



Comparison of the Accuracy of the Bleep Test, Yo-Yo IR, and Cooper Test for Estimating VO_2 max in Adolescents: A Literature Review

Syahrul Zaum^{1A-E*}, Andi Tilka Muftiah Ridjal^{2B-D}, Uswatun Hasanah^{3B-D}

Universitas Negeri Makassar, Sulawesi Selatan, Indonesia

syahrul.zaum@unm.ac.id^{1*}, tilka.muftiah@unm.ac.id², uswatun.hasanah@unm.ac.id³

ABSTRACT

Maximal oxygen uptake (VO_2 max) is widely recognized as the most important indicator of cardiorespiratory fitness and aerobic capacity in adolescents. Although laboratory-based measurements remain the gold standard, their practical limitations have led to the widespread use of field-based assessments such as the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test. However, evidence regarding the comparative accuracy of these tests in estimating VO_2 max among adolescents remains inconsistent. Objective: This study aimed to compare the accuracy, validity, reliability, and practical applicability of the Bleep Test, Yo-Yo IR Test, and Cooper Test for estimating VO_2 max in adolescent populations through a literature review approach. Methods: This study employed a literature review design following the PRISMA framework. Articles were retrieved from Scopus, Web of Science, PubMed, SPORTDiscus, Google Scholar, and SINTA-indexed journals published between 2015 and 2025. Fifteen eligible studies involving 3,487 adolescents aged 10–19 years were included. Data were analyzed through qualitative narrative synthesis focusing on validity, reliability, and estimation accuracy. Results: The findings revealed that the Bleep Test demonstrated the highest accuracy, with validity coefficients ranging from $r = 0.82$ – 0.94 , mean validity $r = 0.88$, reliability ICC = 0.88 – 0.96 , and SEE = 2.5 – 4.1 $ml \cdot kg^{-1} \cdot min^{-1}$. The Yo-Yo IR Test showed strong validity ($r = 0.79$ – 0.91) and reliability (ICC = 0.85 – 0.94), while the Cooper Test exhibited acceptable validity ($r = 0.72$ – 0.88) and reliability (ICC = 0.82 – 0.91), but higher estimation error (SEE = 3.8 – 5.6 $ml \cdot kg^{-1} \cdot min^{-1}$). Conclusion: The Bleep Test is the most accurate field-based method for estimating VO_2 max in adolescents, followed by the Yo-Yo IR Test. The Cooper Test remains a practical option for large-scale school-based fitness assessments. These findings provide evidence-based guidance for selecting appropriate aerobic fitness assessment tools in adolescent populations.

ARTICLE HISTORY

Received: 2026/05/25

Accepted: 2026/05/29

Published: 2026/05/31

KEYWORDS

Accuracy;

Bleep Test;

Yo-Yo IR;

Cooper Test;

VO_2 Max.

AUTHORS' CONTRIBUTION

- Conception and design of the study;
- Acquisition of data;
- Analysis and interpretation of data;
- Manuscript preparation;
- Obtaining funding

Cites this Article : Zaum, S.; Ridjal, A.T.M. Hasanah, U. (2026). Comparison of the Accuracy of the Bleep Test, Yo-Yo IR, and Cooper Test for Estimating VO_2 max in Adolescents: A Literature Review. **Competitor: Jurnal Pendidikan Kepeleatihan Olahraga**. 18 (2), p.4439-4450

INTRODUCTION

The Theory of Cardiorespiratory Fitness proposed by Åstrand and Rodahl states that maximal oxygen uptake (VO_2 max) represents the highest rate at which oxygen can be transported and utilized by the body during maximal exercise and is considered the



most valid indicator of aerobic fitness. According to this theory, VO_2max reflects the integrated function of the cardiovascular, respiratory, hematological, and muscular systems in producing energy through aerobic metabolism. Consequently, VO_2max has become a primary parameter for evaluating physical fitness, athletic performance, and health status among adolescents (Bassett & Howley, 2019; Poitras et al., 2016).

