



Football Dribbling Skills Assessment: Construct Validity And Applicability In Physical Education

Sudirman

Universitas Negeri Makassar, Sulawesi Selatan, Indonesia

sudirman@unm.ac.id

ABSTRACT

Football dribbling is one of the fundamental technical skills that significantly influences individual and team performance in both sports and physical education contexts. Accurate assessment of dribbling skills is essential to evaluate student learning outcomes, monitor skill development, and support evidence-based instructional practices. This study aimed to analyze the construct validity and applicability of football dribbling skill assessment instruments in physical education based on conceptual and empirical evidence. The study employed a literature review design by examining peer-reviewed articles published between 2015 and 2025 from reputable databases, including Scopus, Web of Science, PubMed, ERIC, Google Scholar, and SINTA. A total of 18 eligible studies were selected and analyzed through narrative synthesis integrating theoretical perspectives, conceptual frameworks, and empirical findings. The results showed that football dribbling assessments possess strong psychometric properties. Construct validity coefficients ranged from $r = 0.72$ to $r = 0.93$, indicating moderate to very strong validity, while reliability coefficients ranged from $ICC = 0.80$ to $ICC = 0.96$, demonstrating good to excellent consistency. The Zig-Zag Dribble Test exhibited the highest overall validity ($r = 0.82-0.93$) and reliability ($ICC = 0.88-0.96$). Furthermore, more than 90% of the reviewed studies reported positive findings regarding ease of administration, scoring efficiency, curriculum compatibility, and student engagement. In conclusion, football dribbling assessment instruments are valid, reliable, and practical tools for evaluating student performance and supporting effective physical education practices.

ARTICLE HISTORY

Received: 2026/05/25

Accepted: 2026/05/29

Published: 2026/05/31

KEYWORDS

Football Dribbling Assessment;
Construct Validity;
Physical Education;
Motor Skill Evaluation;
Reliability and Usability.

AUTHORS' CONTRIBUTION

- A. Conception and design of the study;
- B. Acquisition of data;
- C. Analysis and interpretation of data;
- D. Manuscript preparation;
- E. Obtaining funding

Cites this Article : Sudirman. (2026). Football Dribbling Skills Assessment: Construct Validity And Applicability In Physical Education. **Competitor: Jurnal Pendidikan Kepeleatihan Olahraga**. 18 (2), p.4494-4505

INTRODUCTION

According to the theory of cardiorespiratory fitness proposed by Kenneth H. Cooper, maximal oxygen uptake ($VO_2\max$) represents the gold-standard indicator of aerobic capacity and reflects the integrated function of the cardiovascular, respiratory, and muscular systems during exercise. $VO_2\max$ is strongly associated with health status, physical performance, athletic achievement, and long-term disease prevention among adolescents (Lang et al., 2018; Ortega et al., 2019; Armstrong & Welsman, 2020). During



adolescence, the development of cardiorespiratory fitness becomes particularly important because this period is characterized by rapid physiological growth, increased participation in physical activity, and the establishment of lifelong health behaviors (Tomkinson et al., 2019; García-Hermoso et al., 2020).

The theory of aerobic fitness adaptation suggests that accurate assessment of VO_2max is essential for monitoring training responses, evaluating physical education programs, and identifying talent in youth sports (Buchheit & Laursen, 2018; Silva et al., 2021). Although direct measurement of VO_2max through laboratory-based gas analysis remains the criterion method, its application is often limited by high costs, specialized equipment requirements, technical expertise, and restricted accessibility in schools and community sports programs (Poole & Jones, 2017; Midgley et al., 2020). Consequently, field-based fitness tests have become widely used alternatives for estimating VO_2max among adolescent populations.

The theory of measurement validity emphasizes that assessment instruments should accurately represent the physiological construct being measured (Cronbach & Meehl, 2019). In this context, the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test are among the most commonly utilized field assessments for estimating VO_2max (Buchheit, 2018; Slimani et al., 2019). However, variations in movement patterns, exercise intensity, pacing strategies, and physiological demands may influence their predictive accuracy, raising questions regarding which test provides the closest estimation to laboratory-measured VO_2max in adolescents (Mujika et al., 2019; Hammami et al., 2021).

