Development of Basketball Physical and Technical Test Instruments: Validity and Reliability Analysis

Patriana Nurmansyah Awwaludin[1], Masayu Rizka Risjanna [2], Muhammad Naufal Abdurahman [3], Mona Fiametta Febrianty [4]

[1] Pendidikan Kepelatihan Olahraga, Fakultas Pendidikan Olahraga dan Kesehatan,Universitas Pendidikan Indonesia, Indonesia
[2,4] Program Studi Kepelatihan Fisik Olahraga, Fakultas Pendidikan Olahraga dan Kesehatan, Universitas Pendidikan Indonesia, Indonesia
[3] Program Studi Ilmu Keolahragaan, Fakultas Pendidikan Olahraga dan Kesehatan, Universitas Pendidikan Indonesia, Indonesia

email: patriananurmansyah@upi.edu[1]

Abstract

The development of measurement instruments for basketball skills and physical performance is essential to support athlete training programs based on systematic and data-driven approaches. Although basketball in Indonesia has grown steadily, the transition from amateur to professional levels still faces challenges, one of which is the absence of standardized measurement tools that are both valid and reliable across different stages of athlete development. This study aimed to develop and examine the validity and reliability of basketball skill and physical test instruments intended for athletes aged 12–18 years. A quantitative methodological approach was employed, focusing on content validation through expert judgment. The validation process involved ten experts in basketball and sport science who assessed the relevance of each test item. Content validity was analyzed using the Content Validity Ratio (CVR), while inter-rater reliability was examined with Fleiss' Kappa. The results indicated that of the 11 skill test items, 7 were valid and reliable, with coefficients ranging from moderate to almost perfect. However, several items, such as passing version II, right- and left-hand dribbling, and speed dribbling shoot, were deemed invalid and showed low reliability (slight). In the physical test domain, 8 of 10 items were validated and demonstrated almost perfect reliability, while the standing long jump and hexagon test were found invalid and yielded low reliability (poor). These findings suggest that the developed instruments meet validity and reliability standards for most items, and thus hold potential as reference tools for assessing physical and technical performance in youth basketball. Furthermore, the instruments provide a foundation for future studies involving field trials to confirm their effectiveness in athlete development programs.

Keywords: *Basketball, Measurement Instrument, Validity, Reliability, Athlete Development*

INTRODUCTION

The development of measurement tools or instruments is an essential component in the process of sports training. These tools not only serve as a means for monitoring and evaluating athletes' conditions but also provide a basis for making strategic decisions in designing training schemes (Hoffman, 2014). Well-designed and specific instruments enable the collection of accurate and valid data, which ultimately helps coaches develop training programs tailored to athletes' needs (Tanner & Gore, 2013; Abrar & Fitroni, 2021). The presence of standardized measurement tools also plays a crucial role in facilitating consistent monitoring and evaluation of athletic performance. Clear standards allow coaches, researchers, and educators to employ the same methods to assess and track athletes' progress (Balyi et al., 2020). This provides the foundation for building data-driven approaches to enhance sports performance (Mancha-Triguero et al., 2019).

Basketball is one of the most popular sports in Indonesia, experiencing rapid growth in recent years. However, the athlete development process from amateur to professional levels still faces several challenges, one of which is the limited availability of standardized, valid, and reliable measurement instruments to assess technical skills and physical conditioning. Accurate measurement is vital within the framework of the Athlete Development Model (ADM), which emphasizes stages of growth from building fundamental physical literacy, developing the game, to competing at the national level (Balyi, 2008). Furthermore, the increasing number of naturalized players in professional competitions highlights the need to optimize the development of local athletes through more systematic, data-driven evaluation systems (Asmawi et al., 2022). A good measurement instrument must meet the criteria of validity (the ability to measure what it is intended to measure) and reliability (the consistency of measurement results) (Anastasi & Urbina, 1988).

