needs to be considered in measuring the reliability of the instrument because it produces reliability that satisfies the criteria of good test.

REFERENCES

Anderson, L., 2003. *Classroom assessment: enchanting the quality of teacher decision making. Mahwah*, NJ: Lawrence Erlabaum associates.

Arter, J., 1996. *Performance criteria: the heart of the matter. Ed. R. E. Blum & J. A. Arter. A handbook for student performance assessment.* Alexandria, VA: Association for Supervision and Curiculum Development.

Bob, C., 1994. *Assessment in physical education: a teacher's guide to the issue*. London: The Flamer.

Brennan, R. L., 2006. *Educational measurement*. Westport: Praeger.

From, M., 2012. *Measurement of education evaluation and evaluation*. Yogyakarta: Nuha Litera.

Gronlund, N. E., Linn, R. L., 1990. *Measurement and evaluation in teaching.* New York: MacMillan Publishing Company.

Guskey, T., 1996. *Alternative ways to document and communicate student learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Herman, L.P., Aschbacher, Winters, L., 1992. *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curiculum Development.

Huba, M.E., Freed, J. E., 2000. *Learner-centered assessment on college campuses. Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.

Lund. J. L., & Mary, F. K., 2002. *Performance based assessment for middle and high school physical education.* USA: Human Kinetics.

Morrow, J. R., et al. 2005. *Measurement and evaluation in human performance*. United States of America: Human Kinetics.

Metzler, Michael W., 2005. *Instructional models for physical education second edition*. USA: Holcomb Hathaway Publisher.

Mueller, J., 2006. *Authentic assessment toolbox*. Taken in December 2011.

Nitko, A.J., Brookhart, S.M., 2007. *Educational assessment of students (5th Edition)*. Upper Saddle River, NJ: Pearson Education.

Popham, W. J., 1995. *Classroom assessment*. Boston: Allyn and Bacon.

Reid, L., 2010. *Here's a quick, easy way to improve your volleyball team's skills, without wasting time or effort*. Taken February 12, 2012.

Stiggins, R., 1997. *The design and development of performance assessments. Educational Measurement: Issues and Practice 2nd ed*. Upper Saddle River, NJ: Prentice Hall.

Zainul, A., 2005. *Alternative assessment*. Jakarta: Director General of Higher Education.