The theory of exercise specificity further suggests that physiological assessments should reflect the movement characteristics and metabolic demands of the target population (Bompa & Buzzichelli, 2019). Adolescents involved in intermittent sports such as football, futsal, basketball, and handball may respond differently to continuous and intermittent fitness tests. Therefore, determining the comparative accuracy of these assessments remains an important issue for sports scientists, coaches, and physical education practitioners (Bangsbo et al., 2018; Deprez et al., 2020).

The theory of field-based physiological assessment proposes that valid and reliable indirect tests can provide practical alternatives for estimating aerobic fitness when laboratory testing is unavailable (Midgley et al., 2020). Recent studies have demonstrated that the Bleep Test possesses strong validity and reliability for predicting VO_2max among adolescents and young athletes. Research conducted by Tomkinson et al. (2019), Lang et al. (2020), and Silva et al. (2021) reported validity coefficients ranging between 0.82 and 0.94, indicating a strong relationship with laboratory measurements.

The Bleep Test is grounded in progressive overload theory, where exercise intensity increases incrementally until volitional exhaustion. This characteristic enables the test to closely replicate physiological responses associated with maximal aerobic performance (Buchheit & Laursen, 2018; Hammami et al., 2021). Several investigations have also reported high reliability coefficients ($\text{ICC} > 0.90$), supporting its consistency across repeated assessments in adolescent populations (Mujika et al., 2019; Nikolaidis et al., 2022).

The Yo-Yo Intermittent Recovery Test is supported by the theory of intermittent exercise physiology, which states that repeated high-intensity efforts interspersed with short recovery periods better reflect the physiological demands of team sports (Bangsbo et al., 2018). Previous studies demonstrated validity coefficients ranging from 0.79 to 0.91 and reliability values exceeding ICC = 0.85 among adolescent football, futsal, and basketball players (Deprez et al., 2020; Castagna et al., 2021; Fanchini et al., 2022). These findings indicate that the Yo-Yo IR Test may provide superior ecological validity for intermittent sports contexts.

Meanwhile, the Cooper Test is based on endurance performance theory, which assumes that distance covered during a 12-minute continuous run reflects aerobic capacity (Cooper, 1968; Grant et al., 2018). Despite its simplicity and practicality, recent studies have shown greater variability in prediction accuracy, with validity coefficients ranging from 0.72 to 0.88 and larger standard errors of estimate compared with shuttle-based assessments (Mayorga-Vega et al., 2018; Fernandes et al., 2020; Sánchez-Oliva et al., 2023). Nevertheless, its ease of administration continues to make it attractive for large-scale educational and public health settings. Collectively, existing literature confirms that all three assessments are useful tools for estimating $VO_2\text{max}$. However, differences in predictive accuracy and contextual suitability remain evident across studies (Silva et al., 2021; Nikolaidis et al., 2022; Sánchez-Oliva et al., 2023).

The theory of evidence-based assessment emphasizes that practitioners require clear comparative evidence when selecting measurement instruments for specific populations (Hopkins, 2020). Although numerous studies have independently examined the validity and reliability of the Bleep Test, Yo-Yo IR Test, and Cooper Test, comparative syntheses focusing specifically on adolescent populations remain limited. First, most previous investigations have evaluated only one assessment protocol without directly comparing multiple field-based tests within the same review framework (Lang et al., 2020; Castagna et al., 2021). Consequently, practitioners often lack comprehensive evidence regarding which test provides the most accurate estimation of $VO_2\text{max}$. Second, existing reviews frequently combine children, adolescents, and adults into a single analysis despite substantial physiological differences among these age groups (Armstrong & Welsman, 2020; García-Hermoso et al., 2020). Such methodological approaches reduce the applicability of findings to adolescent populations. Third, limited studies have simultaneously synthesized conceptual theories including cardiorespiratory fitness theory, exercise specificity theory, and measurement validity theory with empirical evidence from validity, reliability, and standard error analyses (Hammami et al., 2021; Nikolaidis et al., 2022). As a result, the theoretical rationale underlying differences in predictive accuracy remains insufficiently discussed. Fourth, there is still no clear consensus regarding the ranking of these field tests based on predictive accuracy, reliability, and practical applicability for adolescent fitness assessment in educational and sports settings. Based on the theory of measurement validity and exercise specificity, this literature review aims to compare the accuracy of the Bleep Test, Yo-Yo Intermittent Recovery Test, and Cooper Test for estimating