Several previous studies have underscored the importance of performance measurement instruments in sports. For example, Barnett et al. (2016) emphasized the role of skill assessment tools in supporting talent identification. In the context of basketball, Fox et al. (2017) found that technical skills such as shooting, passing, dribbling, and defensive movement correlate strongly with competitive performance. Likewise, physical components such as strength, speed, and endurance have been shown to be significant predictors of athletic performance (Drinkwater et al., 2008). However, in Indonesia, research on the development of basketball measurement instruments that

comprehensively test validity and reliability remains limited. Therefore, this study aims to develop basketball-specific test instruments for technical and physical performance and to examine their validity and reliability through expert involvement, thus providing a basis for further refinement.

METHOD

This study employed a methodological research design, focusing on the development and testing of measurement instruments. The primary objective of methodological research is to ensure that the instruments developed demonstrate adequate validity and reliability, thereby allowing their application in accurate and consistent measurement (Anastasi & Urbina, 1988). The instruments developed in this study were designed to assess the technical and physical abilities of basketball athletes aged 12–18 years, particularly those at the regional development level. Accordingly, the target population for these instruments consisted of athletes from cities or districts that achieved first place in the West Java Division I Basketball Championship. However, the current stage of this study focused on expert validation; therefore, the direct participants involved were experts serving as instrument validators.

The research procedure was conducted in two main stages. First, content validity testing was carried out through expert judgment. A total of ten experts were involved, comprising five basketball coaches and five sport science specialists. Basketball coaches were selected based on the following criteria: possession of a recognized basketball coaching license, a minimum of ten years of coaching experience, and experience in training athletes aged 12–18 years. Meanwhile, sport science specialists were required to hold a national or international strength and conditioning certification and to have experience in training elite athletes. The experts were asked to evaluate the relevance of each instrument item using the Content Validity Ratio (CVR) and the Content Validity Index (CVI), in accordance with Lawshe's (1975) method. An item was considered valid if it achieved a CVR score of ≥ 0.99.

Second, reliability testing was conducted to determine inter-rater consistency regarding the validated instruments. Reliability analysis employed Fleiss' Kappa, with interpretation guided by the criteria proposed by Landis and Koch (1977), namely: <0 (poor), 0.00–0.20 (slight), 0.21–0.40 (fair), 0.41–0.60 (moderate), 0.61–0.80 (substantial),

and 0.81–1.00 (almost perfect). Through these stages, the developed instruments were expected to demonstrate strong content validity and high inter-rater consistency, thereby ensuring readiness for subsequent field trials in future research.
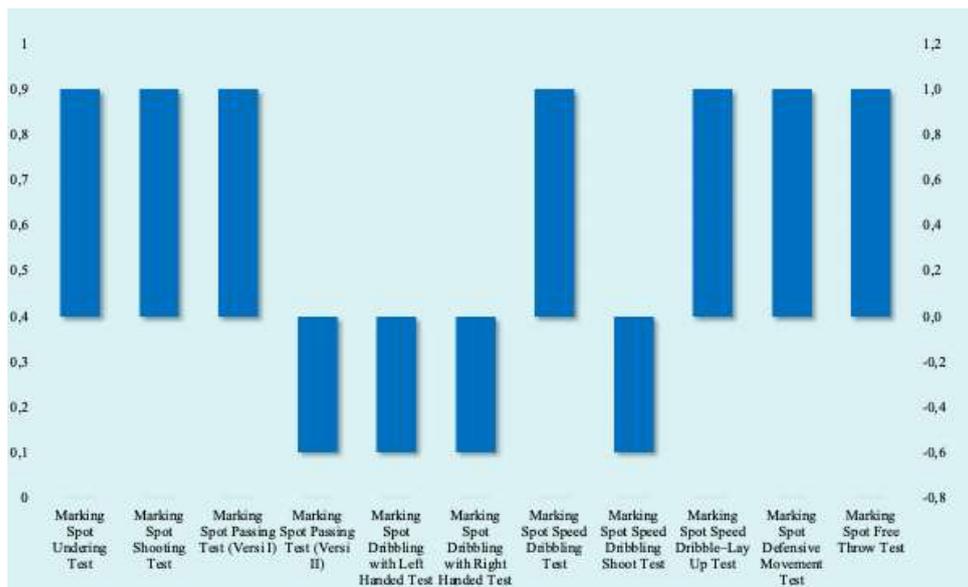
RESULT

