Browse By Title

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

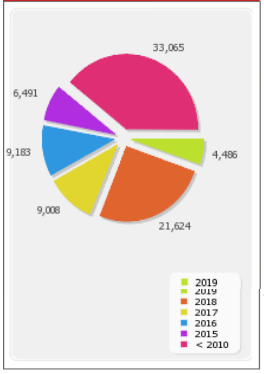
Browse By Subject

- Arts & Humanities (191)
- Engineering & Technology (160)
- Medical & Health Sciences (87)
- Sciences (179)
- Social Sciences (357)

Statistics

Latest Uploads
Most Viewed Articles
Most Downloaded Articles
Most Viewed Journals

Articles uploaded in MyJurnal



Administrator Login

International Journal of Assessment and Evaluation in Education
Username:
Password:
Login
Kindly contact MCC if you forgot your password.

International Journal of Assessment and Evaluation in Education

Volume 6, December 2016

1
Olutola, Adekunle Thomas.
Empirical analysis of item difficulty and discrimination indices of senior school certificate multiple choice biology tests
Download Full Text
View abstract
View references

2
Thamburaj, Kingston Pal; Jabamoney, Samikkanu Isaac Samuel; Shakunthala, R.
Standardization of question items to test Tamil case markers – a study
Download Full Text
View abstract
View references

3
Mostafa Janebi Enayat; Seyed Mohammad Reza Amirian.
Vocabulary levels test and word associates test: can they measure language proficiency?
Download Full Text
View abstract
View references

4
Sadegh Soheili Anaraki; Seyed Mohammad Reza Amirian; Moslem Zolfagharkhani.
Development and validation of a self-efficacy scale with writing and social factors
Download Full Text
View abstract
View references

5
Tomoliyus; Sumaryanti; Herka Maya Jatmika.
Development of validity and reliability of net game performance-based assessment on elementary students' achievement in physical education
Download Full Text
View abstract
View references

6
Ling, Weay Ang; Mona Masood; Siti Hawa Abdullah.
Analysing the relationship of sequential and global learning styles on students' historical thinking and understanding: a case study on form four secondary schools students in Malaysia
Download Full Text
View abstract
View references

Back

Journal Coverage

Volume 8, December 2018
Volume 7, December 2017
Volume 6, December 2016
Volume 5, December 2015
Volume 4, December 2014

Archive
Latest Issues

Indexation

Indexed by	MyJurnal (2017)
H-Index	0
Immediacy Index	0.000
Rank	0

Other journals

- International Journal of Property Sciences
- JUMMEC
- Journal of Surveying, Construction and Property
- SARJANA
- SEJARAH : Jurnal Jabatan Sejarah Universiti Malaya
- Jurnal Pengajian Media Malaysia

International Journal of Assessment and Evaluation in Education
Vol 6/Dec 2016 (41-49) ISSN 2232-1926
Submission Date: 19th July 2016
Acceptance Date: 12th August 2016

Development of Validity and Reliability of Net Game Performance-Based Assessment on Elementary Students' Achievement in Physical Education

Tomoliyus[1]
tomoliyus@yahoo.com

Sumaryanti[2]
sumaryanti@uny.ac.id

Herka Maya Jatmika[3]
herka_mayajatmika@uny.ac.id
Faculty of Sports Science, Yogyakarta State University, Indonesia[1,2,3]

Abstract

The purpose of this study was to seek the validity and reliability of net game's performance-based assessment on the context of elementary school physical education students' achievement. Research and development used as a primary research method. The subject were forty students in Yogyakarta's table tennis school club which chose through random sampling technique. Content Validity Ratio (CVR) is used to test the validity, on the other hand, Anava-general Multifacet model for the inter-rater reliability. the results were as follows: CVR of the decision making is 0.714; CVR of the skill acquisition process is 1; CVR of the skill acquisition product is 1; and CVR of the returning on beginning position is 0.714. as for the decision making inter-rater reliability r = 0.785; skill acquisition process reliability r = 0.872; skill acquisition product r = 0.785; and returning on beginning position r = 0.885. As the conclusion, performance-based assessment instrument of net game, particularly table tennis, has high scores of validity and inter-rater reliability.