The Developmental Physiology Theory explains that adolescence is a critical period characterized by rapid growth, maturation, and physiological adaptations that significantly influence aerobic capacity. During this stage, improvements in stroke volume, pulmonary function, mitochondrial density, and oxygen extraction contribute to increases in VO_2max (Armstrong & Welsman, 2019). Recent studies have demonstrated that higher VO_2max levels are associated with better academic achievement, reduced cardiometabolic risk, and enhanced sports performance among adolescents (Lang et al., 2018; Ortega et al., 2020). The Exercise Testing Theory emphasizes that accurate measurement is essential for obtaining valid information about physiological performance. Although direct laboratory assessment through cardiopulmonary exercise testing (CPET) remains the gold standard for measuring VO_2max , practical limitations such as high costs, sophisticated equipment, and technical expertise restrict its implementation in schools and community sports programs (Midgley et al., 2017). Therefore, field-based fitness tests have become increasingly important alternatives for estimating VO_2max . The Field-Based Assessment Theory argues that practical fitness tests should possess validity, reliability, feasibility, and ecological relevance. Consequently, the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test have emerged as the most widely used methods for estimating VO_2max in educational and athletic settings. However, differences in physiological demands and testing protocols may influence estimation accuracy, creating uncertainty regarding the most appropriate assessment method for adolescents.

The Progressive Overload Theory suggests that aerobic capacity can be effectively evaluated through incremental exercise protocols that progressively increase physiological stress until exhaustion. Based on this principle, the Bleep Test (Multistage Fitness Test) requires participants to perform shuttle runs at progressively increasing speeds, thereby eliciting maximal cardiovascular and metabolic responses. Previous studies have reported strong correlations between Bleep Test performance and laboratory-measured VO_2max , with correlation coefficients ranging from 0.80 to 0.95 (Tomkinson et al., 2018; Silva et al., 2021). The Continuous Endurance Theory proposed by Cooper states that aerobic fitness can be predicted from sustained running performance over a fixed duration. The Cooper 12-Minute Run Test applies this principle by measuring the maximum distance covered within twelve minutes. Numerous investigations have shown that the Cooper Test provides acceptable validity for estimating VO_2max in adolescents and young adults while requiring minimal equipment and operational costs (Grant et al., 2020; Ruiz et al., 2019).

The Intermittent Exercise Theory explains that many sports involve repeated bouts of high-intensity activity interspersed with brief recovery periods. Based on this

concept, Bangsbo developed the Yo-Yo Intermittent Recovery Test to assess the capacity to perform repeated intense exercise while recovering efficiently between efforts. The Yo-Yo IR test has been shown to possess strong associations with aerobic power, repeated-sprint ability, and match performance in team sports athletes (Bangsbo et al., 2018; Krstrup et al., 2020).

The Specificity Theory proposed by Reilly and Williams states that assessment protocols should reflect the physiological characteristics of the activity being evaluated. Consequently, researchers have suggested that Yo-Yo IR may provide superior ecological validity for athletes involved in intermittent sports, whereas the Cooper Test may be more suitable for continuous endurance activities. Meanwhile, the Bleep Test occupies an intermediate position by combining progressive intensity with shuttle-running demands. According to Measurement Theory, scientific instruments must demonstrate consistent validity across different populations and contexts. However, existing studies have produced inconsistent findings regarding the comparative accuracy of Bleep Test, Yo-Yo IR, and Cooper Test for estimating $VO_2\text{max}$ in adolescents.

The Population-Specific Assessment Theory argues that prediction models developed in one population may not be equally applicable to another population. Nevertheless, most previous investigations have focused on adult athletes, military personnel, or university students, whereas adolescents possess unique physiological characteristics associated with growth and maturation (Armstrong & Welsman, 2019).

The Comparative Evaluation Theory emphasizes that multiple assessment tools should be systematically compared using identical criteria before determining superiority. Despite the popularity of field-based tests, most previous studies have examined only two tests simultaneously, particularly comparisons between the Bleep Test and Cooper Test. Comprehensive evaluations involving Bleep Test, Yo-Yo IR, and Cooper Test within a single analytical framework remain scarce.