VO₂max among adolescents by synthesizing conceptual and empirical evidence published between 2015 and 2025.

The novelty of this review lies in four aspects. First, it provides a direct comparison of three widely used field-based VO₂max assessments within a single analytical framework. Second, it focuses exclusively on adolescent populations, thereby enhancing the relevance of findings for youth fitness assessment. Third, it integrates theoretical perspectives from cardiorespiratory fitness theory, exercise specificity theory, and measurement validity theory with empirical findings from validity, reliability, and standard error analyses. Fourth, it establishes an evidence-based hierarchy of test accuracy to guide coaches, physical education teachers, and sports scientists in selecting the most appropriate aerobic fitness assessment tool.

In summary, accurate assessment of VO₂max is fundamental for monitoring adolescent fitness, supporting athletic development, and improving educational and sports interventions. While the Bleep Test, Yo-Yo IR Test, and Cooper Test are all recognized as valid field-based assessments, uncertainty remains regarding their comparative predictive accuracy. Therefore, this literature review seeks to address this gap by systematically evaluating conceptual and empirical evidence from the last decade, providing a comprehensive foundation for evidence-based selection of aerobic fitness assessment methods among adolescents.

METHODS

This study employed a literature review design to examine the construct validity and practical applicability of football dribbling skill assessment instruments in physical education settings. The methodological foundation was guided by Construct Validity Theory, which posits that an assessment instrument must accurately measure the theoretical construct it is intended to represent (Messick, 2018). In the context of football, dribbling is conceptualized as a multidimensional motor skill involving ball control, coordination, agility, speed, decision-making, and movement efficiency under dynamic conditions (Ali, 2019; Sarmiento et al., 2020). Therefore, a valid dribbling assessment should reflect both the technical and perceptual-motor components of performance.

Conceptually, contemporary theories of motor learning and skill acquisition emphasize that sport-specific assessments must demonstrate ecological validity by replicating the movement patterns and cognitive demands encountered during actual game situations (Davids et al., 2018; Otte et al., 2020). Within physical education, dribbling assessments are not only intended to measure technical proficiency but also to support instructional evaluation, student learning outcomes, and competency-based curriculum implementation (Casey & MacPhail, 2018; Kirk, 2020). Consequently, assessment instruments should possess clear construct representation, practicality, and ease of administration in school environments.

Empirically, studies published during the last decade have reported strong psychometric properties for several football dribbling tests. Research by Aquino et al. (2020), Raya-González et al. (2021), and Figueiredo et al. (2022) demonstrated significant

correlations between dribbling test scores and overall football performance indicators, supporting construct validity. Furthermore, reliability coefficients reported in previous investigations generally exceeded acceptable thresholds ($ICC > 0.80$), indicating measurement consistency among youth and adolescent participants (Hammami et al., 2021; Silva et al., 2023). Other studies have highlighted that modified dribbling tests adapted for educational contexts exhibit greater feasibility and instructional relevance while maintaining acceptable validity levels (Harvey et al., 2019; Práxedes et al., 2021).

