INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND ANALYSIS

ISSN(print): 2643-9840, ISSN(online): 2643-9875

Volume 06 Issue 11 November 2023

DOI: 10.47191/ijmra/v6-i11-15, Impact Factor: 7.022

Page No. 5111-5119

Test of Validity, Reliability and Relevance of the "ITN 10.3 ITF Test" on the Level of Basic Technical Ability in Forehand and Backhand Groundstrokes for Tennis Players Aged 12 Years and Under

Ngatman[1], Guntur[2], Danang Pujo Broto[3], Muhamad Nanang Solikhin[4], Sunaryo[5]

[1,2,3] Departement Sport Education, Universitas Negeri Yogyakarta & Jl. Colombo No.1 Yogyakarta. Indonesia

[4,5]Vocational Faculty, Universitas Negeri Yogyakarta & Wates Yogyakarta. Indonesia

ABSTRACT: Several skill tests have been developed to assess the level of proficiency in forehand and backhand groundstroke techniques several decades ago, including the International Tennis Number (ITN) 10.3 Test by the ITF. However, the validity, reliability, and suitability of these tests may not necessarily be guaranteed when used to measure proficiency in these two techniques. This research aims to examine the level of validity, reliability, and relevance of the ITN 10.3 Test in relation to the proficiency level of forehand and backhand groundstroke techniques among junior tennis players aged 12 and below in the Special Region of Yogyakarta Province.

This study is a quantitative descriptive research using a survey method. Data collection techniques involve tests and measurements. The population in this study consists of junior tennis players in the Special Region of Yogyakarta. The research sample was selected using purposive sampling technique, consisting of 30 male and female junior tennis players aged 12 and below in Yogyakarta. The research instrument uses the "ITN 10.3 Test" to measure the basic technical proficiency of forehand and backhand groundstroke. Validity test data analysis is performed by correlating the best results from the ITN 10.3 Test with the results of half-competition matches using Pearson Product Moment Correlation, reliability test using Pearson Product Moment Correlation analysis technique, and relevance test is analyzed using T-Score scores.

The research results show that the validity of the ITN 10.3 Test is -0.976. The test reliability is 0.982, and the test relevance is categorized as moderate, good, and very good at 69.99%. Thus, it can be concluded that the ITN 10.3 ITF Test can be used as a standardized assessment instrument to measure the proficiency level of forehand and backhand groundstroke techniques among junior tennis players aged 12 and below in DIY.

KEYWORDS: Validity, Reliability, Relevance, ITN 10.3 Test

I. INTRODUCTION

The tennis skill assessment test used to measure the level of basic technical skills of tennis players in a protocol manner was designed and developed by experts several decades ago. Some of these tennis skill assessment tests include the Dyer Tennis Test, Ronning Tennis Test (revised Dyer Tennis Test), Scott and Frech Tennis Test, Broer Miller Tennis Test, Jonnes Service Tennis Test, Purcell Tennis Test, Hewitt Achievement Tennis Test, and Kemp-Vincent Rally Test (Strand, 1993). In 2004, the International Tennis Federation (ITF) compiled an assessment instrument for basic tennis playing skills called the International Tennis Number (ITN). One of the 10 graded tests within ITN is the ITN 10.3 Test. The ITN 10.3 Test is designed to measure the level of accuracy and consistency in the basic forehand and backhand groundstroke techniques (ITF, 2004).

However, the existence of the ITN 10.3 Test from the ITF does not necessarily guarantee its level of accuracy when used as an assessment instrument to measure the proficiency in basic forehand and backhand groundstroke techniques among junior tennis players in the Special Region of Yogyakarta Province (DIY) who are 12 years old and below. This is because at the time the ITN 10.3 Test was developed, it may not have been aligned with the rapidly evolving trends of modern tennis, where the game now relies on speed and power. The skill level of the tennis players chosen as research samples may not necessarily share the same characteristics, and the anatomical structure between foreign tennis players and junior DIY tennis players can be

significantly different. Additionally, the hegemony of skills between foreign tennis players and junior DIY tennis players who are 12 years old and below is notably distinct. Therefore, based on these considerations, a thorough re-evaluation is needed to determine the extent to which the ITN 10.3 Test possesses validity, reliability, and relevance when used as an assessment instrument to measure the proficiency level of forehand and backhand groundstroke techniques among junior DIY tennis players who are 12 years old and below.