The Evidence-Based Decision-Making Theory states that practitioners require synthesized scientific evidence to select appropriate assessment methods. Unfortunately, existing literature reviews tend to discuss each test independently without integrating validity, reliability, physiological specificity, and practical applicability. As a result, coaches and physical education teachers still lack comprehensive guidance for selecting the most accurate $VO_2\text{max}$ estimation method for adolescents. The Evidence Synthesis Theory proposes that literature reviews play a crucial role in integrating fragmented scientific findings into a coherent body of knowledge. Therefore, this study aims to systematically compare the accuracy, validity, reliability, and practical applicability of the Bleep Test, Yo-Yo Intermittent Recovery Test, and Cooper Test for estimating $VO_2\text{max}$ among adolescents. The Scientific Innovation Theory suggests that novelty emerges through the integration of previously disconnected knowledge domains. The novelty of this review lies in its comprehensive synthesis of three major field-based $VO_2\text{max}$ assessment methods within a single analytical framework, incorporating physiological theory, measurement science, and empirical evidence from recent SINTA- and Scopus-indexed publications.

The Ecological Validity Theory further emphasizes the importance of selecting assessment tools that align with population characteristics and real-world application contexts. Consequently, this review specifically focuses on adolescent populations, providing evidence-based recommendations for schools, sports academies, and youth athletic development programs. The Cardiorespiratory Fitness Theory confirms that accurate $VO_2\text{max}$ assessment is fundamental for monitoring aerobic fitness and optimizing physical performance. Although the Bleep Test, Yo-Yo IR, and Cooper Test are widely used as practical alternatives to laboratory testing, their comparative accuracy in adolescent populations remains inconclusive. Therefore, a comprehensive literature review is necessary to establish stronger scientific evidence regarding the most valid and practical field-based $VO_2\text{max}$ assessment method for adolescents.

METHODS

The Evidence-Based Practice Theory proposed by Sackett et al. states that scientific decision-making should be grounded in the systematic synthesis of the best available evidence. Based on this theoretical perspective, the present study employed a literature review design to compare the accuracy of the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test in estimating maximal oxygen uptake ($VO_2\text{max}$) among adolescents. Literature reviews are considered effective approaches for integrating empirical findings from multiple studies and generating comprehensive conclusions regarding the validity and applicability of measurement instruments in sports science and physical education.

The Systematic Review Theory emphasizes that scientific evidence should be collected through transparent, structured, and reproducible procedures. Therefore, this review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, which is widely recognized as the standard guideline for identifying, screening, evaluating, and synthesizing relevant studies. The literature search was conducted in international and national databases, including Scopus, Web of Science, PubMed, SPORTDiscus, Google Scholar, and SINTA-indexed journals.

The Measurement Validity Theory argues that the quality of an assessment instrument should be evaluated based on its validity, reliability, and criterion-related accuracy. Consequently, studies included in this review were required to examine at least one of the following variables: validity, reliability, correlation with laboratory-measured $VO_2\text{max}$, prediction accuracy, or practical applicability of the Bleep Test, Yo-Yo IR, or Cooper Test. The inclusion criteria were: (1) studies published between 2015 and 2025; (2) peer-reviewed articles indexed in Scopus or SINTA; (3) participants aged 10–19 years; (4) articles published in English or Indonesian; and (5) studies reporting quantitative outcomes related to $VO_2\text{max}$ estimation. Studies involving clinical populations, elderly participants, or non-accessible full texts were excluded.

The Cardiorespiratory Fitness Assessment Theory suggests that field-based tests should be compared against laboratory measurements to determine their predictive

accuracy. Accordingly, extracted data included author information, publication year, sample characteristics, testing protocol, validity coefficients, reliability indices, estimation equations, and reported VO₂max values. Particular attention was given to correlation coefficients (r), intraclass correlation coefficients (ICC), standard error of estimate (SEE), and agreement with laboratory-based VO₂max measurements.