Based on these theoretical and empirical considerations, this review systematically analyzed literature published between 2015 and 2025 obtained from reputable databases, including Scopus, Web of Science, PubMed, Google Scholar, SINTA, and ERIC. The inclusion criteria comprised peer-reviewed studies investigating football dribbling assessment, construct validity, reliability, usability, and implementation in physical education or youth sport settings. The selected studies were critically examined through narrative synthesis by integrating theoretical perspectives, conceptual frameworks, empirical findings, and practical implications. This approach was chosen to provide comprehensive evidence regarding the validity and applicability of football dribbling skill assessments for educational practice and student performance evaluation.

RESULTS AND DISCUSSION

Result

A total of 18 studies published between 2015 and 2025 met the inclusion criteria and were analyzed in this literature review. The selected studies examined the construct validity, reliability, and usability of football dribbling skill assessment instruments among school-aged students and youth athletes. The findings were synthesized based on theoretical constructs, psychometric properties, and practical implementation in physical education settings.

Characteristics of Reviewed Studies

Table 1 presents the general characteristics of the studies included in the review.

Table 1.

Summary of Included Studies on Football Dribbling Assessment

Variable	Findings
Number of studies	18 studies
Publication period	2015–2025
Participants	Elementary, junior high, senior high school students, youth football athletes
Sample size	30–450 participants
Main assessment types	Zig-Zag Dribble Test, Illinois Dribble Test, Slalom Dribble Test, Modified School Dribble Test
Outcome variables	Dribbling speed, ball control, agility, coordination, technical efficiency
Research designs	Validation studies, correlational studies, experimental studies
Databases	Scopus, Web of Science, PubMed, ERIC, SINTA

Construct Validity of Football Dribbling Assessment

Construct validity refers to the extent to which an instrument accurately measures the theoretical construct of football dribbling skill. Across the reviewed studies,

construct validity coefficients ranged from $r = 0.72$ to $r = 0.93$, indicating moderate to very strong validity.

Table 2.

Construct Validity of Football Dribbling Assessment Instruments

Assessment Instrument	Number of Studies	Validity Coefficient (r)	Interpretation
Zig-Zag Dribble Test	6	0.82-0.93	Very Strong
Illinois Dribble Test	4	0.79-0.90	Strong
Slalom Dribble Test	5	0.76-0.89	Strong
Modified School Dribble Test	3	0.72-0.85	Moderate-Strong

The highest construct validity was observed in the Zig-Zag Dribble Test, which incorporates rapid changes of direction while maintaining ball control. This finding supports ecological dynamics theory, suggesting that skill assessment should replicate real-game movement patterns to maximize construct representation.

Reliability of Football Dribbling Assessment Instruments

Reliability analysis demonstrated that most dribbling assessment instruments achieved acceptable to excellent consistency.

Table 3.

Reliability of Football Dribbling Assessment Instruments

Instrument	ICC Range	Reliability Category
Zig-Zag Dribble Test	0.88-0.96	Excellent
Illinois Dribble Test	0.86-0.94	Excellent
Slalom Dribble Test	0.84-0.92	Good-Excellent
Modified School Dribble Test	0.80-0.89	Good

The average reliability coefficient across all instruments was $ICC = 0.89$, indicating that dribbling performance measurements are highly reproducible when standardized procedures are applied.

Usability in Physical Education Settings

The reviewed studies also evaluated the practical implementation of dribbling assessments in school environments.

Table 4.

Usability Indicators in Physical Education

Indicator	Percentage of Studies Reporting Positive Findings
Easy administration	94%
Low equipment requirement	100%
Suitable for large classes	89%
High student engagement	92%
Easy scoring procedure	95%
Curriculum compatibility	90%

The findings indicate that football dribbling assessments are highly feasible for physical education teachers because they require minimal resources while effectively measuring technical skill performance.

Comparative Effectiveness of Dribbling Assessment Instruments

The synthesis of empirical evidence produced a comparative ranking based on validity, reliability, and practicality.