According to Miller (2009), one of the assessment instruments to maintain accuracy and remain reliable over time should undergo periodic review. This periodic review is intended to assess whether the test is still suitable for use as a standard assessment instrument or not. Furthermore, Miller (2009) states that through periodic retesting, inaccurate assessments and outdated evaluations can be avoided. Testing the assessment instrument should always be done to ensure that the instrument used meets the criteria of being a good assessment instrument. An assessment instrument is considered to meet the requirements of a good instrument if it is at least: (1) valid, (2) reliable, (3) objective, and (4) relevant (Miller, 2009; Morrow, 2005; Lacy, 2013; Ngatman, 2017).

Assessment instruments in measuring various sports, including tennis, must have a good level of validity. Validity represents empirical and theoretical support for interpreting test scores in accordance with the intended purpose of the assessment instrument, obtained through continuous evidence that supports the interpretation of test scores from the use of the assessment instrument (Mardapi, 2008; Barnett, et.al., 2023). The data collected (test scores) through the measurement process encompass the actual investigation area to be measured (Ghauri and Gronhaug, 2005). Essentially, validity means "measuring what is intended to be measured" (Field, 2005). Additionally, assessment instruments for sports skills must have a level of stability/consistency to estimate what needs to be measured (Morrow, 2010; Kirk, 2010; Lacy, 2013). Testing the reliability of a sports skill test is important because it refers to consistency throughout the measurement instrument (Huck, 2007). A sports skill test is considered reliable if it produces relatively stable, consistent, and measurement-stable results (Morrow, 2005; Mohajan, 2017).

Sports skill tests must also have objective and relevant requirements. This means that an assessment instrument for tennis skills must be able to depict real situations and objectively evaluate the player's abilities according to the basic technical skills possessed by the trainee. Miller (2009) states that a sports skill test has high objectivity if two or more assessors can administer the test in the same group, obtaining results that are not significantly different or are approximately the same. Furthermore, Miller (2009) and Guntur, et. al., 2020, also state that sports skill assessment instruments must also have a level of compatibility with the situations faced in the field.

Tennis in the DIY Province is a popular sport, alongside football, badminton, and volleyball. Many clubs, schools, and tennis academies have emerged in the DIY Province, specifically to nurture junior tennis players. In the process of nurturing young tennis players, accurate databases are needed to determine both the physical abilities and technical quality possessed by the players. With this data, coaches can use it as a basis for designing training programs accurately. To periodically assess the extent of progress in mastering the basic technical skills of tennis players in clubs/schools, skill assessments must be conducted. However, a current issue in assessing the skill mastery progress of trainees is how a coach can measure a player's skill using an assessment instrument/skill test that is accurate, reliable, and appropriate to the player's skill level.

Based on observations in tennis clubs/schools in the DIY Province, it turns out that the assessment of the level of basic technical skills of junior DIY tennis players has been based solely on observations during training sessions in clubs/schools. The assessments made do not use accurate and truly reliable instruments. This is because they do not yet have standardized and relevant assessment instruments. Therefore, the assessments made by these coaches are less trustworthy, so the accuracy, reliability, and relevance of their assessment results are still in question. To delve further into this issue, the author conducted a preliminary study by interviewing several junior tennis coaches in the DIY Province. Based on the interviews with several coaches in tennis clubs/schools, it turns out that they: (1) do not yet have a standardized test that can be used to objectively measure the ability of basic technical skills, (2) greatly need an assessment instrument that can be used to measure the level of basic tennis playing skills according to the trainee's skill level, (3) want to obtain a reference assessment instrument that can truly be used as a reference to assess the basic technical skill according to the actual tennis playing situation.