This study produced measurement instruments consisting of seven technical test items and nine physical test items for basketball. The content validity testing, conducted by ten experts (five in strength and conditioning and five in basketball coaching specializing in technical skills), revealed varying levels of appropriateness across test items. Of the eleven technical items, seven were deemed valid based on the Content Validity Ratio (CVR). The validated instruments included essential fundamental movements such as shooting, passing, and dribbling, as well as advanced skills such as speed dribbling, speed dribble lay-up, defensive movement, and free throw. Meanwhile, four items—passing test version II, left-handed dribbling, right-handed dribbling, and speed dribbling shoot test—did not meet the validity criteria (CVR < 0.99). The results of the technical validity testing are presented in Table 3.1.

Table 3.1. Results of Validity Testing for Basketball Technical Instruments (Ages 12–18 Years)

Nomor	Technical Item Test	CVR Value	Interpretation
1	Marking Spot Undering Test	1,0	Valid
2	Marking Spot Shooting Test	1,0	Valid
3	Marking Spot Passing Test (Versi I)	1,0	Valid
4	Marking Spot Passing Test (Versi II)	-0,6	Invalid
5	Marking Spot Dribbling with Left-Handed Test	-0,6	Invalid
6	Marking Spot Dribbling with Right-Handed Test	-0,6	Invalid
7	Marking Spot Speed Dribbling Test	1,0	Valid
8	Marking Spot Speed Dribbling Shoot Test	-0,6	Invalid
9	Marking Spot Speed Dribble Lay-Up Test	1,0	Valid
10	Marking Spot Defensive Movement Test	1,0	Valid
11	Marking Spot Free Throw Test	1,0	Valid

The table 3.1 shows that items 1, 2, 3, 7, 9, 10, and 11 (a total of seven items) achieved CVR scores above 0.99, indicating validity. These items included: (a) undering test, (b) shooting test, (c) passing test version I, (d) speed dribbling test, (e) speed dribble lay-up test, (f) defensive movement test, and (g) free throw test. Conversely, items 4, 5, 6, and 8—namely (a) passing test version II, (b) dribbling with left hand, (c) dribbling with right hand, and (d) speed dribbling shoot test—were deemed invalid (CVR < 0.99). A graphical representation of these findings is presented in Figure 3.1.

Figure 3.1 Results of the Validity Test of the Basketball Technique Instrument

(Ages 12-18 Years)

For the physical instrument, validation results showed that eight out of ten items were valid. The validated items covered essential biomotor components, including flexibility, power, speed, agility, strength endurance, and aerobic fitness. However, the standing long jump and hexagon test did not meet the validity criteria. The physical test validity results are presented in Table 3.2.

Table 3.2. Results of Validity Testing for Basketball Physical Instruments

(Ages 12–18 Years)

No Butir	Item Tes Fisik	Nilai CVR	Interpretasi
1	Sit and Reach	1	Valid
2	No-step Vertical Jump	1	Valid
3	One-step Vertical Jump	1	Valid
4	Standing Long Jump	-0,6	Invalid
5	Hexagon Test	0,6	Invalid
6	3/4 Court Sprint	1	Valid

7	Push-up	1	Valid
8	Sit-up	1	Valid
9	Line Agility Drill	1	Valid
10	Multistage Fitness Test	1	Valid

The table indicates that items 1, 2, 3, 6, 7, 8, 9, and 10 (a total of eight items) achieved CVR scores above 0.99, indicating validity. These items included: (a) sit and reach, (b) no-step vertical jump, (c) one-step vertical jump, (d) ¾ court sprint, (e) push-up, (f) sit-up, (g) line agility drill, and (h) multistage fitness test. Conversely, items 4 and 5 (standing long jump and hexagon test) were deemed invalid (CVR < 0.99). A graphical summary of these results is presented in Figure 3.2.