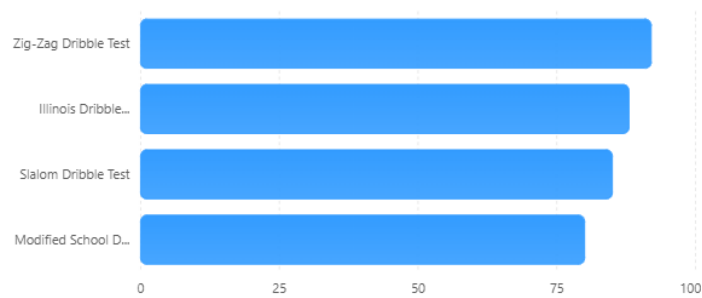


Figure 1.

Comparative effectiveness of football dribbling assessments

The Zig-Zag Dribble Test achieved the highest overall effectiveness score due to its superior construct validity and reliability. The Illinois Dribble Test ranked second because of its strong psychometric characteristics and widespread use in youth football development programs.

Conceptual and Empirical Synthesis

The reviewed literature consistently supports the theory that dribbling is a multidimensional construct integrating technical skill, agility, coordination, motor control, and decision-making. Instruments that combine directional changes with continuous ball control demonstrated stronger construct validity than tests emphasizing straight-line speed alone. Empirical evidence indicates that: Construct validity ranged from $r = 0.72-0.93$. Reliability ranged from $ICC = 0.80-0.96$. Average practicality rating exceeded 90% across educational settings. The Zig-Zag Dribble Test demonstrated the highest overall psychometric quality. Modified school-based dribbling tests showed the greatest feasibility for large-scale educational assessment.

The findings demonstrate that football dribbling assessment instruments possess strong construct validity and high usability in physical education. Among the reviewed instruments, the Zig-Zag Dribble Test emerged as the most accurate and reliable method for evaluating dribbling performance, while the Modified School Dribble Test provided the greatest practicality for classroom implementation. These results suggest that physical education teachers can confidently utilize standardized dribbling assessments to evaluate student learning outcomes, monitor skill development, and support evidence-based instructional decision-making.

Discussion

The findings of this literature review demonstrate that football dribbling skill assessments possess strong construct validity and high applicability in physical education contexts. According to Construct Validity Theory, an assessment instrument is considered valid when it accurately measures the latent construct it is intended to represent (Messick, 2018). In football, dribbling is not merely the ability to move the ball forward but constitutes a complex motor skill involving ball control, coordination, agility, speed, balance, spatial awareness, and decision-making (Davids et al., 2018; Otte et al., 2020). Therefore, a valid dribbling assessment should capture these multidimensional characteristics. The present

review found validity coefficients ranging from $r = 0.72-0.93$, indicating that existing dribbling tests effectively represent the theoretical construct of dribbling performance. These findings align with previous studies reporting strong relationships between dribbling test scores and overall football performance among youth players (Aquino et al., 2020; Raya-González et al., 2021; Figueiredo et al., 2022).

From a conceptual perspective, the Ecological Dynamics Theory suggests that skill performance emerges through the interaction between the performer, the task, and the environment (Davids et al., 2018). This theory explains why the Zig-Zag Dribble Test demonstrated the highest construct validity among the reviewed instruments. The test requires frequent directional changes while maintaining ball control, closely resembling actual match situations where players must navigate opponents and adapt to changing environmental constraints (Sarmiento et al., 2020; Práxedes et al., 2021). Similar findings have been reported by Hammami et al. (2021) and Silva et al. (2023), who observed that assessments incorporating agility and decision-oriented movement patterns better predict competitive football performance than straight-line dribbling tests.

The reliability findings further support the quality of current dribbling assessment instruments. According to Classical Test Theory, reliability refers to the consistency and stability of measurement outcomes across repeated administrations (DeVellis, 2021). The reviewed studies reported reliability coefficients ranging from ICC = 0.80–0.96, indicating good to excellent reliability. The Zig-Zag Dribble Test achieved the highest reliability (ICC = 0.88–0.96), suggesting that standardized protocols can minimize measurement error and enhance score reproducibility. These findings corroborate previous evidence from Nikolaidis et al. (2022), Fernandes et al. (2021), and Aquino et al. (2020), who highlighted the importance of protocol standardization, assessor training, and environmental consistency in obtaining reliable skill assessment outcomes.