Based on this background, the researcher aims to investigate whether the ITN 10.3 Test has accuracy/validity, reliability, and relevance when used as an assessment instrument for basic tennis playing techniques for junior DIY tennis players. By conducting a study on the ITN 10.3 ITF basic skill test, it is hoped that it will be determined whether the ITN 10.3 assessment instrument is accurate, reliable, and relevant when used as a standard instrument to measure the level of basic tennis playing skills.

II. METHODS

A. Research Design

This study employs a quantitative descriptive research design using a survey method, with data collection techniques involving tests and measurements of basic tennis playing skills.

B. Population and Sample

Population

The population for this study comprises Junior Tennis Players in the Special Region of Yogyakarta Province.

Sample

The sampling technique used in this study is purposive sampling, specifically selecting junior tennis players aged 12 and below. This sample was chosen based on the assumption that these players fall within the junior tennis category as defined by the age group regulations set by the International Tennis Federation and PB Pelti (Indonesian Tennis Association).

C. Instruments and Data Collection Techniques

Data Collection Technique

Data collection in this study involves conducting tests and measurements of the level of basic tennis playing skills. This is done using the ITN 10.3 ITF test, which assesses the accuracy and consistency of forehand and backhand groundstrokes. The tennis skills test is conducted in both indoor and outdoor tennis courts at the FIKK UNY (Faculty of Sports Science, Yogyakarta State University).

Data Collection Instrument

The skill assessment to measure tennis playing skills is conducted using the Accuracy and Consistency Test for Forehand and Backhand Groundstrokes in the International Tennis Number (ITN) 10.3. This test aims to measure the level of consistency and accuracy of strokes for beginner-level tennis players in performing forehand and backhand groundstrokes (ITF, 2004).

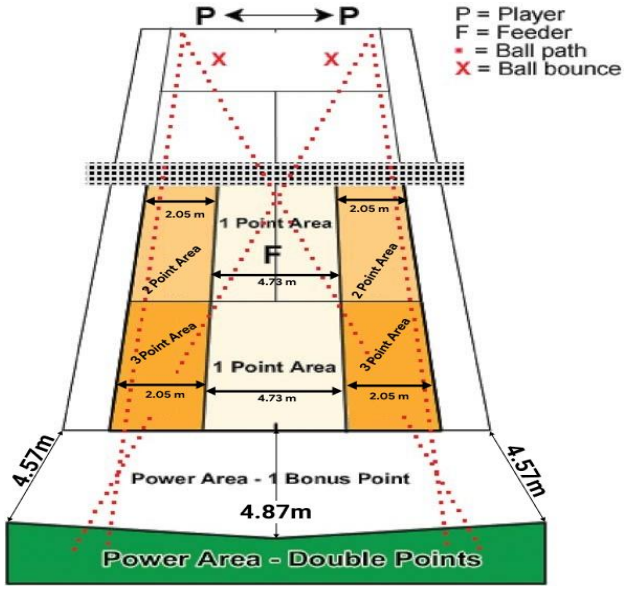

Figure 3.1. Area of Accuracy and Consistency Test for Forehand and Backhand Groundstroke in the International Tennis Number (ITN) 10.3 (ITF, 2004)

Test Implementation Instructions:

- The trainee stands at the baseline (around the center mark area) and is ready to receive balls fed by the feeder.
- The feeder delivers six consecutive balls towards the trainee. The trainee must hit each ball once with a forehand and once with a backhand groundstroke, alternating between the two.
- Each ball hit with a forehand and backhand groundstroke must be directed straight (down the line).
- The feeder then delivers another six balls towards the trainee, who remains in the baseline area.
- The trainee must hit each ball once with a forehand and once with a backhand groundstroke, alternating between the two. Each ball hit with a forehand and backhand groundstroke must cross the court.

Test of Validity, Reliability and Relevance of the "ITN 10.3 ITF Test" on the Level of Basic Technical Ability in Forehand and Backhand Groundstrokes for Tennis Players Aged 12 Years and Under

Scoring System:

The score for each hit is determined by where the ball first lands and where it bounces for the second time. The feeder must ensure that all delivered balls fall between the service line and the baseline, as shown in the diagram. If a ball delivered by the feeder is not favorable, the trainee is not obligated to hit it.

