Jurnal Sains Sukan dan Pendidikan Jasmani Vol 10, No 2, 2021 (8-18)
ISSN: 2232-1918 / eISSN: 2600-9323
https://ejournal.upsi.edu.my/journal/JSSPJ

THE SELECTED PHYSICAL FITNESS FIELD TESTS FOR YOUTH RUGBY PLAYER: VALIDITY AND RELIABILITY

Zarizi Ab Rahman[1*], Azlan Ahmad Kamal[1], Ahmad Hashim[2] & Mohad Anizu Mohd Nor[3]

[1]Faculty of Education, Universiti Teknologi MARA, Kampus Puncak Alam, Puncak Alam, Selangor, Malaysia
[2]Faculty of Sport Science and Coaching, Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak, Malaysia.
[3]Faculty of Sport Science and Recreation, Academic Building 3, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Published online: 30 August 2021

To cite this article (APA): Ab Rahman, Z., Ahmad Kamal, A., Hashim, A., & Mohd Nor, M. A. (2021). The Selected Physical Fitness Field Tests for Youth Rugby Player: Validity and Reliability. *Jurnal Sains Sukan & Pendidikan Jasmani*, *10*(2), 8-18. https://doi.org/10.37134/jsspj.vol10.2.2.2021

To link to this article: https://doi.org/10.37134/jsspj.vol10.2.2.2021

Abstract

The purpose of the present study is to investigate the validity and reliability of the test instruments for selected physical fitness components (agility, power, and speed) among youth rugby players in Selangor, Malaysia. Even though previous researchers have conducted the validity and reliability for the fitness test, investigating the reliability and validity is ongoing and needs to be established according to specific populations to ensure valid, reliable interpretation or precise measurement. Sixty youth rugby players aged 13 and 14 years old underwent measurements of agility (4x10 meters run), power (standing long jump), speed (30 meters sprint test). In addition, two testers independently assessed all the tests towards subjects using standardized skill criteria. The descriptive statistics of the variables were; i) 4x10 meters test for elite players (M = 11.9, SD = .81) and non-elite players (M = 12.9, SD = .83), ii) 30 meters sprint test for elite players (M = 5.56, SD = .31) and non-elite players (M = 5.90, SD = .33), iii) standing long jump for elite players (M = 211.8, SD = 8.20) and non-elite players (M = 206.53, SD = 6.89). Significant ($p < 0.05$) differences were detected for all the 3 tests. The study found that the test-retest and Pearson correlation of reliability for the 3 test instruments were high (r = .80) for 4x10 meters test, 30 meters sprint test (r = .82), and standing long jump (r = .72). These findings suggested that all the 3 test instruments are reliable and valid for the studied population. This study will provide comprehensive evidence for all practitioners in terms of appropriate and quality measurement to ensure the training program and talent selection are meaningful. These findings also provide direction for future researchers to explore the different research methods, populations, and other fitness components.

Keywords: Reliability, validity, skill related physical fitness components, youth rugby player.

INTRODUCTION

Skill-related physical fitness components such as speed, power, agility, balance, coordination, and reaction time are essential in sports as a fundamental basis of performance. Each athlete needs to have a good condition of skill-related fitness to ensure they can perform better in their sport. The requirement of the skill-related physical fitness components may be different according to the demand of the sports. Rugby is a field-based contact team sport that requires both high and low-intensity activity. This game consists of 2 periods of 20 minutes (under 13), 30 minutes (under 16), 35 minutes (under 18), and 40 minutes (seniors) (Oliver & Lloyd, 2012, Bailey, Collins, Ford, MacNamara, Toms & Pearce, 2010; Quarrie, & Hopkins, 2008; Malina, Cumming, Kontos, Eisenmann, Ribeiro & Aroso, 2005). Typically rugby player spends most of the time with low-intensity activity despite high-intensity activities such as combinations of sprinting, physical body contact, and high force generation from the lower and upper body place utmost demands on the anaerobic system (O'Connor, 1996; Brewer & Davis, 1995). Sprinting, tackling, rucking, scrummaging, mauling are among the high-intensity activities while the lower intensities such as standing, walking, jogging (Duthie, Pyne, Marsh & Hooper, 2006). Smart, Hopkins, Quarrie, and Gill (2014) explained that strength, power, speed, agility, aerobic and anaerobic conditioning, and body composition among the physical attributes of a rugby player.