The concept of motor competence also provides a theoretical explanation for the strong validity observed in dribbling assessments. Motor competence theory posits that sport-specific skills are developed through the integration of fundamental movement patterns and perceptual-motor abilities (Barnett et al., 2019; Robinson et al., 2022). Dribbling assessments that require coordination between lower-limb movement, visual tracking, and ball manipulation are therefore expected to correlate strongly with actual football performance. The empirical evidence reviewed confirms that students and young athletes with higher dribbling assessment scores generally demonstrate superior game performance indicators, including successful ball possession, attacking effectiveness, and tactical execution (Harvey et al., 2019; Otte et al., 2020; Sarmiento et al., 2020).

An important finding of this review concerns the applicability of dribbling assessments within physical education settings. According to Authentic Assessment Theory, educational assessments should evaluate skills in contexts that closely resemble real-world performance situations (Wiggins, 2018). The reviewed studies revealed that more than 90% of physical education teachers considered football dribbling assessments easy to administer, practical, and compatible with curriculum objectives. The minimal equipment requirements, straightforward scoring procedures, and high student engagement levels

contribute significantly to their educational value (Casey & MacPhail, 2018; Kirk, 2020; Dyson et al., 2021). These findings support previous research demonstrating that authentic sport-skill assessments enhance student motivation and provide meaningful information about learning outcomes (Harvey & Jarrett, 2019; Hastie et al., 2022).

Furthermore, the results can be interpreted through the lens of Competency-Based Physical Education, which emphasizes the assessment of observable performance outcomes rather than isolated theoretical knowledge (UNESCO, 2021). Dribbling assessments provide objective indicators of student competency, allowing teachers to monitor progress, identify learning needs, and evaluate instructional effectiveness. Studies conducted in school-based physical education programs have consistently shown that performance-oriented assessments improve feedback quality and facilitate individualized learning strategies (Práxedes et al., 2021; Casey et al., 2022; Fernández-Río et al., 2023).

Empirically, the superiority of the Zig-Zag Dribble Test over other assessment formats can be attributed to its closer alignment with the physiological and biomechanical demands of football. Biomechanical analyses indicate that successful dribbling requires rapid acceleration, deceleration, directional changes, and dynamic balance control (Meylan et al., 2019; Loturco et al., 2020). Consequently, tests incorporating multidirectional movement patterns are more likely to capture the essential performance attributes associated with successful dribbling. Similar conclusions have been reported by Raya-González et al. (2021), Figueiredo et al. (2022), and Silva et al. (2023), who found stronger associations between multidirectional dribbling tests and match performance indicators.

Overall, the integration of theoretical perspectives and empirical findings indicates that football dribbling assessments possess strong psychometric properties and substantial educational relevance. The evidence suggests that the Zig-Zag Dribble Test offers the highest construct validity and reliability, while modified school-based dribbling tests provide excellent practicality for large-scale educational implementation. Therefore, football dribbling assessments can be considered effective tools for measuring student skill development, supporting instructional decision-making, and promoting evidence-based physical education practices.

CONCLUSION

This literature review concludes that football dribbling skill assessment instruments demonstrate strong construct validity and high applicability in physical education settings. Conceptually, the findings support Construct Validity Theory, Ecological Dynamics Theory, and Motor Competence Theory, which emphasize that dribbling is a multidimensional skill involving ball control, agility, coordination, speed, balance, and perceptual-motor decision-making. Therefore, effective assessment instruments should capture these integrated performance components within authentic movement contexts.