Scoring Criteria:

Score 1: If the ball lands in the middle of the service box or in the area between the service line and the baseline (4.73 meters).

Score 2: If the ball lands in the service box area (before the service line) to the right or left (2.05 meters) of the single game area.

Score 3: If the ball lands behind the service box and before the baseline, to the right or left of the single game area (2.05 meters).

Examples:

Point 4 is awarded if the first bounce lands in point area 3 and the second bounce lands behind the baseline.

Point 6 is awarded if the ball lands in point area 3 and the second bounce lands behind the bonus point line.

Point 0: If the first bounce of the ball falls outside the single game area.

Consistency: An additional point is awarded for each hit if no errors occur (the ball does not touch the net or go out of bounds in the single game area).

The maximum possible score for this accuracy and consistency test of basic forehand and backhand groundstroke technique in ITN is 84, with the breakdown as follows:

Straight forehand and 3 straight backhand groundstrokes X 6 (maximum points) = 36

Cross court forehand and 3 cross court backhand groundstrokes X 6 (maximum points) = 36

12 bonus points if the trainee does not make any unforced errors during the 12 forehand and backhand groundstroke hits.

The earned points will increase based on the overall achievement of the total score points in the accuracy assessment of forehand and backhand groundstroke (ITF, 2004).

A. Data Analysis Technique

To assess the accuracy of the measurement results of the test in the study, the following data analysis techniques are used:

a. To test the validity of the ITN 10.3 Test as a criterion test, the results are compared with the outcomes of half-competition matches using the pro set tiebreak system. Each player will participate in a half-competition match to achieve a tiebreak score of 10 points using the rally point system.

b. To test the reliability, a test-retest method is used. Research data from the test-retest are then analyzed using the Pearson Product Moment Correlation analysis technique. The formula for Pearson Product Moment Correlation is as follows

$$r_{XY} = \frac{N.\sum XY - (\sum X)(\sum Y)}{\sqrt{\left\{N.\sum X^2 - (\sum X)^2\right\}\left\{N.\sum Y^2 - (\sum Y)^2\right\}}}$$

(Huck, 2007; Ngatman, 2022)

a. To assess the appropriateness of the test, normative assessment categories are established for the ITN 10.3 test data. The obtained data is then analyzed using T-Score. The formula for T-Score is as follows:

TScore = 10 Z + 50

Dengan diketahui informasi berikut:

Harga setiap satu standard deviasi dari distribusi TScore = 10

Rerata dari distribusi TScore = 50

Z = (Rerata - X)/(Standar Deviasi)

Next, from the TScore value, the average value (mean) and standard deviation (SD) are calculated and then converted into the following assessment norm category table.:

Norm Range	Category
$\overline{X}$ + 1,5 SD ke atas	Very Good

$\overline{X}$ + 0,5 SD s/d < $\overline{X}$ + 1,5 SD	Good
$\overline{X}$ - 0,5 SD s/d < $\overline{X}$ + 0,5 SD	Average
$\overline{X}$ - 1,5 SD s/d < $\overline{X}$ - 0,5 SD	Bad
kurang dari $\overline{X}$ - 1,5 SD	Very Bad

(Ngatman, 2022)

RESEARCH RESULTS AND DISCUSSION

A. RESEARCH RESULTS

This study involves two variables: the independent variable (X) which is the International Tennis Number (ITN) 10.3 Test from the ITF (used as a predictor), and the dependent variable (Y) which is the outcome of half-competition matches in search of a score of 10 using the pro set tiebreak system. The data collection methods used in this study include the test method and retest method for the independent variable, and half-competition matches for the dependent variable as the criterion score. From the half-competition matches, rankings will be obtained for each research sample, ranging from the highest to the lowest ranking.

To determine the level of test validity, the best data from the ITN 10.3 Test and retest will be correlated with the rankings from the half-competition matches. The data analysis technique used is the Pearson Product Moment Correlation, which will determine the coefficient of test validity (rX1Y). To determine the coefficient of reliability of the International Tennis Number 10.3 Test (rX1X2), it will be obtained from the results of the test-retest. The statistical analysis technique used is the Pearson Product Moment Correlation.