In line with the needs of the sport, rugby players need to possess and master all the skills in the game and have a high level of physical fitness to ensure they can maintain their performance and compete during training or competition. Duthie, Pyne & Hooper (2003) also suggested that it is crucial for a rugby player to develop and maintain physical fitness from an early age because of the physical demands and reduce injury and prolong their playing careers. This requirement is also essential to ensure rugby development at the grass-roots level and remain until for the senior team. Therefore, rugby practitioners, coaches, and teachers can ensure sustainable participation and player development. Furthermore, Till, Weakley, Read, Phibbs, P., Darrall-Jones, Roe, Chantler, Mellalieu, Hislop, Stokes, Rock and Jones (2020) also stressed the importance of fitness; for instance, speed, change of direction speed, high intensity running ability, muscular strength, and power are essential for player development alongside injury prevention and should be strongly considered when developing any training program for grass root. Consistently to Henriques-Neto, Minderico, Peralta, Marques & Sardinha (2020), monitoring fitness in young athletes is vital to improve physical performance, identify new talents and develop injury prevention programs.

The evaluation of young athletes in rugby is crucial for training prescription, long and short-term goal setting, including talent development. Besides, fitness testing in rugby is conducted to ensure that fitness performance is at the required level for competition purposes. In terms of fitness achievement, rugby players should be asked to perform periodically to generate data. It can be an effective procedure, including revealing a detailed and appropriate evaluation of the athlete's physical abilities, health, strengths, and weaknesses. Furthermore, all the information related to the validity and reliability of the test will help the coaches and teachers to monitor any changes for athletes, either performance, talent development, injury prevention, and so on. With this in mind, the evaluation of components of physical fitness for rugby players is crucial to assure all the practitioners can use all the data for significant interpretation in terms of player performance, injury prevention, increase motivation, and so forth. Inappropriate and poor measurement will fail any planning to improve performance, reduce injury, and prepare for talent development. When involving the quantifying physical fitness, the quality used of measurement or test battery used in collecting data is important to make sure the data can be used for any purposes of interpretation for the player, including physical and skill performance, quantify player responsiveness to training interventions, talent identification, and so forth. Accordingly, it is not surprising that a strong interest exists to establish the validity and reliability of the test battery to prepare and develop the player.

The quality of the test battery is relying on two crucial components in measurements like validity and reliability. Regardless of whether a test is used for fitness assessment, exercise prescription, or for estimating the effectiveness of an exercise program, it is essential to determine its reliability (Currell & Jeukendrup, 2008) as well as validity because these two terms are independent of each other, a measurement may be valid but not reliable, or reliable but not valid. Miller (2013) explained that it must have a high degree of reliability for a test to have a high degree of validity. If the test score is not reliable, the validity of a test is limited. Moreover, a test may be consistent, but it may not measure what it claims to measure (Miller, 2013). A few testing methods or test batteries are used in rugby (Haycraft, Kovalchik, Pyne & Robertson, 2017) to measure physical fitness. Nevertheless, all the test batteries need to be valid and reliable in a particular population. Baumgartner, Jackson, Mahar & Rowe (2006) reported a test that yields scores for beginners and may not do so for advanced performers. A valid interpretation and information only for individuals similar in age, gender, and experience to those on whom the initial validation was conducted (Miller 2013; Baumgartner et al., 2006). Even though previous researchers have conducted the validity and reliability for the physical fitness test, investigating the reliability and validity is ongoing and needs to be established according to specific populations to ensure valid, reliable interpretation or precise measurement. Specific characteristics are essential when administering the test and measurement. Baumgartner, Jackson, Mahar & Rowe (2006) also suggested, it is a good idea to determine validity because the populations cannot be exactly like those initially used in the validation study.

In addition, Baumgartner et al., (2006) also asserted, a test must have reliability first before validity. However, to have reliability, a test must have objectivity. Hence, three components that are important during measurement are reliability, validity, and objectivity. Validity exists if the interpretation of the test scores is accurate (Baumgartner, Strong & Hensley, 2005). The essential characteristic of measurement is validity, the second most important quality is reliability, and another vital characteristic is objectivity. The objectivity, reliability, and validity of the test batteries should be determined before administering any measurement. The objectivity, reliability, and validity must be acceptable for a test to be considered a good test. Without them, the data is useless (Thomas & Nelson, 2001; Baumgartner et al., 2006; Miller, 2013). Thus, this study attempted to determine the validity and reliability of selected skill-related physical fitness field tests for a youth rugby player.