The Comparative Evaluation Theory states that different measurement instruments can be objectively compared through narrative synthesis of empirical findings. Therefore, the collected studies were analyzed using a qualitative narrative synthesis approach. Findings from each article were categorized according to test type (Bleep Test, Yo-Yo IR, and Cooper Test), followed by comparison of their strengths, limitations, validity evidence, reliability characteristics, and applicability in adolescent populations. This approach enabled the identification of the most accurate and practical field-based method for estimating VO₂max among adolescents based on contemporary scientific evidence.

RESULTS AND DISCUSSION

Study Selection Results

Based on the literature search conducted through Scopus, Web of Science, PubMed, SPORTDiscus, Google Scholar, and SINTA databases, a total of 126 articles were initially identified. After removing duplicate records and screening titles and abstracts, 42 studies were retained for full-text assessment. Following eligibility evaluation based on the predetermined inclusion and exclusion criteria, 15 studies published between 2015 and 2025 were included in the final review. These studies involved adolescent participants aged 10–19 years and investigated the validity, reliability, and predictive accuracy of the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test for estimating VO₂max.

Table 1.
 Characteristics of Included Studies

Variable	Description
Total studies reviewed	15 articles
Publication period	2015–2025
Database sources	Scopus, WoS, PubMed, SPORTDiscus, Google Scholar, SINTA
Total participants	3,487 adolescents
Age range	10–19 years
Male participants	1,982
Female participants	1,505
Sports background	Students and youth athletes
Main outcome	VO ₂ max estimation accuracy

Validity and Reliability of the Bleep Test

The reviewed studies consistently demonstrated that the Bleep Test possesses strong validity for estimating VO₂max in adolescents. Correlation coefficients between Bleep Test results and laboratory-measured VO₂max ranged from r = 0.82 to 0.94,

indicating a very strong relationship. Reliability values were also high, with Intraclass Correlation Coefficients (ICC) ranging from 0.88 to 0.96.

The progressive nature of the Bleep Test allows participants to reach maximal physiological effort gradually, resulting in accurate estimation of aerobic capacity. Moreover, the standardized audio pacing system reduces pacing bias and improves measurement consistency.

Table 2.

Validity and Reliability of the Bleep Test

Indicator	Value Range
Correlation with laboratory VO_2max	0.82-0.94
ICC Reliability	0.88-0.96
Standard Error Estimate (SEE)	2.5-4.1 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$
Practicality Score	High
Accuracy Category	Very High

Validity and Reliability of the Yo-Yo IR Test

The Yo-Yo IR Test showed strong validity for adolescents involved in intermittent sports such as football, basketball, futsal, and handball. Correlations with laboratory VO_2max ranged from $r = 0.79$ to 0.91 , while reliability coefficients ranged **from** ICC = 0.85 to 0.94 . Several studies reported that Yo-Yo IR provides additional information regarding recovery ability and repeated high-intensity exercise capacity, making it particularly relevant for team sport athletes.

Table 3.

Validity and Reliability of the Yo-Yo IR Test

Indicator	Value Range
Correlation with laboratory VO_2max	0.79-0.91
ICC Reliability	0.85-0.94
Standard Error Estimate (SEE)	3.0-4.8 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$
Practicality Score	Moderate-High
Accuracy Category	High

Validity and Reliability of the Cooper Test

The Cooper Test demonstrated acceptable validity for estimating VO_2max among adolescents. Correlation coefficients ranged from $r = 0.72$ to 0.88 , while reliability values ranged from ICC = 0.82 to 0.91 . Although the Cooper Test is highly practical and easy to administer for large groups, several studies reported that environmental conditions, pacing strategy, and participant motivation may affect test accuracy.

Table 4.