The detailed description of the research data will be presented as follows:

Validity of the International Tennis Number 10.3 Test

Based on the data from the ITN 10.3 Test and retest, as well as the rankings from the half-competition matches of junior tennis players aged 12 and below in the DIY Province, the following values were obtained: Mean X = 58.60, Standard Deviation X = 10.84. Mean Y = 15.50, while Standard Deviation Y = 8.80. Table 1 presents the research data for the results of the ITN 10.3 Test and rankings from the half-competition matches (Y).

Table 1. International Tennis Test Results Number 10.3 (X) and Ranking of Match Results 1/2 of the competition (Y)

No.	***Test International Tennis Number* 10.3 (X)**	**Match Results Ranking in 1/2 competition (Y)**
1.	72	3
2.	69	8
3.	58	17
4.	64	9
5.	74	2
6.	54	21
7.	48	24
8.	72	5
9.	70	4
10.	64	16
11.	55	20
12.	47	23
13.	71	6
14.	63	12
15.	63	15
16.	54	19
17.	46	27
18.	56	18
19.	72	7
20.	41	28
21.	45	26
22.	50	25
23.	76	1
24.	39	30

25.	47	22
26.	62	11
27.	59	13
28.	57	14
29.	44	29
30.	66	10

Diketahui:

N = 30 ∑XY = 24547

∑X = 1758 ∑Y = 465

∑X2 = 106424 ∑Y2 = 9455

rXY = - 0,976706734

Upon analyzing the research data using the Pearson Product Moment Correlation, the coefficient of validity for the International Tennis Number 10.3 Test (X) and the rankings from the half-competition matches (Y) was found to be -0.977. Based on the normative range of validity coefficients according to Strand (1993); Azwar (2009); Mohajan (2017), the validity of the International Tennis Number 10.3 Test in measuring the skill level of junior tennis players aged 12 and below in the DIY Province is considered to be high. Therefore, the ITN 10.3 Test demonstrates a high level of accuracy when used as an assessment tool to measure the skill level of forehand and backhand groundstrokes in junior tennis players aged 12 and below in the DIY Province.

Reliability of the International Tennis Number 10.3 Test

Based on the data from the ITN 10.3 Test and retest of junior tennis players aged 12 and below in the DIY Province, the following values were obtained: Mean X1 = 54.87, Standard Deviation X1 = 10.77. Mean X2 = 58.60, while Standard Deviation X2 = 10.84. The research data for the test-retest of the International Tennis Number 10.3 (X1) and (X2) are presented in Table 2 below.

Table 2. International Tennis Number 10 test and retest results.3

No.	Tes I (X_1)	Tes II (X_2)
1.	69	72
2.	66	69
3.	52	58
4.	60	64
5.	72	74
6.	49	54
7.	48	48
8.	70	72
9.	70	70
10.	61	64
11.	53	55
12.	47	47
13.	68	71
14.	57	63
15.	60	63
16.	48	54
17.	41	46
18	50	56
19.	65	72
20.	39	41
21.	42	45
22.	44	50
23.	72	76
24.	38	39
25.	44	47
26.	57	62
27.	53	59
28.	51	57
29.	40	44
30.	60	66

Diketahui:

N = 30 ∑X1X2 = 99780

∑X1 = 1646 ∑X2 = 1758

∑X12 = 93676 ∑X22 = 106424

rX1X2 = 0,982017697

After conducting the analysis using the Pearson Product Moment Correlation, the coefficient of reliability for the International Tennis Number 10.3 Test was found to be 0.982. Based on the normative range of reliability coefficients according to Strand (1993); William and Lacy (2018), the reliability of the International Tennis Number 10.3 Test in measuring the skill level of forehand and backhand groundstrokes in junior tennis players aged 12 and below in the DIY Province falls into the category of "excellent". Therefore, the ITN 10.3 Test demonstrates an exceptionally high level of consistency and reliability, making it suitable for use as an assessment tool to measure the ability of forehand and backhand groundstrokes in junior tennis players aged 12 and below in the DIY Province.