Keywords *performance-based assessment instruments, content validity, inter-rater reliability, net game, table tennis*

INTRODUCTION

The efforts in achieving competency for the net games extracurricular learning outcomes in elementary school can be reached through increased learning and assessment. Both are interlinked with each other, because a good learning will produce good judgment. Therefore, in an effort to improve the learning outcomes of students there must be some balance among assessment, learning materials, and learning strategies. In other words, a good assessment instrument is the one that the judge can assess the learning outcomes of the learning material which are contextual and real (authentic). The contextual and authentic assessment is called the performance-based assessment or an alternative assessment (Lund, 2010: 19).

According to Stiggins (Mueller, 2009: 102), performance-based assessment is the one that asks the students to demonstrate certain skills and competencies as it is the application of essential knowledge. According to Mueller (2009: 106), a performance-based assessment is a form of a task that requires students to demonstrate the real-world performance significantly which is the essence of the application of knowledge and skills. Suzann (2000: 3) states that a performance-based assessment is applied in real-life situations, which require students to use higher-level thinking skills, such as problem solving and decision making skills. Lund (2010: 19) states that the characteristics-based assessment of performance or authentic assessment, require the specific tasks which are designed to represent the performance, emphasize higher level thinking and learning more complex, articulate criteria in advance so that students know how to evaluate, expect the students to present the results work in public whenever possible, and involves the examination process and learning products.

From some of these opinions, it can be concluded that the performance-based assessment on the net game extracurricular is emphasizing performance-based assessment, to do something with the application of science which has been mastered theoretically. Moreover, performance-based assessment requires the students to demonstrate the knowledge, skills and strategy with the creation of an answer or products. Performance-based assessment also requires students to use the higher-level thinking skills, such as problem solving and decision making skills. Performance-based assessment is an assessment of the performance-based learning refers to the situation or context of the "real world" that requires a variety of approaches to solve the problem that the possibility could have more than one solution. In other words, performance-based assessment is to monitor and measure students' skills in a variety of possibilities for solving the problems encountered in real situations or contexts. In a learning process, the performance-based assessment is trying to measure, monitor and assess all aspects of learning outcomes (included in the domain of cognitive, affective, and psychomotor), which is seen as a process of learning, as well as in the form of change and development activities, and the acquisition of learning during the learning process in the classroom and outside the classroom.

On the basis of these descriptions, then a Physical Education teacher must master the physical education assessment for the student learning outcomes, and one of them is the performance-based assessment. Question, Should a Physical Education teacher master a performance-based assessment? There are some opinions that they should master it but there are also some opinions that they should not master the performance-based assessment. In addition, there is a question if the net games assessment instruments used by Physical Education teachers are in accordance with the competencies being taught? There are some opinions that it is in accordance with the competencies taught, some other opinions show that it is a little fit with the competencies taught, and also there is an opinion that it is not in accordance with the competencies taught.

The survey results (2010) interviewed 43 Physical Education teachers in elementary schools in Yogyakarta indicated that 96 percent of Physical Education teachers in elementary schools did not master yet the performance-based assessment. In addition, 93 percent was generated that the teacher assessment instrument was not in accordance with the competencies taught. The Physical Education teachers in elementary school in assessing the learning outcomes of some net games extracurricular were not using the parallel instrumental assessment of the net games learning outcomes towards the learning materials, they used some tests of physical fitness and skill tests existing in the books, for example, volleyball game learning outcomes was using Brady Tests, and the results of learning techniques using test sprint athletics 30 meters or 50 meters. Some physical education teachers have criticized that the test measurement in the book cannot measure the competency of the tangible net games learning outcomes.

From the description above, there is a gap in the assessment of student learning outcomes, as they all do not understand and master the Physical Education performance assessment of students. It causes the gap, because of the unavailability of Physical Education teachers' performance assessment instrument, they did not master it well on the performance assessment, and they also arrange their own biased assessment instruments. Based on the cause of the gap, it is necessary to look for a solution to address these gaps. One solution is by developing a performance-based assessment instruments on the net games material in elementary school that are valid and reliable.

