Validity and Reliability of Performance Assessment Index (PAI) for Invasion Games in Football

Mohammad Zarid Zaini & Norkhalid Salimin

Validity and Reliability of Performance Assessment Index (PAI) for Invasion Games in Football

Mohammad Zarid Zaini & Norkhalid Salimin
Faculty of Sports Science and Coaching, Universiti Pendidikan Sultan Idris, Malaysia

Abstract
The aim of this study is to determine the validity and reliability of Performance Assessment Index (PAI). This instrument was developed to assess football players under 14 years old. The PAI is a holistic instrument comprise of psychomotor, cognitive, and affective. The findings showed high validity and reliability for PAI for football in invasion games. The final validity analysis shows the agreement between expert panel is 91%. While the reliability analysis using interobserver agreement showed the percentage of agreement between two coaches is 82.22%. In conclusion, the finding of this study found that PAI is useful to assess football players under 14 years old from three aspects which is psychomotor, cognitive, and affective.
Keywords: Football, Invasion Games, Performance Assessment Index (PAI), Validity, Reliability

Introduction

A comprehensive assessment tools will provide a good insight to coaches. Assessment is one of the efforts in learning (Raplh Tyler, 1964). According to Buck et al. (2007) assessment will assess the learning stage or the performance stage, strength and weaknesses in certain learning process and provide information about the effectiveness of a training program. However, the assessment instrument is validated when it's really measures what the needed (Ahmad, 2004). This study aim to determine the validity and reliability of PAI in assessing football players under 14 years old.

The PAI instrument was given to the experts to determine the validity. According to Pallant (2007) validity refers to how the instrument assessing the study items. Validity was run to ensure the indicators that been use is measuring the actual purpose of the research (Sabitha, 2006). Wee (2009) says that the assessment aspect in physical education should consists of psychomotor, cognitive, and affective. Therefore, to assess full aspect of assessment it must consist of psychomotor, cognitive, and affective. These three domains will affect individual in completed their tasks (Ramasamy, 2000).

Cognitive domain refers to the individual way of thinking and individual intellectual. The cognitive domain is to assess the knowledge level, understanding and intelligence of an individual (Hajar, 2004). The cognitive domain is an important aspect to help athlete to

324

perform in sports (Regnier et al., 1993; Brown, 2001; Abbot & Collins, 2002, 2004; Abbot et al., 2005). Cognitive also about perception, memory and remembering (Pithers, 2000).

Psychomotor domain includes the physical movement, coordination, and motor skills. Psychomotor is about the skills and performance of muscle movement (Rupani & Buttho, 2011). Affective domain is referring to the changes of feeling, attitude and value that contribute to the thinking and behaviour (Allen & Friedman, 2010).

Next, to assess the reliability of PAI, two coaches have involved in the interobserver agreement analysis. Medina and Noguera (1999) suggested it is essential to include more than one rater to ensure data consistency are not affected by the subjective judgment if one rater is involved. Lange (2011) and Rink (2014) agreed that interobserver agreement refers to the agreement between two or more raters. Salimin et al (2014) concluded that at least 70 percent agreement is acceptable for reliability.

In conclusion, the combination of psychomotor, cognitive, and affective domain will produce a comprehensive and holistic instrument to assess football players. The suitable, right, and holistic instrument may help to provide meaningful information to coaches from time to time. The right instrument able to determine the player's strength and weakness. This information will help coach to plan the next steps (training) to improve player's performances.

The Objective of the Study
The objective of this study is to identify the validity and reliability of PAI instrument. Below is the objectives of this study:
  i) To identify the validity of PAI instrument towards the performance assessment of football players under 14 years old.
  ii) To identify the reliability of PAI instrument toward the performance assessment of football players under 14 years old.

Methodology
The PAI instrument is use by the coaches to assess the football player during the training. This study involves two coaches and 30 football players under 14 years old from one of Sekolah Menengah in Johor Bahru. Besides, there are five expert panels involved to identify the validity of the instrument. The PAI instrument will use the observation from video recording to assess the performance of football players.