1. Relevance of International Tennis Test Number 10.3

Based on the test results data from the International Tennis Number 10.3 (Y), the skill level of forehand and backhand groundstrokes in junior tennis players aged 12 and below in the DIY Province obtained a mean (X2) of 58.60 and a Standard Deviation (SD) of 10.84. The data was further analyzed using T-Score. The complete data analysis results can be seen in Table 3 below.

Table 3. Relevance of International Tennis Number 10.3

Norm Stretch	Frequency	Percentage	Category
75 ke atas	1	3,33 %	Very Good
64 - 74	10	33,33 %	Good
53 - 63	10	33,33 %	Average
42 - 52	7	23,33%	Bad
41 ke bawah	2	6,68 %	Very Bad
Jumlah	**30**	**100%**	

From the data analysis, it can be concluded that the International Tennis Number 10.3 Test is relevant when used as an assessment tool to measure the skill level of forehand and backhand groundstrokes in junior tennis players aged 12 and below in the DIY Province. This is evidenced by 69.99% of junior tennis players aged 12 and below possessing forehand and backhand groundstroke skills assessed by the ITN 10.3 Test falling into the categories of moderate, good, and excellent. The detailed breakdown of categories is as follows: moderate (33.33%), good (33.33%), and excellent (3.33%).

B. DISCUSSION

Validity is a prerequisite for an assessment instrument and should be the primary consideration when using the instrument for evaluation. This is because if the assessment instrument meets this requirement, it will have a degree of assessment accuracy (Azwar, 2009; Mohajan HK, 2017; Lutfi Nur, et. al, 2018; Barnett, et. al., 2023). The degree of accuracy or validity of the instrument does not always serve as the assessment standard. Periodically, the instrument must be tested to review whether it is still suitable for use or not (Miller, 2009; William and Lacy, 2018; Lutfi & Maslacki, 2020). Based on the data analysis, the validity coefficient of the International Tennis Number 10.3 Test (X) and the ½ competition result ranking = -0.977. This negative validity coefficient indicates that one of the research variables (variable X or Y) correlated is inverted data (Y = ½ competition result ranking). This indicates that the validity coefficient of this test is still acceptable or appropriate to be used as one of the assessment instruments to measure the level of forehand and backhand groundstroke skills for junior tennis players aged 12 and under in DIY Province (Strand (1993:11). Rink (2002); Martinez (2014) state that if the validity coefficient of sports skill tests is -0.70 or above, then the assessment instrument is suitable.

The reliability of the sports skill assessment instrument is one of the very important criteria that must be considered when using the skill instrument in a sports branch. This is because reliability is the level of stability/consistency to what extent the assessment instrument can perform its measuring function. To what extent if the instrument is used more than once to measure

the same child's skills, it produces no different scores (O'Donoghue, 2010; Azwar, 2011; Fraenkel, J. R., Wellen, N., & Hyun, H.H., 2012; Barnett, et. al., 2023). The magnitude of the reliability coefficient of the International Tennis Number 10.3 Test for junior tennis players aged 12 and under in DIY Province = 0.982, thus including an excellent category. This reliability coefficient of 0.982 indicates that the level of forehand and backhand groundstroke skills of junior tennis players aged 12 and under in DIY Province has a very high level of consistency/stability. This is because the forehand and backhand groundstroke test scores from the first test to the second test are not significantly different. Rink (2002) says that if a sports skill test has a reliability coefficient of 0.70 or above, then the instrument is very good for use as an instrument to measure sports skills, including tennis.