Figure 3.2 Results of the Validity Test of the Physical Instrument for Basketball Players (Aged 12-18 Years)
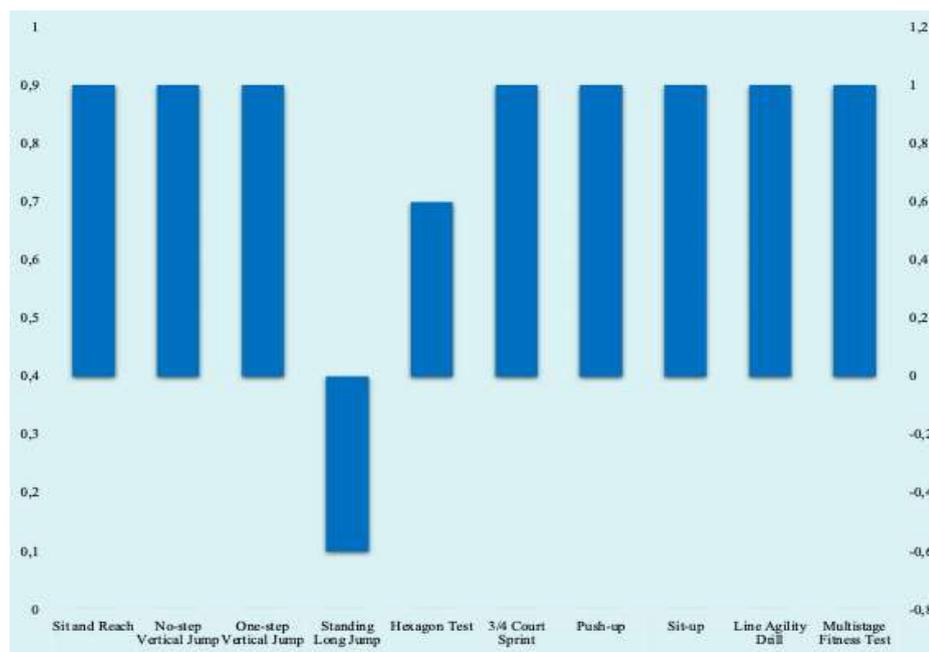
Reliability testing using Fleiss' Kappa demonstrated high consistency for the majority of items. For the technical instruments, six items achieved reliability classified as *almost perfect* (k > 0.81), while four items fell into the *slight* (k = 0.00–0.20) to *moderate* (k = 0.41–0.60) categories. This indicates that fundamental and advanced technical test items could be applied consistently across different raters. For the physical instruments, eight items demonstrated *almost perfect* reliability, while the standing long jump and hexagon test showed low reliability (*poor*, K < 0.20). These findings confirm that

most physical test items yield stable and consistent results, though revisions are required for the less reliable items. The results of reliability testing for both technical and physical instruments are shown in Tables 3.3 and 3.4.

Table 3.3. Results of Reliability Testing for Basketball Technical Instruments (Ages 12–18 Years)

Nomor	Item Test	Nilai Reliabilitas	Interpretasi (Landis & Koch, 1977)
1	Marking Spot Undering Test	0,6	Moderate
2	Marking Spot Shooting Test	1	Almost Perfect
3	Marking Spot Passing Test (Versi I)	1	Almost Perfect
4	Marking Spot Passing Test (Versi II)	0,2	Slight
5	Marking Spot Dribbling with Left Handed Test	0,2	Slight
6	Marking Spot Dribbling with Right Handed Test	0,2	Slight
7	Marking Spot Speed Dribbling Test	1	Almost Perfect
8	Marking Spot Speed Dribbling Shoot Test	0,2	Slight
9	Marking Spot Speed Dribble Lay-Up Test	1	Almost Perfect
10	Marking Spot Defensive Movement Test	1	Almost Perfect
11	Marking Spot Free Throw Test	1	Almost Perfect

The results presented in Table 3.3 show that six technical test items achieved *almost perfect* reliability: shooting test, passing test version I, speed dribbling test, speed dribble lay-up test, defensive movement test, and free throw test. The undering test achieved *moderate* reliability (k = 0.6), while passing test version II, left-handed dribbling, right-handed dribbling, and speed dribbling shoot test demonstrated only *slight* reliability.