Development of Footballer Assessment Instrument (PAI)
The PAI instrument was developed with three main domains which is psychomotor, cognitive, and affective. Donabedian Model (1985) in Figure 1 was used and it's consisted of three main items which is input, process and output. Besides, Model Morrow et al (2005) was used in constructing the PAI as per Figure 2.

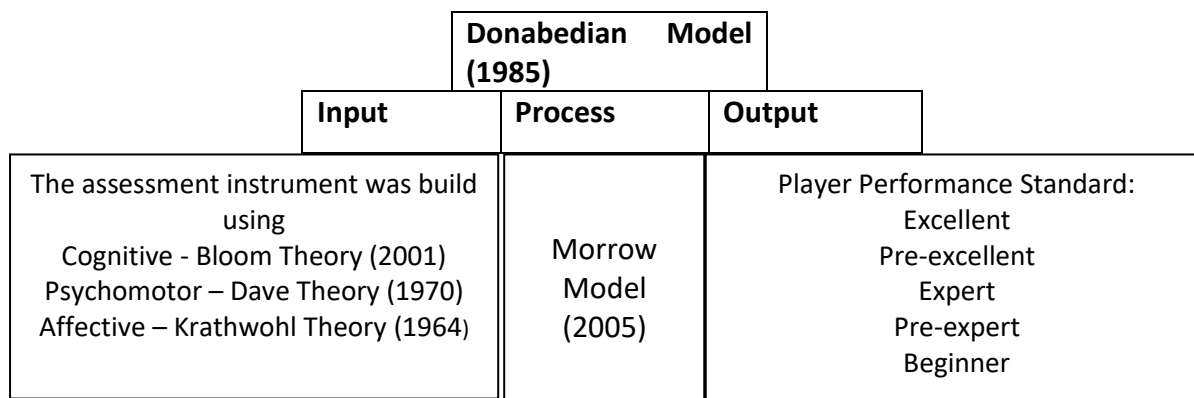| Donabedian Model (1985)		
Input	**Process**	**Output**
The assessment instrument was build using Cognitive - Bloom Theory (2001) Psychomotor – Dave Theory (1970) Affective – Krathwohl Theory (1964)	Morrow Model (2005)	Player Performance Standard: Excellent Pre-excellent Expert Pre-expert Beginner

Figure 1. The development of PAI based on Donabedian Model (1985)

Figure 1 show the development processes of PAI instrument to assess performance of football players. The process is based on Donabedian Model which had been introduced in year 1985. This model consists of three main aspects which is input, process and output. (Donabedian, 1985).

The first step of development the PAI instrument, researcher set the objectives to be achieved in this model. In this study, the input is the assessment of football players. The assessment is normally depending on the research objective. Focus on this study is to assess football players from three main aspects which is psychomotor, cognitive, and affective. Several numbers of theory were referred to build the instrument. The assessment items were developing according three theories including cognitive aspect was referred to Bloom (2001), psychomotor aspect referred to Dave (1970) and affective aspect referred to Krawthwohl et al (1964).

Second step in developing the PAI instrument is by referring to Morrow Model (2005). There are eight steps to develop the instrument as per figure 2. The process (Morrow Model) covers the initial stage of development of items in the assessment instrument until the process of getting the validity and reliability.