In addition to validity and reliability, relevance or appropriateness of sports skill assessment instruments is one of the requirements that must be met when using the assessment instrument. According to Ngatman and Andriyani, (2016); Guntur et al., (2020), an assessment instrument is considered relevant if testing is carried out on the sample used to test the instrument. 70% have assessment norm categories of moderate, good, and excellent. The research data analysis shows that the International Tennis Number 10.3 Test for junior tennis players aged 12 and under in DIY Province is relevant for use as an assessment instrument to measure the level of proficiency in forehand and backhand groundstroke. The majority of forehand and backhand groundstroke skill levels using the International Tennis Number 10.3 assessment instrument fall under the categories of moderate, good, and excellent, at 69.99%. The detailed categories are as follows: moderate = 33.33%, good = 33.33%, and excellent = 3.33%.

Accuracy in choosing an assessment instrument relevant to the skill level of trainees will have implications for drawing conclusions. The conclusion of the assessment results is greatly determined by how well the measurement tool used in the assessment matches the trainee's skill level and age. If the measurement tool used as an assessment instrument is suitable for the trainee's skill level and age, it will be very useful in developing assessment norms for the trainee's skill level. With this assessment norm, it can be used to differentiate skill levels between trainees (Strand, 1993; Morrow, 2005; Oluwatayo, et.al., 2012; Izwan, et.al., 2015; Ngatman, 2017).

V. CONCLUSIONS

The validity coefficient of the ITN 10.3 Test (X) and the ½ competition result ranking for junior tennis players aged 12 and under in the DIY Province (Y) is -0.977. Therefore, the ITN 10.3 Test is accurate in assessing the level of forehand and backhand groundstroke skills for junior tennis players aged 12 and under. The reliability coefficient of the ITN 10.3 Test is excellent. Based on the data analysis, the reliability coefficient is 0.982. This indicates that the ITN 10.3 Test demonstrates a high level of consistency when used as an assessment instrument to measure the level of forehand and backhand groundstroke skills for junior tennis players aged 12 and under in the DIY Province. The ITN 10.3 Test exhibits a very good level of relevance when used as an assessment instrument to measure the level of forehand and backhand groundstroke skills for junior tennis players in the DIY Province aged 12 and under. The relevance level of the ITN 10.3 Test falls under the categories of moderate, good, and excellent, at 69.99%.

ACKNOWLEDGMENT

I wish to extend my heartfelt gratitude to the esteemed Rector of UNY, Prof. Dr. Sumaryanto, M. Kes., and the distinguished Dean of FIKK UNY, Prof. Dr. Wawan S. Suherman, M. Ed., for their invaluable guidance and mentorship throughout this research endeavor. My sincere thanks also go to Wisnu Nugroho, M. Or., M. Ilham, S. Pd.Jas, Damas Adi Surya, Igor Sudiyaksa, and Budi Setyo Wahadi, S. Pd., for their generous assistance and facilitation during the course of this study."

REFERENCES

1) Azwar, S. (2009) Reliabilitas dan Validitas. Jakarta: Pustaka Pelajar
2) _____. (2011) Reliabilitas dan Validitas. Jakarta: Pustaka Pelajar
3) Barnet, L.M., Jerebine, A., Keegan, R., Watson-Mackie, K., Arundell, L., Ridgers, N. D.,
4) Salmon, J., & Dudley, D. (2023) Validity, Reliability, and Feasibility of Physical Literacy Assessments Designed for School Children: A Systematic Review. Journal Sport Medicine (Juni 2023).
5) Field, A. P. (2005). Discovering Statistics Using SPSS, Sage Publications Inc.
6) Fraenkel, J. R., Wellen, N., & Hyun, H.H. (2012). How to Design and Evaluate Research in Education (eighth). New York: McGraw-Hill.
7) Ghauri, P. & Gronhaug, K. 2005. Research Methods in Business Studies, Harlow, FT/Prentice Hall