Assessment instrument plays a very important role in determining the quality of a learning outcomes, because the validity of the data obtained will be determined by the validity of the instruments used, in addition to data collection procedures were adopted. If the assessment instruments have a valid and reliable quality in high category then the data obtained will correspond to the facts or the real situation. While the quality of assessment instruments has a low validity and reliability, the data obtained is not valid or not in accordance with the facts, which can lead to erroneous conclusions.

Validity is the main thing that must be owned by an assessment instruments. It is because the instrument has a high validity implication for the result itself. According to Nitko and Brookhart (2011: 38), the validity of the interpretation is the accuracy and usefulness of the results of the assessment. According to Anwar (2012: 112), validity is the ability of a test to measure what should be measured. Based on these opinions, it can be concluded that the validity is the degree of accuracy of a relevant instrument to measure what should be measured.

According Sugiyono (2006: 45), the validity test is a step of tests performed on the content of an instrument. The aim of validity test is to determine the extent to which the precision and accuracy of a measurement instrument in performing measuring function.

Although the statistical and psychometric correlation coefficient cannot be used to assess the validity of the content, several approaches have been proposed by experts to help to measure the content validity for instance, the approach developed by Lawshe (1975), which proposes the ratio of content validity (content validity ratio / CVR). It can determine whether it is the content, the validity of the instrument qualifies or not used 2 formula which is from the formula of Schultz & Whitney (2005) to see the content validity of each item and formula of Gregorry (2007) to see the overall validity of the content.

Reliability according to Sukadji (2000: 34) is a test to measure the degree of target measured consistently. According to Sugiyono (2006: 48), reliability is a series of measurements or series of gauges that have consistency as the measurements made by the measuring instrument are done repeatedly.

Reliability is the degree of consistency / regularity assessment of a repetition of the assessment procedure. The degree of reliability of the assessment results may determine the level of confidence of the results. Reliability is expressed in the form of numbers, usually as a coefficient. High coefficient means high reliability. Reliability of an assessment does not guarantee the validity of the assessment results. Only reliability may increase the confidence in determining the relevant decision outcome assessment.

Reliability test is the measurement of the accuracy (consistent) of an instrument (Husaini, 2003: 64). This test is intended to ensure that the instrument used is an instrument that has the consistency of the scores given by the assessor one another, so that when used repeatedly it can generate the same data. There are two kinds of inter-rater reliability test, namely the test agreement Inter-rater correlation coefficient of Kappa and inter-class correlation coefficient test (intraclass correlation Coeffisients/ICC) (Wahyu Widhiharso, 2006: 150). Test agreement Inter-rater correlation coefficient of Kappa will be used if there are 2 raters. Inter-class correlation coefficient test (Intraclass Correlation Coefficients) will be used if the raters are more than two people.

According to Griffin (2007: 10), Grehaigne (2005: 3), Memmert, D. & Harvey, S. (2008) and Lund (2010: 225) net game includes tennis, table tennis, badminton, and volleyball. The game of table tennis, tennis, badminton and ball games is in the small groups in the curriculum. The net game has such problems in tactics such as the tactic in scoring and preventing the score. Problems in scoring include the scoring and creating the space to attack, use the space to attack. Preventing includes maintaining the space, maintaining the score, and returning the ball.

To create a score and defend score in the game required skills base (the base), decision making (decision making), skill execution (implementation skills), support (support), marking (care), cover (cover), and adjust (adjustment) (Lund, 2010: 225; Griffin, 2007: 220).

Of the seven aspects of the game described above, in the net game performance the most important component is that the student is able to carry out the process and product skills for efficient and effective (skill execution) and able to think effectively in taking the right decision to choose or execute movements with ball in accordance with the conditions of the game (decision making), and returned to the position after receiving and returning the ball (base).

AIM OF THE RESEARCH

This research aimed in finding the content validity and inter-rater reliability value of the performance-based assessment instrument for the Physical Education learning outcomes of net game material for the elementary school students.

METHODS

The research employed research and development design. The steps of instrument development, namely: preliminary studies, developing initial assessment tools, expert validation, testing, analyzing the validity (validity) and reliability (reliability). The research subjects were for about 40 table tennis extracurricular club members in elementary school, taken by random.