LITERATURE REVIEW

An instrument or test batteries are any physical performance test, mechanical, electronic equipment, paper-pencil test used to collect information or data on the variable under study. The choice of the instrument to be used in the data collection process involves deciding whether it will have already been developed and will be used as it is one that already exists but will be revised, or one that will be new and needs to be developed (Baumgartner, Strong & Hensley, 2005). Researchers also need to consider the validity and reliability of the test battery. Validity of measurement indicates the degree to which the test or instrument measures what it is supposed to measure (Thomas & Nelson, 2001). Nevertheless, Miller (2013) stressed that validity is about the agreement between the test measures and the performance, skill, or behavior the test is designed to measure. Hence, it can be concluded that validity is specific to a particular use and group and how well the test battery should provide valid evidence for interpretation regarding the measurement of the construct or trait. Whereas Baumgartner et al., (2006) stated that validity refers to the appropriateness of inferences made from specific measures, validation is the process of obtaining evidence to support these inferences (Baumgartner et al., 2006). There are three basic types of validity evidence, namely content validity, criterion validity, and construct validity (Miller, 2013; Baumgartner et al., 2006). Validity can be estimated either logically or statically. Exercise science work predominantly with a physical performance test. Thus, validity evidence for exercise science depends more on the statistical approach.

An integral part of validity is reliability, which pertains to the consistency, or repeatability, of a measure (Thomas and Nelson 2001). Typically, reliability refers to the consistency of the test (Miller, 2013). A reliable test should obtain approximately the same results regardless of the number of times it is given. A test given to a group of individuals on one day should yield the same results if given on another day to the same group. Even though there are no identical scores due to measurement error, the order of the scores will be around the same if the test has reliability. Mainly, reliability is a question of data quality, whereas validity is a question of inferential quality (Zumbo & Rupp, 2004). Reliability and validity theory are interconnected; a test cannot be considered valid if it is not reliable. For instance, if the test is not consistent, then the test cannot be trusted. In line with Thomas and Nelson (2001), it can be concluded that test batteries should be reliable first before a test can be valid. In other words, for a test to have a high degree of validity, it must have a high degree of reliability.

Baumgartner et al., (2006) stated that individuals who conduct research on physical activity must choose a measure with evidence of validity specific to the purpose of the measurement and the particular population under investigation. Hence, researchers should decide on a test battery-related specifically to the particular purpose of the study and the particular type of respondents with whom the measure will be used. Mahar & Rowe (2016) explained some parts of validity theory in psychology, and educational measurement does not seem fit to the different types of research in exercise science. Therefore, a strong program and method of construct validation that will fit a wide variety of constructs and contexts, especially those relating to the study of physical activity, need to be established.

Baumgartner et al., (2006), construct validity evidence for exercise science, and physical education can be determined based on judgment by expertise in related area of the variables, comparison of the performance of the group before and after instruction or training, and statistical procedure namely factor analysis to identify constructs and the test that yield score leading to valid interpretation.

Miller (2013) reported that comparing the mean difference for elite and non-elite performers also one of the procedures to determine to construct validity. That method is similar to a known group difference method referring to a test that can discriminate between two groups known to differ on the variable interest (Davidson, 2014; Thomas & Nelson, 2001). Thomas and Nelson (2001) explained this method used for establishing construct validity in which the test scores of groups that should differ on a trait or ability to compare. Construct validity evidence can be demonstrated when there are significant mean differences between elite and non-elite groups (Miller, 2013). Test instrument that has construct validity, the elite player will most likely score higher than the non-elite players. Previous researcher has used this method to determine the construct validity evidence. Hachana, Chaabene, Rajeb, Khlifa, Aouadi, Chamari, & Gabbett (2014) also compared performance between elite and sub-elite groups in soccer to determine the validity of the agility test. The comparison between two groups of samples also used a study to determine validity evidence of cognitive instrument performance for badminton players (van de Water, Huijgen, Faber & Elferink-Gemser, 2017).