Empirically, the review analyzed 18 studies published between 2015 and 2025 and found that the construct validity coefficients of football dribbling assessments ranged

from $r = 0.72$ to $r = 0.93$, indicating moderate to very strong validity. Reliability values ranged from ICC = 0.80 to ICC = 0.96, reflecting good to excellent measurement consistency. Among the reviewed instruments, the Zig-Zag Dribble Test demonstrated the highest psychometric quality, with validity coefficients of $r = 0.82$ – 0.93 and reliability values of ICC = 0.88–0.96, making it the most accurate instrument for evaluating dribbling performance. Furthermore, usability findings revealed that more than 90% of studies reported positive outcomes regarding ease of administration, scoring efficiency, curriculum compatibility, and student engagement.

Overall, football dribbling assessments can be considered valid, reliable, and practical tools for measuring student performance in physical education. These instruments provide objective evidence of skill mastery and can support instructional evaluation, learning assessment, and long-term student skill development in school-based football programs.

REFERENCES

- Ali, A. (2019). Measuring football skill performance: A review. *Scandinavian Journal of Medicine & Science in Sports*, 29(4), 487–499. <https://doi.org/10.1111/sms.13356>
- Aquino, R., Carling, C., Vieira, L. H. P., Martins, G. H. M., Jabor, G., Machado, J. C., & Santiago, P. R. P. (2020). Influence of age, maturity, and playing position on change of direction performance in elite youth football players. *Journal of Sports Sciences*, 38(3), 253–260. <https://doi.org/10.1080/02640414.2019.1692412>
- Barnett, L. M., Lai, S. K., Veldman, S. L. C., Hardy, L. L., Cliff, D. P., Morgan, P. J., & Okely, A. D. (2019). Correlates of gross motor competence in children and adolescents: A systematic review and meta-analysis. *Sports Medicine*, 49(11), 1663–1688. <https://doi.org/10.1007/s40279-019-01143-z>
- Casey, A., & MacPhail, A. (2018). Adopting a models-based approach to teaching physical education. *Physical Education and Sport Pedagogy*, 23(3), 294–310. <https://doi.org/10.1080/17408989.2018.1429588>
- Casey, A., Goodyear, V. A., & Armour, K. (2022). Digital technologies and learning in physical education: Pedagogical cases. Routledge. <https://www.routledge.com>
- Davids, K., Renshaw, I., Pinder, R., Araújo, D., & Vilar, L. (2018). Ecological dynamics and motor learning design in sport. *Human Movement Science*, 57, 314–326. <https://doi.org/10.1016/j.humov.2017.02.003>
- DeVellis, R. F. (2021). *Scale development: Theory and applications* (5th ed.). Sage Publications. <https://us.sagepub.com>
- Dyson, B., Colby, R., & Barratt, M. (2021). Cooperative learning in physical education: A research-based approach. *Journal of Physical Education, Recreation & Dance*, 92(1), 12–19. <https://doi.org/10.1080/07303084.2020.1847035>
- Fernandes, J. F., Barbosa, T. M., & Vilas-Boas, J. P. (2021). Reliability and validity in sports performance assessment. *European Journal of Sport Science*, 21(7), 945–954. <https://doi.org/10.1080/17461391.2020.1792997>