The next step is the output of this study which is the PAI instrument. The football players will be categories into five level which is excellent, pre-excellent, expert, pre-expert, and beginner.
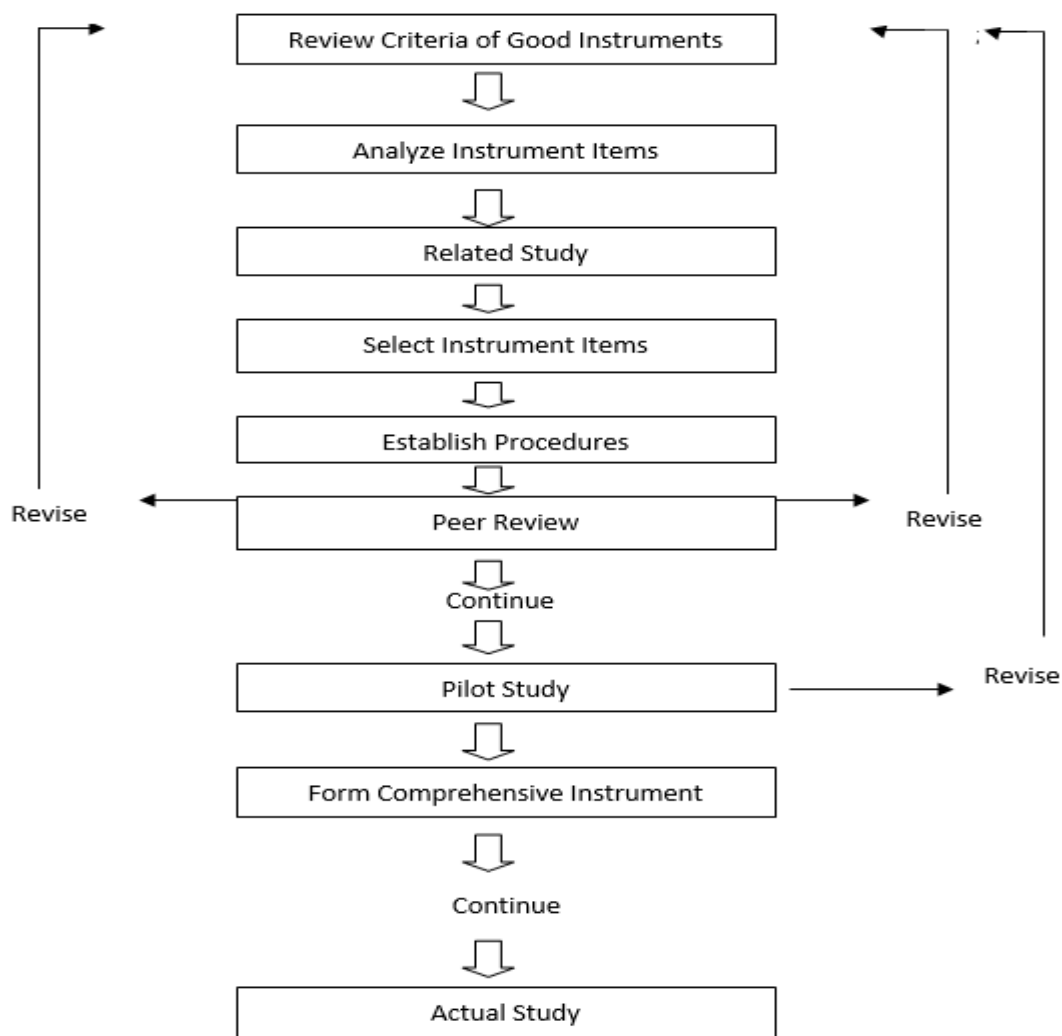
Figure 2. The flow chart of Developing Learning Assessment Model (Morrow et al., 2005)

There are several steps to be complied in the study to ensure that the instrument can achieve the best quality to measure football player's performance. The steps are as follows:

Step One
The first step is to choose the right and suitable criteria of good instrument. The PAI instrument is the assessment of football players based on the best evaluation criteria in learning which is the three domain learning which is psychomotor, cognitive and affective.

Step Two
The second process is to design a right or suitable criteria and rubrics for PAI Instrument. The criteria is divided into three categories which is Part A for Psychomotor, Part B for Cognitive and Part C for Affective. There are five items in Part A, 2 items in Part B and 2 items in Part C.

Step Three
The third process is to study the previous related literatures. The PAI instrument was developed based on the three domain of learning which is Cognitive domain refers to Bloom

et al. (1956), Psychomotor domain refers to Dave (1970) and Affective domain refers to Krathwohl et al. (1964).

Step Four
The fourth process is to choose the most suitable items for the instrument to assess the performance of football players. In this case, the items were chosen based on the three domains of learning.

Step Five
The fifth process is to set the procedures of assessment the PAI instrument. Coaches need to evaluate football players based on the rubrics. The rubrics for PAI instrument is using the 5 point likert scale. The value of 1 represent very poor performance, 2 represent poor performance, 3 represent medium performance, 4 represent good performance and 5 represent very good performance.