Baumgartner, Jackson, Mahar & Rowe (2006) suggested that to have validity; a test must have reliability. However, to have reliability, a test must first have objectivity. According to Ahmad Hashim (2015), objectivity depends on the credibility of the testers. Bishop (2008) suggest that the results between testers should be consistent. There are a few methods to determine the objectivity in a study, such as the intraclass correlation coefficient using the scores among the testers and Pearson Correlation Hashim, Ariffin, Hashim and Yusof (2018) used intraclass correlation coefficient using the scores of the two testers to estimate the objectivity between two testers in 90º Push-Ups Test Protocol. This correlation coefficient suggests high objectivity between testers. In this study, data will be correlated using Pearson Correlation between testers one and testers 2. Baumgartner et al., (2006) defined reliability as the consistency of test scores; a test has objectivity if the scores are not dependent on who administered the test. At least two scores for each person being tested must be gathered to provide evidence of reliability or objectivity. These two scores can be collected from two different scorers on two different trials in one day or two different days (Baumgartner et al., 2006). When two score sets are measured on a continuous scale, and one is the interested inconsistency of measurement, the Pearson's r is calculated to estimate reliability, otherwise termed an interclass reliability coefficient (Odom & Morrow, 2006). Examples of reliability coefficients that are based on Pearson's r include the stability coefficient, equivalence coefficient, and the objectivity (both interrater and intrarater) coefficient. A few methods can be used to calculate the reliability coefficient, such as Cronbach's alpha, Pearson's correlation, the Spearman-Brown formula, and Cohen's Kappa. Has used Pearson's correlation to estimate the reliability of Gottman Questionnaires of "Couple Trust Measurement." In Kooiman, Krijnen, Van der Schans, Cees and De Groot (2016), test-retest reliability was determined by calculating the intraclass correlations (ICC) between Session 1 and Session 2 for reliability of Ten Consumer Activity Trackers Depend on Walking speed, and paired-sample t-tests and Wilcoxon signed-rank tests investigated significant mean differences.

Rugby is a team sport that involves high-intensity activity such as sprinting, tackling, collisions, and low-intensity activity like walking and jogging. These various physiological demands, rugby players require muscular strength and power, speed, agility, in addition to well-developed aerobic and anaerobic capacity (Argus, Gill, Keogh, Hopkins & Beaven, 2010; Gabbett, Kelly & Sheppard, 2008; Granados, Izquierdo, Ibanez, Ruesta & Gorostiaga, 2008). There a several test instruments to test skill-related fitness among rugby players. Gabbett et al. (2008) used 5 meters, 10 meters, and 20 meters sprint tests to evaluate the speed of rugby players. A youth rugby player needs to have a good speed performance since rugby is a game that needs speeds to attack and defense. Moreover, Agar-Newman and Klimstra (2015) has used the standing long jump to test the explosive of leg power for female rugby players in Canada. Meir, Holding and Hetherington (2014) suggested L Run test evaluate the agility of rugby players while Jarvis, Sullivan, Davies, Wiltshire and Baker (2009) used T-Test and Illinois Test as a test instrument for agility. Agility helps performance in activities that require the player to change direction quickly while keeping balance, strength, speed, and body control. Wang, Hoffman, Tanigawa, Miramonti, Monica and Stout (2016) assumed that the relationship between strength, speed, and agility does seem to be synergistic. Hence, a youth rugby player will have good agility when they have good strength and speed. Even though there are a few test batteries to measure skill-related components for a rugby player, all the practitioners need to ensure the validity and reliability of all the test batteries. Data collection is influenced by random errors of measurement, systematic error, or random error. Measurement error may come from any sources such as the participants, the scoring, the testing, and the instruments. Besides, the practitioners should consider the data's meaningfulness, which comes from the validity of the test instrument. All these components will reduce the measurement error as well as the quality of data.

METHODOLGY

Participants and Procedures

Data were collected from youth rugby player (n=60) consisted of elite player (n = 30) and non-elite player (n = 30) age 13 to 14 in Selangor. All the 60 subjects were measured six components of skill-related fitness by using 4 x 10-meter shuttle run (agility), 30-meter sprint (speed), and standing long jump (explosive leg power). Objectivity for the test instruments was estimated from the score between two testers. Two testers (tester 1, n=30, tester 2, n= 30) were administered the same test instruments with the standard procedure to the sample in this study. Test-retest was used to determine the reliability of the instruments. The time interval for 48 hours' time interval for 4 x 10-meter shuttle run, standing long jump, 30-meter sprint. In terms of validity purposes, all the test instruments were administered to samples on a different day. Samples were divided into the elite and non-elite groups. Each group was measured the skill-related fitness using the same test instruments.