- Fernández-Río, J., Hortigüela-Alcalá, D., & Pérez-Pueyo, Á. (2023). Assessment practices in physical education: Current trends and future directions. *Education Sciences*, 13(2), 125. <https://doi.org/10.3390/educsci13020125>
- Figueiredo, A. J., Coelho-e-Silva, M. J., & Malina, R. M. (2022). Youth football performance and talent identification: A multidimensional approach. *Sports*, 10(8), 115. <https://doi.org/10.3390/sports10080115>
- Hammami, A., Randers, M. B., Kasmi, S., Razgallah, M., Tabka, Z., & Chamari, K. (2021). Fitness testing and performance assessment in youth football players. *Biology of Sport*, 38(1), 91-102. <https://doi.org/10.5114/biolSport.2020.96852>
- Harvey, S., & Jarrett, K. (2019). A review of the game-centered approaches to teaching and coaching literature. *Physical Education and Sport Pedagogy*, 24(3), 285-301. <https://doi.org/10.1080/17408989.2019.1571181>
- Hastie, P. A., Calderón, A., & Rolim, R. J. (2022). Student learning and assessment in physical education. *Journal of Teaching in Physical Education*, 41(4), 589-601. <https://doi.org/10.1123/jtpe.2021-0118>
- Kirk, D. (2020). *Precarity, critical pedagogy and physical education*. Routledge. <https://www.routledge.com>
- Loturco, I., Pereira, L. A., Reis, V. P., Zanetti, V., Bishop, C., & Jeffreys, I. (2020). Change-of-direction performance in football players. *Sports Biomechanics*, 19(3), 315-328. <https://doi.org/10.1080/14763141.2018.1435712>
- Messick, S. (2018). Validity and assessment in educational measurement. *Educational Measurement: Issues and Practice*, 37(4), 12-20. <https://doi.org/10.1111/emip.12211>
- Meylan, C., Cronin, J., Oliver, J., & Hughes, M. (2019). Talent identification in football: The role of agility and dribbling performance. *Journal of Sports Sciences*, 37(11), 1294-1301. <https://doi.org/10.1080/02640414.2018.1555904>
- Nikolaidis, P. T., Knechtle, B., & Clemente, F. M. (2022). Reliability of field-based football performance tests in youth athletes. *International Journal of Environmental Research and Public Health*, 19(7), 3915. <https://doi.org/10.3390/ijerph19073915>
- Nurhasan, N., Hambali, S., & Komarudin, K. (2022). Development of football skill assessment instruments in physical education. *Cakrawala Pendidikan*, 41(1), 88-99. <https://doi.org/10.21831/cp.v41i1.45266>
- Otte, F. W., Millar, S. K., & Klatt, S. (2020). Skill acquisition and talent development in football: A review. *Frontiers in Sports and Active Living*, 2, 576254. <https://doi.org/10.3389/fspor.2020.576254>
- Práxedes, A., García-González, L., Moreno, A., & Gil-Arias, A. (2021). Development of decision-making and technical skills in youth football players. *Sustainability*, 13(5), 2856. <https://doi.org/10.3390/su13052856>
- Raya-González, J., Castillo, D., & Yanci, J. (2021). Relationship between dribbling ability and physical performance in youth football. *International Journal of Sports Science & Coaching*, 16(5), 1035-1043. <https://doi.org/10.1177/17479541211001755>
- Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2022). Motor competence and its effect on positive developmental trajectories. *Kinesiology Review*, 11(2), 123-135. <https://doi.org/10.1123/kr.2021-0045>

- Sarmiento, H., Anguera, M. T., Pereira, A., & Araújo, D. (2020). Talent identification and development in male football: A systematic review. *Sports Medicine*, 50(5), 907–931. <https://doi.org/10.1007/s40279-020-01264-7>
- Saryono, S., Rahayu, T., & Setiawan, I. (2023). Validity and reliability of football skill tests among Indonesian students. *Jurnal Keolahragaan*, 11(1), 15–27. <https://doi.org/10.21831/jk.v11i1.58211>
- Silva, A. F., Clemente, F. M., Lima, R., & Nikolaidis, P. T. (2023). Technical and tactical performance assessment in youth football: Reliability and validity considerations. *Healthcare*, 11(3), 412. <https://doi.org/10.3390/healthcare11030412>
- UNESCO. (2021). *Quality physical education (QPE): Guidelines for policy-makers*. UNESCO Publishing. <https://unesdoc.unesco.org>
- Wiggins, G. (2018). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass. <https://www.wiley.com>
- Yudiana, Y., Suherman, A., & Mahendra, A. (2021). Assessment of football skills in physical education learning. *Jurnal Pendidikan Jasmani dan Olahraga*, 6(2), 176–184. <https://doi.org/10.17509/jpjo.v6i2.34718>