Statistical Analysis

Analysis of the data for the validity test was using the Content Validity Ratio (CVR) and for inter-rater reliability test it employed Anava-General Multifacet Model.

RESULTS

Results of the Guide Book for the Performance-based Assessment

The results have shown the guide book assessment instrument based on performance results of learning materials game net primary school students containing factors and indicators of the net game (mental, process skills, product skills and attitudes and behavior), the task of the net games, instructions and guidelines for filling the observation sheet, observation sheets, and assessment criteria.

Table 1 Factor dan indicator of net game (individual)

	Factor	Indicator
Net Games	Decision Making	The right *decision making* is done right before showing the skills (have no doubt)
	Process of skills	The process of *skill execution* is efficient in the phase of preparation, implementation, and further steps.
	Product Skills	It is effective in gaining the score or defending the score
	Return to Base	Return to the base or retraced on the early position (the "T" position) right after hitting the ball and energetic in doing it.

Task

Athletes or the students are given the task in playing single table tennis, badminton, and tennis for one game.

Directions and Guidance on the Observation Sheets Filling

The filling directions on the Observation Sheets of Decision Maker
Put a mark "+" (plus) and a score of 3, if the decision makers always show the right decision during the play between 90% -100%.
Put a mark "V" (tick) and a score of 2, when the decision-makers show the decision making score between 66% -89% during the play.
Put a mark "-" (minus) and a score of 1, if the decision makers only show the tiny score between 0-65% for the decision making during the play.

Guidance

Mark	Indicate
+(3)	The decision makers are always right during the play
√(2)	The right decision making appear a lot during the play
- (1)	The good decision making never show up

The filling directions for the Observation Sheets of Process of Skills

Put a mark "+" (plus) and a score of 3, if the process of skills always appear during the play between 90% -100%.

Put a mark "V" (tick) and a score of 2, when the process of skills appears during the play between 66% - 89%.

Put a mark "-" (minus) and a score of 1, if the process of skills only shows the tiny score between 0-65% during the play.

Guidance

Mark	Indicate
+(3)	The efficient process of skills appears a lot during the play
√(2)	The efficient process of skills show up quite a lot during the play
- (1)	The efficient process of skills never appear during the play

The filling directions for the Observation Sheets of Product of Skills

Put a mark "+" (plus) and a score of 3, if the product of skills always appear during the play between 90% -100%.

Put a mark "V" (tick) and a score of 2, when the product of skills always appears between 66% -89% during the play.

Put a mark "-" (minus) and a score of 1, if the product of skills only shows the tiny score between 0-65% for the decision making during the play.

Guidance

Mark	Indicate
+(3)	The effective product of skills appear a lot during the play
√(2)	The effective product of skills shows up quite a lot during the play
- (1)	The effective product of skills never appear during the play

The filling directions for the Observation Sheets of Return to the Base

Put a mark "+" (plus) and a score of 3, if the position of return to the base during the play between 90% - 100%.

Put a mark "V" (tick) and a score of 2, when the position of return to the base shows the score between 66% -89% during the play.

Put a mark "-" (minus) and a score of 1, if the position of return to the base only shows the tiny score between 0-65% during the play.

Guidance

Mark	Indicate
+(3)	Always return to the base during the play
√(2)	Returning to the base quite a lot during the play
- (1)	Never returning to the base during the play

Assessment Criteria

Besides the observation sheet for the decision makers, process of skills, product skills and attitudes and behavior during play, teachers or coaches also use the scoring sheet for the final assessment of each instrument. For this purpose, the researchers use the following criteria

The Criteria

Interval	Indication
2,25<Score≤3,00	Good
1,50<Score≤2,25	Enough Mediocre
1,00≤ Score≤1,50	Less

Model of Observation Sheets

No	Name	Factor					
		Dm	Pros	Prod	Rb	Total	Criteria
1	Yanto	2	3	2	3	2,5	Good
2	Edy	2	2	1	3	2.0	Enough
3	Tono	2	3	3	3	2.75	Good
4	Tatik	1	2	1	1	1.25	Less
dst							

Content Validity Results

The data result from the assessment of 7 experts shows 3 different categories: important, quite important, not really important. Look at the table below.