Data Collection

During the initial meeting with the participants, all the participants signed an informed consent form, were told the purpose of the study, and were familiarized with tests they would be performing. The participants have tested for a total of 4 days. All the participants will perform the warmup and follow by the measurement for each instrument. On the first day, all participants have performed all the test instruments. On the third day, all the subjects were measured standing long jump, 4 x 10 meters shuttle run, and 30-meter sprint. One-hour rest will be given to respondents before another test. On the fourth day, 60 subjects were divided by the elite player and non-player. All the subjects were measured by two testers for all the test instruments for the validity evidence purpose. Below are the test protocols for each instrument.

4x10 meters shuttle run: Mark two lines 10 meters apart using marking tape or cones. The two blocks are placed on the line opposite the line they are going to start at. On the signal "ready," the participant places their front foot behind the starting line. On the signal, "go!" the participant sprints to the opposite line, picks up a block of wood, runs back, and places it on or beyond the starting line. Then turning without rest, they run back to retrieve the second block and carry it back across the finish line. Two trials are performed.

30-meter sprint: The test involves running a single maximum sprint over 30 meters, with the time recorded. A thorough warm-up was performed by the samples, including some practice starts and accelerations. Start from a stationary position, with one foot in front of the other. The front foot must be on or behind the starting line. This starting position should be held for two seconds prior to starting, and no rocking movements are allowed. The tester should provide hints for maximizing speed, such as keeping low, driving hard with the arms and legs, and encourage them to continue running hard through the finish line.

Standing Long Jump: The participant stood behind the starting line, with parallel feet, and pushed off vigorously, and jumped forward as far as they can. The distance is measured from the take-offline to the point where the back of the heel nearest to the take-offline lands on the mat or non-slippery floor. The test was repeated twice, and the best score was retained (in cm).

Data Analysis
Objectivity for the test instruments skill-related fitness was estimated using an inter-class correlation coefficient (R) based on the Pearson Product Moment model presented by Baumgartner et al., (2006). Reliability for the test instruments skills related fitness was estimated by correlating the test-retest data using the Pearson Product Moment model suggested by Baumgartner et al., (2006). For construct validity, an Independent Sample t-Test was calculated to determine the mean difference between the elite and non-elite players. The IBM SPSS version 27.0 package of statistical computer programs was used to do all the calculations and analysis.

RESULTS AND DISCUSSION

A total of 60 participants consisted of youth rugby players included in this study, were tested with all the test instruments for skill-related fitness. All participants were used in the objectivity, reliability, and construct validity of the test instruments' skill-related fitness in this study. Descriptive information for the scores on the test instruments skill related fitness are presented in Table 1. Results indicated that the value of mean for objectivity of 4x10 meters tester 1 (M = 11.95, SD = .812), tester 2 (12.01, SD = .733), 30 meters Sprint test tester 1 (M = 5.56, SD = .318), tester 2 (M = 5.59, SD = .367), Standing long jumper 1 (M = 213.56, SD = .318) tester 2 (M = 212.59, SD = .367).

Table 1. Descriptive Statistics for The Skill Related Fitness Measurement for Objectivity.

Variables		N	M	SD
4x10 meters(agility)	Tester 1	30	11.95	.812
	Tester 2	30	12.01	.733
30-meter sprint (speed)	Tester 1	30	05.56	.318
	Tester 2	30	05.59	.367
Standing long jump (explosive leg power)	Tester 1	30	213.56	.318
	Tester 2	30	212.59	.367

Table 2 showed descriptive statistics for the skill related fitness measurement for reliability. Results indicated that the value of mean for 4x10 meters Test (M = 11.95, SD = .812), retest (12.16, SD = .825), 30 meters Sprint test test (M = 5.56, SD = .318), retest (M = 5.60, SD = .392), Standing long jump 1 (M = 212.16, SD = 8.32) retest (M = 210.46, SD = 10.23).