Table 3.4. Results of Reliability Testing for Basketball Physical Instruments (Ages 12–18 Years)

Nomor	Item Test	Reliabilitas	Interpretasi (Landis & Koch, 1977)
1	Sit and Reach	1	Almost Perfect
2	No-step Vertical Jump	1	Almost Perfect
3	One-step Vertical Jump	1	Almost Perfect
4	Standing Long Jump	-0,24	Poor
5	Hexagon Test	-0,24	Poor

6	3/4 Court Sprint	1	Almost Perfect
7	Push-up	1	Almost Perfect
8	Sit-up	1	Almost Perfect
9	Line Agility Drill	1	Almost Perfect
10	Multistage Fitness Test	1	Almost Perfect

Table 3.4 demonstrates that eight physical test items achieved *almost perfect* reliability: sit and reach, no-step vertical jump, one-step vertical jump, ¾ court sprint, push-up, sit-up, line agility drill, and multistage fitness test. By contrast, standing long jump and hexagon test were categorized as *poor* (k < 0.20).

DISCUSSION

The findings of this study reinforce the importance of developing validated and reliable instruments in basketball, particularly for athlete development programs. The validity testing indicated that most of the developed tests aligned with the intended constructs. For instance, the shooting test and defensive movement test were deemed valid, as they reflect essential basketball skills. This is consistent with the findings of Fox et al. (2017), who emphasized that fundamental skills such as shooting, passing, and defense are critical indicators of competitive performance. Conversely, several items, such as passing test version II and dominant-hand dribbling (right and left), were found invalid. This could be attributed to test designs that did not sufficiently reflect real-game contexts or the unique roles of different player positions (Bompa & Buzzichelli, 2015). Memmert and Harvey (2008) highlighted that skill assessment should ideally be game-based to capture the competitive context more accurately. Future instrument development should therefore integrate aspects of decision-making and situational pressure to enhance validity.

In the physical domain, the validity and reliability results confirmed that core components—such as sit and reach, vertical jumps, ¾ court sprint, and multistage fitness test—are appropriate measures for flexibility, lower-body power, speed, and aerobic endurance. Drinkwater et al. (2008) similarly reported that these parameters are significant predictors of elite basketball performance. However, the standing long jump and hexagon test showed weaknesses in both validity and reliability. Horizontal jump

tests may be less suitable for evaluating vertical power, which is more relevant to basketball movements (Cormack et al., 2008). Meanwhile, the hexagon test appears less representative because its jumping patterns do not closely replicate basketball-specific movement demands. According to expert feedback, agility tests such as shuttle runs may provide a more suitable alternative, as they involve rapid acceleration, deceleration, and directional changes, which better simulate actual basketball demands. Pauole et al. (2000) also confirmed that shuttle run variations are valid and reliable indicators of agility in team sports. Thus, future development should consider adopting shuttle runs to obtain more relevant and applicable agility measurements.

The practical implications of this study are significant. First, valid and reliable instruments allow coaches to make more objective, data-driven decisions in athlete selection. Second, these instruments enable periodic monitoring of athletes' technical and physical performance, thereby allowing coaches to adjust training programs according to individual and team needs. Overall, the findings emphasize the importance of valid and reliable measurement tools in basketball. The developed instruments proved effective in measuring most technical and physical components, making them applicable for both athlete selection and development. Moreover, this study provides a foundation for further refinement and competitive implementation. This aligns with Barnett et al. (2016), who stressed the need for repeated testing to achieve universal recognition of assessment instruments. In summary, this study represents an important step toward providing standardized basketball measurement tools in Indonesia, serving as a national reference for youth athlete development and supporting policies for science- and technology-based sports development.

CONCLUSION

This study successfully developed technical and physical test instruments for basketball athletes aged 12–18 years. Based on content validity testing by experts, seven out of eleven technical test items and eight out of ten physical test items were found valid. Meanwhile, the reliability testing using Fleiss' Kappa demonstrated that the majority of the instruments showed very high inter-rater consistency (*almost perfect*). However, several items—such as standing long jump, hexagon test, passing test version II, dominant-hand dribbling (left and right), and speed dribbling shoot—displayed low validity and reliability,

thus requiring further refinement.