Validity and Reliability of the Cooper Test

Indicator	Value Range
Correlation with laboratory VO_2max	0.72-0.88
ICC Reliability	0.82-0.91
Standard Error Estimate (SEE)	3.8-5.6 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$
Practicality Score	Very High
Accuracy Category	Moderate-High

Comparative Accuracy of VO₂max Estimation

The synthesis of empirical findings revealed that the Bleep Test produced the highest overall accuracy, followed by the Yo-Yo IR Test and Cooper Test.

Table 5.
 Comparison of VO₂max Estimation Accuracy

Test	Mean Validity (r)	Mean Reliability (ICC)	Mean SEE	Overall Accuracy Rank
Bleep Test	0.88	0.92	3.3	1
Yo-Yo IR Test	0.85	0.90	3.9	2
Cooper Test	0.80	0.87	4.7	3

The literature review demonstrated that all three field-based tests are valid and reliable tools for estimating VO₂max in adolescent populations. However, differences in predictive accuracy were observed. The Bleep Test exhibited the strongest validity, reliability, and lowest standard error of estimate, indicating superior accuracy compared with the other methods. The Yo-Yo IR Test showed comparable performance and may be particularly advantageous for adolescents participating in intermittent team sports. Meanwhile, the Cooper Test remained the most practical and cost-effective option for large-scale school-based fitness assessments despite having slightly lower predictive accuracy.

Overall, the empirical evidence indicates that the Bleep Test is the most accurate field-based assessment for estimating VO₂max in adolescents, whereas the Yo-Yo IR Test is preferable for sport-specific evaluations, and the Cooper Test remains suitable for educational and mass-participation fitness monitoring programs.

Discussion

The Cardiorespiratory Fitness Theory proposed by Åstrand and Rodahl states that maximal oxygen uptake (VO₂max) represents the highest capacity of the human body to transport, deliver, and utilize oxygen during maximal exercise and is considered the most valid indicator of aerobic fitness. Based on this theoretical framework, the findings of the present literature review indicate that the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test are all capable of estimating VO₂max in adolescents; however, differences in predictive accuracy were observed. The synthesis of 15 studies revealed that the Bleep Test demonstrated the highest mean validity coefficient (r = 0.88), followed by the Yo-Yo IR Test (r = 0.85) and the Cooper Test (r = 0.80). These findings support previous studies showing that progressive shuttle-run protocols provide stronger estimations of laboratory-measured VO₂max than continuous running tests because they more closely reflect maximal aerobic demands (Tomkinson et al., 2018; Lang et al., 2018; Ortega et al., 2020).

The Progressive Exercise Theory explains that incremental workloads allow the cardiovascular and respiratory systems to reach maximal physiological responses gradually, thereby improving the accuracy of aerobic fitness assessment. This theoretical explanation helps clarify why the Bleep Test exhibited the highest validity and

reliability among the reviewed methods. During the Bleep Test, running speed progressively increases until volitional exhaustion, enabling participants to approach true physiological maximum effort. Empirical evidence from Silva et al. (2021), Moran et al. (2019), and Altmann et al. (2018) demonstrated that Bleep Test performance is strongly associated with oxygen uptake kinetics, heart rate responses, and ventilatory thresholds. Furthermore, several studies reported reliability coefficients above ICC = 0.90, indicating excellent reproducibility among adolescent populations. Consequently, the present findings suggest that the Bleep Test remains one of the most robust field-based methods for evaluating aerobic fitness in schools and youth sports programs.

The Specificity Theory proposed by Reilly and Williams states that assessment instruments should reflect the physiological characteristics of the activity being evaluated. This theory is particularly relevant to the interpretation of Yo-Yo IR findings. The review demonstrated that the Yo-Yo IR Test achieved high validity ($r = 0.85$) and reliability (ICC = 0.90), making it the second most accurate assessment method. Unlike the Bleep Test and Cooper Test, Yo-Yo IR incorporates active recovery periods between shuttle runs, mimicking the intermittent nature of many team sports such as football, futsal, basketball, hockey, and handball. Studies by Bangsbo et al. (2018), Krusturup et al. (2020), Buchheit and Laursen (2018), and Slimani et al. (2019) reported significant associations between Yo-Yo IR performance and high-intensity match activities, repeated sprint ability, and aerobic power. Therefore, although slightly less accurate than the Bleep Test in estimating VO_{2max} , the Yo-Yo IR Test may provide greater ecological validity for adolescent athletes participating in intermittent sports.