Table 2. Descriptive Statistics for The Skill Related Fitness Measurement for Reliability

Variables		N	M	SD
4x10 meters(agility)	Test	30	11.95	.812
	Retest	30	12.16	.825
30-meter sprint (speed)	Test	30	5.56	.318
	Retest	30	5.60	.392
Standing long jump (explosive leg power)	Test	30	212.16	8.32
	Retest	30	210.46	10.23

Table 3 showed descriptive statistics for the skill related fitness measurement for construct validity**.** Results indicated that the value of mean for 4x10 meters elite athlete (M = 11.95, SD = .812), non-elite athlete (12.44, SD = .567), 30 meters Sprint test elite athlete (M = 5.56, SD = .318), non-elite athlete (M = 5.90, SD = .333), Standing long jump elite athlete (M = 212.16, SD = 8.32) non-elite athlete (M = 206.53, SD = 6.89).

Table 3. Descriptive Statistics for The Skill Related Fitness Measurement For Construct Validity

Variable		N	M	SD
4x10 meters(agility)	Elite	30	11.95	.812
	Non-Elite	30	12.44	.567
30-meter sprint (speed)	Elite	30	5.56	.318
	Non-Elite	30	5.90	.333
Standing long jump (explosive leg power)	Elite	30	212.16	8.32
	Non-Elite	30	206.53	6.89

Table 4 showed the objectivity and reliability value of 4x10 meters Test, 30 meters Sprint test, and Standing long jump. Findings indicate that 4x10 meters test has high objectivity r = .87, reliability r = .80, 30 meters Sprint test is r = 88 for objectivity, r = .82, for reliability and Standing long jump is r = .88 for objectivity, r = .72 for reliability.

Table 4. Pearson Correlation Coefficients Calculated as Objectivity and Reliability

Variables	Evidence Use	r	N
4x10 meters (agility)	Objectivity	.87	30
	Reliability	.80	30
30 meters sprint (speed)	Objectivity	.88	30
	Reliability	.82	30
Standing long jump (explosive leg power)	Objectivity	.88	30
	Reliability	.72	30

Tables 5, 6, 7, and 8 showed descriptive analysis and independent t test to determine the construct validity of skill related test instrument for 4x10 meters test, 30 meters sprint test, and standing long jump among youth rugby player.

Table 5. Descriptive Analysis of Construct Validity for skill related test instrument among youth rugby player

Variables	Evidence Use	Group	M	SD	N
4x10 meters (agility)	Construct Validity	Elite	11.9	.81	30
		Non-Elite	12.9	.83	30
30-meter Sprint test (speed)	Construct Validity	Elite	5.56	.31	30
		Non-Elite	5.90	.33	30
Standing long jump (explosive leg power)	Construct Validity	Elite	211.8	8.20	30
		Non-Elite	206.53	6.89	30

Independent Sample T Test analysis for 4x10 meters Test was used to compare mean score between elite and non-elite rugby player. Result showed t (58) = -4.634, p = .000 was significant. Result also showed mean score for elite player (M=11.9, SD = .81), higher than non-elite player (M= 12.9, SD = .83). Thus, 4x10 meters Test was valid among youth rugby player in Selangor.

Table 6. Independent Sample T Test for 4x10 meters Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Agility	Equal variances assumed	.075	.785	-4.634	58	.000	-.98300	.21212	-1.40760	-.55840
	Equal variances not assumed			-4.634	57.971	.000	-.98300	.21212	-1.40760	-.55840

Independent Sample T Test analysis for 30 meters Sprint test was used to compare mean score between elite and non-elite rugby player. Result showed t (58) = -4.027, p = .000 was significant. Result also showed mean score for elite player (M=5.56, SD = .318), higher than non-elite player (M= 5.90, SD = .333). Thus, 30 meters Sprint test was valid among youth rugby player in Selangor.

Table 7. Independent Sample T Test for 30 meters Sprint test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Speed	Equal variances assumed	.003	.954	-4.027	58	.000	-.33900	.08418	-.50750	-.17050
	Equal variances not assumed			-4.027	57.868	.000	-.33900	.08418	-.50750	-.17050

Independent Sample T Test analysis for Standing long jump was used to compare mean score between elite and non-elite rugby player. Result showed t (58) = 2.709, p = .009 was significant. Result also showed mean score for elite player (M=211.8, SD = 8.20), higher than non-elite player (M= 206.5, SD = 6.89). Thus, standing long jump was valid among youth rugby player in Selangor.