Step Six
The next process is to refer the criteria and rubrics to five experts' panels. The suggestions and feedbacks received from the expert panels are collected for improvising of the PAI Instrument.

Step Seven
Next, the pilot study will be done to the selected football players under 14 years old (N = 30). The purpose of pilot study is to obtain the reliability of PAI Instrument. There are few changes has been made based on feedback during the pilot study. Changes was made n order to best suit the feedback from the pilot study.

Last Step
After the validity and reliability of PAI Instrument is obtained, it can be used as a standard instrument to assess the football player's performance.

Analyses and Results
Identify the Validity of PAI
The validity concept was introducing by Kelley (1927) whereby the question of validity, aimed at investigating "whether a test really measures what it purports to measure" (Kelley, 1927, p. 14). Ahmad (2004) also concluded that validity is when the instrument really measuring what its intended to measure. According to Russel (1974) a good instrument must consider the background and behaviour, learning situation and the implementation is good and satisfied, sufficient time to do the assessment and there is a positive change for the subject involved. Sidek and Jamaludin (2005); Rink (2014); Tuckman and Waheed (1981) agreed that the validity of instrument is classified as achieve the high level if r = 0.70.

In this study, the validity of the instrument has been done for two rounds. For the first-round validity the instrument was sent to the expert panels for review and the result as per Table 1. It's involved five expert panels which is content expert, field expert, football coach with B license, football coach from school level and language expert.

The content of the instrument is submitted to the expert panels for review and evaluation. The questionnaire of content validity is in semantic-scale form with 5 points. Based on the feedback from expert panels, all data is calculated using the formula as follow:

$$\frac{\text{Total Expert Score (x)}}{\text{Maximum Score}} \quad X \quad 100\% = \text{Content Validity Achievement}$$

Table 1. Content Validity by Expert Panels of Pilot Study (First Time)

Item/Expert	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Σ	M
1	4	3	4	4	4		
2	5	3	5	5	4		
3	5	4	4	5	4		
4	4	4	4	4	4		
5	5	4	5	5	5		
6	5	3	4	4	4		
7	5	3	5	4	5		
8	5	4	5	4	4		
9	4	4	4	4	5		
10	5	3	4	4	5		
11	4	4	4	5	4		
12	4	4	4	4	5		
13	4	3	4	4	4		
14	5	3	4	4	4		
Σ	64	49	60	60	61	294	0.84
M	0.91	0.70	0.86	0.86	0.87	4.20	0.84

Table 1 show the validity of pilot study is r = .84 (n = 5). According to Tuckman and Waheed (1981), Rink (2014) and Sidek & Jamaluddin (2005), the value r = .70 is sufficient to proof that the instrument achieves a high-level validity.

Next, the second round of validity was run as refer to Mohd Izwan et al. (2005) to enhance the validity the second process is advisable. The instrument then sent to the expert panels for second review and evaluation. A few items have been reviewed and updated according to the feedback from expert panels.

Table 2. Content Validity by Expert Panels of Pilot Study (Second Time)

Item/Expert	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Σ	M
1	4	4	4	5	4		
2	5	4	5	4	4		
3	5	5	4	5	5		
4	4	5	5	4	4		
5	5	5	5	5	5		
6	5	4	5	4	4		
7	5	4	5	4	5		
8	5	5	5	5	5		
9	4	4	4	5	5		
10	5	4	4	5	4		
11	4	5	4	4	5		
12	4	5	5	5	5		
13	4	4	5	4	4		
14	5	5	5	4	5		
Σ	64	63	65	63	64	319	0.91
M	0.91	0.90	0.93	0.90	0.91	4.56	0.91

Based on Table 2, the validity for the second round validation is r = .91 (n = 5). All views, comments and reviews from the expert panels were taken into account for improvising.