Table 2

Assesor	Item Judgement			
	Mental	Process of Skills	Product of Skills	Attitudes and Behavior
A	Important	Important	Important	Important
B	Important	Important	Important	Important
C	Not really important	Important	Important	Important

D	Important	Important	Important	Not really important
E	Important	Important	Important	Important
F	Important	Important	Important	Important
G	Important	Important	Important	Important

Based on the Table above, it can be concluded that from 7 assessors expert, item 1 (decision making) 6 persons state that it is important, 1 person states that it is not really important, dan there is no one states that it is not useful at all. Item 2 (process of skills), 7 persons state that it is important. Item 3 (product of skills) 7 persons state that it is important. Item 4 (return to the base) from 7 persons, 6 persons state that it is important, 1 person states that it is not really important, and no one states that it is not useful at all. The data is counted by *content validity ratio* (CVR) with Lawshe's formula as follows:

CVR = (ne- N/2)/ N/2)
CVR is *content validity ratio*
Ne is the total persons who choose it important
N is the total persons
The calculation results of CVR is shown as follows:

Tabel 3

Item	CVR
Decision Making	0.714
Process of Skills	1.00
Product of Skills	1.00
Return to the Base	0.714

Based on the table above, it can be concluded that the items of decision making, process of skills, product of skills, and return to the base have the very good content validity. It is because the score of CVR for those items is more than 0.500. So this performance-based instrument is right in measuring the net game learning outcomes.

Inter- Rater Reliability Results

Based on the test of data of performance-based assessment instruments net game material for elementary school students on mental items, it shows that the reliability value is estimated using Alpha coefficient, shows a high coefficient with r = 0971. Analysis is done by using Anova-General multifacet Model (ICC). Interclass correlation coefficient value indicates that the inter-rater reliability values are high, with r = 0. 918.

Based on the test data of performance-based assessment instruments net game material for elementary school students on the item process of skills, it shows that the value of reliability if estimated by using Alpha coefficient, shows a high coefficient with r = 0.953. Analysis is done by using Anova-General multifacet Model (ICC). Interclass correlation coefficient value indicates inter-rater reliability values are high, with r = 0. 872.

Based on the test data of performance-based assessment instruments net game material for the elementary school students on the item of product of skills, it is known that the reliability value when estimated using Alpha coefficient, shows a high coefficient with r = 0916. Analysis is done by using Anova-General multifacet Model (ICC). Interclass correlation coefficient value indicates inter-rater reliability values are high, with r = 0. 785.

Based on the test data of performance-based assessment instruments net game material for elementary school students on items attitudes and behavior, it is known that the reliability value when estimated using Alpha coefficient, shows a high coefficient with r = 0918. Analysis is done by using Anova-General multifacet Model (ICC). Interclass correlation coefficient value indicates inter-rater reliability values are high, with r = 0. 885.

DISCUSSION

A player of the net game (table tennis) in his playing always maintains the score and tries to obtain a score for the victory. Maintain the score includes maintaining the space, maintaining the score and return the ball. Obtain a score includes creating the scores and create space to attack, use the space to attack. To generate a score and maintain the score in the game required skills of decision making (decisions), skill execution (implementation process and product skills) and skills base that is returning to the base after receiving and returning the ball (Griffin, 2007: 220 and Lund, 2010: 225). Therefore, in the learning of table tennis, it always teaches and assesses aspects of decision-making, implementation skills include process and product and returning to the base upon receipt and return the ball.

Assessment is an important part of the learning process of the net game (table tennis). This assessment is not simply done by using a test of skill, since the test of their skills is only trying to measure the product and it cannot measure the mental capabilities, process skills and attitudes during play. Therefore, to assess the mental, skills and attitudes simultaneously while playing table tennis, it needs a performance-based assessment instruments for the net game (table tennis) for elementary school students.