Table 8. Independent Sample T Test for Standing long jump

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Power	Equal variances assumed	.000	.982	2.709	58	.009	5.30000	1.95631	1.38401	9.21599
	Equal variances not assumed			2.709	56.322	.009	5.30000	1.95631	1.38153	9.21847

CONCLUSSION

The main objectives of the present study were to determine the reliability and construct validity of the test instruments for skill-related physical fitness components among youth rugby players aged 13 and 14 years old in Selangor. In conclusion, all three instruments (4x10 meters Agility Test, 30 meters Sprint test, and Standing long jump) have their high reliability and the inter testers objectivity. It is essential to the used quality and precision in collecting data to evaluate youth rugby players' performance, either skill or fitness. It will be drawn meaningful generalization about the player so that all the practitioners can plan the training program based on current performance. Valid and reliable data is essential to develop a training plan such as training periodization. Invalid, questionable, and inappropriate data will cause a poor program due to inaccurate generalization. This situation will lead to bad performances of the team in a match. This study will provide comprehensive evidence for all practitioners in terms of appropriate and quality measurement to ensure the training program and talent selection are meaningful. These findings also provide future direction for the other researcher to explore the different research methods, populations, and other fitness components.

REFERENCES

Agar-Newman, D. J., & Klimstra, M. D. (2015). Efficacy of horizontal jumping tasks as a method for talent identification of female rugby players. *Journal of Strength and Conditioning Research*,*29*(3), 737-743

Ahmad Hashim, (2015). *Pengujian, Pengukuran dan Penilaian Pendidikan Jasmani*. Bandar Baru Bangi: Dubook Press Sdn. Bhd.

Argus, C. K., Gill, N., Keogh, J., Hopkins, W. G., & Beaven, C. M. (2010). Effects of a short-term pre-season training programme on the body composition and anaerobic performance of professional rugby union players. *Journal of Sports Sciences*, *28*(6), 679-686.

Baumgartner, T. A., Jackson, A. S., Mahar, M.T., Rowe, D.A. (2006). *Measurement for Evaluation in Physical Education and Exercise Science* (No. Ed.8).WCB/McGraw- Hill.

Baumgartner, T. A., Strong, C. H., & Hensley, L. D. (2005). *Conducting and Reading Research in Health and Human Performance*. McGraw-Hill Humanities/Social Sciences/Languages.

Bailey, R., Collins, D., Ford, P., MacNamara, A., Toms, M., & Pearce, G. (2010). Participant development in sport: An academic review. *Sports Coach UK, 4*, 1-134.

Brewer, J., & Davis, J. (1995). Applied physiology of rugby league. *Sports Medicine*, *20*(3), 129-135.

Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports Medicine, 38*(4), 297-316.

Davidson M. (2014) Known-Groups Validity. In: Michalos A.C. (eds) *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Dordrecht

Duthie, G. M., Pyne, D. B., Marsh, D. J., & Hooper, S. L. (2006). Sprint patterns in rugby union players during competition. *Journal of Strength and Conditioning Research, 20*(1), 208.

Duthie, G., Pyne, D., & Hooper, S. (2003). Applied physiology and game analysis of rugby union. *Sports Medicine, 33*(13), 973-991.

Fokkema, T., Kooiman, T. J. M., Krijnen, W. P., Van der Schans, Cees P, & De Groot, M. (2017). Reliability and validity of ten consumer activity trackers depend on walking speed. *Medicine and Science in Sports and Exercise, 49*(4), 793-800.

Gabbett, T. J., Kelly, J. N., & Sheppard, J. M. (2008). Speed change of direction speed, and reactive agility of rugby league players. *Journal of Strength and Conditioning Research, 22*(1), 174-181.

Granados, C., Izquierdo, M., Ibanez, J., Ruesta, M., & Gorostiaga, E. M. (2008). Effects of an entire season on physical fitness in elite female handball players. *Medicine & Science in Sports & Exercise, 40*(2), 351-361.

Hachana, Y., Chaabene, H., Rajeb, G. B., Khlifa, R., Aouadi, R., Chamari, K., & Gabbett, T. J. (2014). Validity and reliability of new agility test among elite and subelite under 14-soccer players. *PloS one*, *9*(4), e95773.

Hashim, A., Ariffin, A., Hashim, A. T., & Yusof, A. B. (2018). Reliability and Validity of the 90º Push-Ups Test Protocol. *International Journal of Scientific Research and Management*, *6*(06).