These findings confirm that most of the developed instruments are suitable for use in measuring basketball technical and physical performance. The instruments may be utilized by coaches, teachers, and sports organizations for athlete selection as well as for designing data-driven training programs that are more objective and systematic. Integrating these research outcomes into coaching practices is expected to enhance the quality of youth athlete development, strengthen technical and physical foundations from an early stage, and ultimately contribute to the improvement of basketball performance through evidence-based approaches.

REFERENCE

Abrar, I. I., & Fitroni, H. (2021). Pembinaan Cabang Olahraga Bola Basket Pada Usia Dibawah 20 Tahun Di Jawa Timur. *Jurnal Kesehatan Olahraga* , *9*(4), 145–154.

Anastasi, A., & Urbina, S. (1988). *Psychological Testing*. Upper Saddle River, NJ: Prentice Hall.

Asmawi, Mohammad, F. H. P. Yudho, I. Sina, A. Gumantan, A. Kemala, R. Iqbal, and C. R. (2022). *Desain Besar Olahraga Nasional Menuju Indonesia Emas* (1 ed., hal. 2022). Jejak Pustaka. https://www.researchgate.net/profile/Firdaus-Hendry-Yudho/publication/360096519_Desain_Besar_Olahraga_Nasional_Menuju_Indonesia_Emas/links/626205138cb84a40ac801c91/Desain-Besar-Olahraga-Nasional-Menuju-Indonesia-Emas.pdf

Balyi, I. (2008). *Canadian Basketball: Athlete Development Model*. Department of Canadian Heritage.

Balyi, I., Way, R., & Higgs, C. (2020). Long-Term Athlete Development Model. In *Long-Term Athlete Development*. https://doi.org/10.5040/9781492596318.ch001

Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., Iivonen, S., Miller, A. D., Laukkanen, A., Dudley, D., Lander, N. J., Brown, H., & Morgan, P. J. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education*, *35*(3), 219–225. https://doi.org/10.1123/jtpe.2014-0209

Bompa, T. O., & Buzzichelli, C. A. (2015). *Periodization Training for Sports, 3e*. Human Kinetics.

Cormack, S. J., Newton, R. U., McGulgan, M. R., & Doyle, T. L. A. (2008). Reliability of measures obtained during single and repeated countermovement jumps. *International Journal of Sports Physiology and Performance*, *3*(2), 131–144. https://doi.org/10.1123/ijspp.3.2.131

Drinkwater, E. J., Pyne, D. B., & McKenna, M. J. (2008). Design and interpretation of anthropometric and fitness testing of basketball players. In *Sports Medicine* (Vol. 38, Nomor 7). https://doi.org/10.2165/00007256-200838070-00004

Fox, J. L., Scanlan, A. T., & Stanton, R. (2017). A Review of Player Monitoring Approaches in Basketball: Current Trends and Future Directions. *Journal of Strength and Conditioning Research*, *31*(7), 2021–2029. https://doi.org/10.1519/JSC.0000000000001964

Hoffman, J. (2014). *Physiological Aspects of Sport Training and Performance* (Second

Edi). University of Central Florida.

Landis, J. R., & Koch, G. G. (1977). Landis amd Koch1977_agreement of categorical data. *Biometrics*, *33*(1), 159–174.

Lawshe, C. H. (1975). A quantitative approach to content validity".Personnel Psychology. *Personnel Psychology*, *28*, 563–575.

Mancha-triguero, D., García-rubio, J., Calleja-gonzález, J., & Ibáñez, S. J. (2019). *Physical fitness in basketball players : a systematic review*. *59*(0). https://doi.org/10.23736/S0022-4707.19.09180-1

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*(2), 220–240. https://doi.org/10.1123/jtpe.27.2.220

Pauole, K., Madole, K., Garhammer, J., Lacourse, M., & Rozenek, R. (2000). Reliability and validity of the T-test as a measure of agility, leg power, and leg speed in college-aged men and women Fiabilite et validite du T-test comme mesure de l ' agilite, de la puissance et de la vitesse de la jambe chez des garcons et des filles. *Journal of Strength & Conditioning Research*, *14*(4), 443–450.

Tanner, R. K., & Gore, C. J. (2013). Physiological tests for Elite Athletes 2nd Edition. In *Physiologial tests for Elite Athletes*.