According to Nitko (2011: 38) and Anwar (2012: 112) the assessment instrument is said to be good if the instrument has the high reliability and validity. It generates a validation of the net game contents (table tennis) for elementary school students include aspects of decision making (Content validity ratio (CVR) = 0.714, the process (CVR = 1:00, the product (CVR = 1.00) and the attitudes and behavior (CVR = 0714). Since the value of the CVR all aspects is above 0.500 then the instrument or the instruments has represented a right measurement (Cook & Beckman, 2006). Thus, the performance-based assessment instrument is able to measure the results of the net games learning material (table tennis) or the instrument has content validity in the very good level. It is consistent with Nitko and Brookhart (2011: 38) and Anwar (2012: 112) who state that the level of accuracy of a relevant instrument is able to measure what should be measured.

It is generated the reliability of inter-rater net game (table tennis) for elementary school students include aspects of decision making (alpha coefficient, r = 0971), the process of skills (alpha coefficient, r = 0.953), product of skill (alpha coefficient, r = 0916), return to the base (alpha coefficient, r = 0918). Since the results are approaching to 1.00, then this instrument has the reliability in the high level. It is in accordance with the opinion of Anwar, (2012: 112) who states that the reliability coefficient (rxx ') in the range of 0 to 1.00, where the coefficients reliability is getting close to 1.00, it is said in the high level of reliability, and vice versa.

Model performance-based assessment instruments for the net game material (table tennis), covering the aspects of decision making, process of skills, product of skill, and return to the base, are known to have the high value of inter-rater reliability, then the whole aspect of the instrument is expressed consistently. The performance-based assessment instrument is an instrument that has the consistency of the scores given by the assessors, so that when it is used repeatedly, it generates the same data.

CONCLUSION

Based on the results of research and discussion above, it can be concluded as follows: (1) The model of performance-based assessment instruments on the net game material (table tennis) have the high value of content validity; (2) The model of performance-based assessment instruments on the net game material (table tennis) have the high reliability value.

ACKNOWLEDGE

Thanks authors say sincere to all those who have helped smooth this study was. Similarly to the leadership journals have been willing to load a script ini. Semoga everything got abundant rewards from God.

REFERENCES

Azwar, S. (2012). *Reliabilitas dan validitas.* Yogyakarta: Pustaka Pelajar.

Griffin, L. L., Oslin, J., & Mitchell, S. (2007). *Teaching sport concepts and skill. A tactical game approach*. Champaign, II: Human Kineties.

Grehaigne, J. F., Richard J. F., & Griffin, L. L. (2005). *Teaching and learning team sport and games.* London: Routledge.

Husaini Usman, & R. Purnomo Setiady Akbar. (2003). *Pengantar statistik*. Jakarta: Bumi Aksara.

Lund, J. L., & Kirk, M. F. (2010). *Performance-based assessment for Middleand Highschool physical education*. USA: Human Kinetics.

Memmert, D., & Harvey, S. (2008). The game performance assessment instrument (GPAI): Some concerns and solutions for further development. *Journal of Teaching in Physical Education, 27: 220-240*

Mueller, John. (2009). Authentic assessment toolbox. North Central C ollege *http:*//www.noctrl.ed/Naperville*, http://jonathan.mueller.fakulty.noctrl.edu/toolbox/index.htm* (Diambil 27 Agustus 2009).

Nitko, A, J., & Brookhart, S. M. (2011). *Educational assesment of students (6th Ed).* Boston: Pearson Education, Inc.

Suzann Schiemer. (2000). *Assessment strategies elementary physical education*. USA: Human kinetics.

Sugiyono. (2006). *Metode penelitian pendidikan, pendekatan kuantitatif, kualitatif dan R&D*. Bandung: Penerbit Alfabeta.

Sukadji, Soetarlinah. (2000). *Menyusun dan mengevaluasi laporan penelitian*. Jakarta: UI-Press.

Shultz, K. S., & Whitney, D. J. (2005). *Measurement theory in action: Case studies and exercises.* Thousand Oaks, California: Sage Publications, Inc.

Performance-based assessment instrument developed for the net game, particularly table tennis have been tested with high value content validity and reliability. It can be concluded that the performance-based instrument for table tennis is a valid and reliable instrument in measure elementary student acquisition on the table tennis.