8) Guntur, Ngatman, Sridadi, Broto, D. P, & Pambudi, D. K. (2020). Pengujian Validitas, Reliabilitas, Dan Relevansi Norma Penilaian"Dyer Tennis Test" Terhadap Tingkat Keterampilan Bermain Tenis. Jurnal Keolahragaan Vol. 8, No. 2 (December 2020). Jurnal Keolahragaan 8(2):195-203 DOI:10.21831/jk. V 8i2.35266.
9) Huck, S. W. 2007. Reading Statistics and Research, United States of America, Allyn & Bacon.
10) International Tennis Federation. (2004). ITN on Court Assessment. Bank Lane, Roehampton, London: International Tennis Federation.
11) Izwan, M., Norkhalid, B.S., & Gunathevan, E. (2015). The Validity and Reliability of ISO Test Toward The Performance Assessment of Future Physical Education Teachers in Teaching and Learning Process.
12) Kirk, M. F. & Lund, J. L. (2010). Performance Based Assessment for Middle and Hight School Physical Education. USA. Human Kinetics Publishers
13) Lacy, Alan C. (2013). Measurement & Evaluation in Physical Education and Exercise Science (Sixth Edition). San Fransisco, CA: Pearson Benjamin Cummings Publishing.
14) Lutfi, N., Suherman, A., Subarjah, H., & Budian, D. (2018). Validity and Reliability of Physical Education Learning Motivation Questionnaires. 3rd International Conference on Sport Science, Health, and Physical Education (ICSSHPE 2018). Advances in Health Sciences Research, volume 11.
15) Lutfi Surucu dan Ahmet Maslakci. (2020). Validity and Reliability in Quantitative Research. Journal Business and Management Studies an International Journal, October 2020. DOI: 10.15295/bmij.v8i3.1540.
16) Mardapi, Djemari. (2008). Teknik Penyusunan Instrumen Tes dan Nontes. Yogyakarta: Mitra Cendikia Press.
17) Martínez-Cantó, A., Hernández-Davo, J. L., Sarabia, J.M., Guillé, S., Fernández, J., & Moya, M. (2014). Validity And Reliability of a Specific Repeated Sprint Ability Test for Young Tennis Players. VIII International Congress of the Spanish Association of Sports Sciences (Cáceres, November 13, 14 and 15, 2014) ISBN: 84-697-1248-9.
18) Miller, D.K. (2009). Measurement By The Physical Educator Why And How 5th ed. Dubuque, Iowa: Wm. C. Brown.
19) Mohajan, H. K. (2017). Two Criteria for Good Measurement in Research: Validity and Reliability. Munich Personal RePEc Archive. Annals of Spiru Haret University, 17(3): 58-82. https://mpra.ub.uni-muenchen.de/83458/MPRA Paper No. 83458, posted 24 Dec 2017 08:48 UTC.
20) Morrow, J. R., Jackson, A. W., Disch, J. G., & Mood, D. P. (2005). Measurement And Evaluation In Human
21) Ngatman dan Andriyani, F.D. (2016). Uji Validitas, Reliabilitas, Dan Relevansi "Kemp Vincent Rally Tennis Test" Terhadap Tingkat Keterampilan Bermain Tenis Bagi Mahasiswa FIK UNY. (https://staffnew.uny.ac.id/upload/132093449/penelitian/Uji%20Validitas%20,Reliabilitas,Dan%20Relevansi%20KempVincent%20Rally%20Tennis%20Test%20Terhadap%20Tingkat%20Keterampilan%20Bermain%20Tennis%20Bagi%20Mahasiswa%20Fik%20UNY1.pdf).
22) Ngatman. (2017). Evaluasi Pendidikan Jasmani Olahraga dan Kesehatan. Grobogan: CV. Sarnu Untung.
23) _____. (2022). Statistika Untuk Penelitian Bidang Pendidikan Jasmani Olahraga dan Kesehatan. Yogyakarta: UNY Press.
24) O'Donoghue, P. (2010). Research Methods for Sports Performance Analysis.Oxon: Routledge.
25) Oluwatayo, J. A. (2012). Validity and Reliability Issues in Educational Research. Journal of Educational and Social Research, 2(2), 391-400.
26) Rink, J. E. (2002). Teaching Physical Education for Learning (4th edition). New York: McGraw Hill.
27) Strand, Bradford N., Wilson, Rolayne. (1993). Assesing Sport Skills. United State of America: Human Kinetics Publishers.
28) William, S. M., & Lacy, A. (2018). Measurement and Evaluation in Physical Education and Exercise Science (8th edition): London. Routledge.