Identify the Reliability of PAI

One of the reliability observational tools is interobserver agreement. Interobserver agreement is suitable when there is more than one observer. Reliability is usually determined by comparing the observations of one observer with another observer considered to be an expert. However, according to Rink (2014), in the absent of expert of when no expert is available, reliability is usually determined by agreement between more than one observer on the same event. Therefore, this research conducted an interobserver agreement between two football coaches. The coaches were supplied with 45 videos consist of 3 vs 3 small sided games. The coaches than evaluate the performance in the video based on the rubric scale provided.

The coaches perform the evaluation using PAI form and then analyze using the percentage of agreement method follow Rink (2014) procedures. Formula to compute simple percentage agreement:

$$\text{Percentage of agreement} = \frac{\text{Numbers of agreements}}{\text{Number of agreements + Number of disagreements}} \times 100$$

The data of evaluation from the two coaches showed the percentage of agreements is between 60% to 100% with M = 82.22%. According to Saad et al (2016) and Rink (2014), the acceptable reliability value should be at least 70% agreement between the observers.

Therefore, PAI is fulfilled the condition and can be use as the instrument to assess football players performance.

Conclusion

In conclusion, PAI is suitable to be as an assessment tool to assess football player's performance especially for under 14 years old. The PAI able to evaluate football players from three learning domains which is psychomotor, cognitive, and affective. Moreover, the assessment will help to determine the football player's performance status as reference to the coaches. Additionally, PAI can be used as a guidance for coaches to improve the quality of teaching or coaching football players in future.

Reference

Ahmad, S. M. N. D. J. (2005). *Pembinaan Modul: Bagaimana Membina Modul Latihan dan Modul Akademik*. Serdang: Penerbit Universiti Putra Malaysia.

Allen, K. N., & Friedman, B. D. (2010). Affective Learning: A Taxonomy for Teaching Social Work Values. *Social Work Values and Ethics*, 7(2).

Donabedian, A. (1985). The epidemiology of quality. *Inquiry*. 22:282–292.

Kelley, T. L. (1927). *Interpretation of educational measurements*. New York: World Book Company

Krathwohl, D. R., Bloom, B.S., Masia, D. B. B. (1964). *Taxonomy of educational objectives the classification of educational goals, handbook II: Affective domain.* New York: David McKay Company, Inc.

Lange, R. T. (2011) Inter-rater Reliability. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) *Encyclopedia of Clinical Neuropsychology.* Springer, New York, NY.

Medina, J., & Noguera, M. A. D. (1999). Observer training methodology for research on Physical Education and Sport in which observation is used as a method. *European Journal of Human Movement,* (5), 69-86.

Izwan, M. S., Norkhalid, S., Ginathevan, D. E. (2015). The Validity and Reliability of ISO Test towards the Performance Assessment of Future Physical Education Teachers in Teaching and Learning Process. *Procedia Social and Behavioral Sciences*. Vol 195, pp. 814-820.

Pallant, J. (2007). *SPSS Survival Manual*. Sydney: Ligare Book Printer.

Pithers, R. T. (2000). Cognitive Learning Style: A Review of The Field Dependent-Field Independent Approach. *Journal of Vocational Education and Training*, 54(1), 117-129.

Ramasamy, J. (2000). Model Kompetensi: Wajarkah Sektor Awam Meniru Sektor Swasta? http://www.jpa.gov.my/buletinjpa/J2bil2?/KOMPETENSI.htm

Rink, J. E. (2014). *Teaching physical education for learning* (7th ed.). Boston: McGraw Hill.

Rupani, C. M., & Bhutto, M. I. (2011). Evaluation of Existing Teaching Learning Process on Bloom's Taxonomy. *International Journal of Academic Research in Business and Social Sciences,* 1, 119-128.

Salimin, N., Jani, J., Ishak, N., A., Hasan, S., Elumalai, G., Ong, K. B., Shahril, M. I. (2014). Validity and Reliability of Comprehensive Assessment Instrument for Professional Skills in Field and Courts Sport Among Major Students Coaching Based on Model K. *International Journal of Humanities, Arts, Medicine and Sciences,* 2, 37-44.

Tuckman, B. W., & Waheed, M. A. (1981). Evaluation an individualized science Programme for community college students. *Journal of research in science teaching*. 18:489-495.