Haycraft, J. A., Kovalchik, S., Pyne, D. B., & Robertson, S. (2017). Physical characteristics of players within the Australian Football League participation pathways: a systematic review. *Sports Medicine-Open*, *3*(1), 1-16.

Henriques-Neto, D., Minderico, C., Peralta, M., Marques, A., & Sardinha, L. B. (2020). Test–retest reliability of physical fitness tests among young athletes: The FITescola® battery. *Clinical Physiology and Functional Imaging*, *40*(3), 173-182.

Jarvis, S., Sullivan, L. O., Davies, B., Wiltshire, H., & Baker, J. S. (2009). Interrelationships between measured running intensities and agility performance in subelite rugby union players. *Research in Sports Medicine, 17*(4), 217-230.

Malina, R. M., Cumming, S. P., Kontos, A. P., Eisenmann, J. C., Ribeiro, B., & Aroso, J. (2005). Maturity-associated variation in sport-specific skills of youth soccer players aged 13–15 years. *Journal of Sports Sciences, 23*(5), 515-522.

Meir, R., Holding, R., & Hetherington, J. (2014). Impact of the Two-Handed Rugby Ball Carry on Change of Direction Speed and Reactive Agility: Implications for Sport Specific Testing. *International Journal of Sports Science & Coaching, 9*(5), 1181– 1190

Miller, D. (2013). *Measurement by the physical educator why and how*. McGraw-Hill Higher Education

O'Connor, D. (1996). Physiological characteristics of professional rugby league players. *Strength and Conditioning Coach*, 4(1), 21-26.

Odom, L. R., & Morrow, Jr, J. R. (2006). What's this R? A correlational approach to explaining validity, reliability and objectivity coefficients. *Measurement in Physical Education and Exercise Science*, *10*(2), 137-145.

Oliver, J. L., & Lloyd, R. S. (2012). Long-term athlete development and trainability during childhood: A brief review. *Prof Strength Cond J*, 26, 19-24.

Quarrie, K. L., & Hopkins, W. G. (2008). Tackle injuries in professional rugby union. *The American Journal of Sports Medicine, 36*(9), 1705-1716.

Smart, D., Hopkins, W. G., Quarrie, K. L., & Gill, N. (2014). The relationship between physical fitness and game behaviours in rugby union players. *European Journal of Sport Science, 14*(sup1), S8-S17.

Thomas, J. R., & Nelson, J. K. (2001). *Research Methods in Physical Activity (4th ed.)* Champaign, IL: Human Kinetics.

Till, K., Weakley, J., Read, D. B., Phibbs, P., Darrall-Jones, J., Roe, G., Chantler, S., Mellalieu, S., Hislop, M., Stokes, K., Rock, A., & Jones, B. (2020). Applied sport science for male age-grade rugby union in England. *Sports Medicine-Open*, *6*(1), 1-20.

van de Water, T., Huijgen, B., Faber, I., & Elferink-Gemser, M. (2017). Assessing cognitive performance in badminton players: a reproducibility and validity study. *Journal of Human Kinetics*, 55, 149.

Wang, R., Hoffman, J. R., Tanigawa, S., Miramonti, A. A., La Monica, M. B., Beyer, K. S., Stout, J. R. (2016). Isometric mid-thigh pulls correlates with strength, sprint, and agility performance in collegiate rugby union players. *Journal of Strength and Conditioning Research, 30*(11), 3051-3056

Zumbo, B.D. (2007). *Validity: Foundational issues and statistical methodology. In Handbook of Statistics: Psychometrics. Vol. 26. Edited by C.R. Rao and S. Sinharay*. Elsevier, Amsterdam, Netherlands.

Zumbo, B.D., & Rupp, A.A. (2004). *Responsible Modelling of Measurement Data for Appropriate Inferences: Important Advances in Reliability and Validity Theory.* In the SAGE Handbook of Quantitative Methodology for the Social Sciences. Edited by D. Kaplan. Sage Press, Thousand Oaks, Calif., USA.

Zarizi Ab Rahman
Faculty of Education,
Universiti Teknologi MARA,
Kampus Puncak Alam,
42300 Puncak Alam,
Selangor, Malaysia
Email: zarizi@uitm.edu.my