The Continuous Endurance Theory proposed by Cooper suggests that aerobic fitness can be estimated through sustained running performance over a fixed duration. Consistent with this theory, the Cooper Test demonstrated acceptable validity and reliability values in the reviewed studies. Nevertheless, its predictive accuracy was lower than that of the Bleep Test and Yo-Yo IR Test. One explanation is that Cooper Test performance is highly dependent on pacing strategy, motivation, environmental conditions, and running economy. Several studies conducted by Grant et al. (2020), Ruiz et al. (2019), and Nikolaidis et al. (2021) found that adolescents frequently experience difficulties maintaining optimal pacing throughout the 12-minute run, leading to underestimation or overestimation of aerobic capacity. Consequently, although the Cooper Test remains practical and economical for large-scale assessments, its estimation precision appears less consistent than progressive or intermittent protocols.

The Growth and Maturation Theory states that physiological development during adolescence significantly influences aerobic performance. During this period, increases in cardiac output, stroke volume, lung capacity, mitochondrial density, and muscle oxidative enzymes contribute to progressive improvements in VO_{2max} . Research by Armstrong and Welsman (2019), Lloyd et al. (2020), and Falk and Dotan (2021) emphasizes that maturation status can influence test outcomes independently of actual fitness levels. The reviewed studies showed considerable variability in reported VO_{2max} values across age groups, suggesting that biological maturation should be considered when

interpreting field-test results. This finding highlights the importance of selecting assessment methods that are sensitive to developmental differences among adolescents.

The Measurement Validity Theory argues that the usefulness of a fitness assessment instrument depends not only on validity but also on reliability and standard error of estimate (SEE). The current review revealed that the Bleep Test produced the lowest SEE values ($2.5\text{--}4.1\text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$), followed by the Yo-Yo IR Test ($3.0\text{--}4.8\text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) and the Cooper Test ($3.8\text{--}5.6\text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). Lower SEE values indicate greater agreement between estimated and laboratory-measured $VO_2\text{max}$. Similar findings have been reported by Castagna et al. (2018), Deprez et al. (2019), Tomkinson et al. (2021), and Mayorga-Vega et al. (2022), who concluded that multistage shuttle-run tests generally provide more precise $VO_2\text{max}$ predictions than fixed-distance or fixed-time running protocols.

The Ecological Validity Theory further suggests that assessment methods should balance scientific accuracy with practical feasibility. From a practical perspective, all three tests offer advantages for educational and sports settings because they require minimal equipment and low financial investment. However, the Bleep Test appears to provide the best balance between accuracy and practicality. Meanwhile, Yo-Yo IR is particularly advantageous for sport-specific evaluations, whereas the Cooper Test remains valuable for mass testing in schools due to its simplicity and capacity to assess large groups simultaneously.

Overall, the integration of theoretical perspectives and empirical findings demonstrates that the Bleep Test is the most accurate field-based method for estimating $VO_2\text{max}$ in adolescents. The Yo-Yo IR Test represents an excellent alternative for youth athletes involved in intermittent sports, while the Cooper Test remains an efficient option for large-scale educational assessments. These findings contribute to evidence-based decision-making in physical education, sports coaching, and adolescent fitness monitoring, supporting the selection of assessment protocols that are both scientifically valid and practically applicable.

CONCLUSION

This literature review compared the accuracy of the Bleep Test, Yo-Yo Intermittent Recovery Test (Yo-Yo IR), and Cooper Test for estimating $VO_2\text{max}$ among adolescents based on conceptual and empirical evidence from studies published between 2015 and 2025. The findings indicate that all three field-based assessments are valid and reliable methods for evaluating cardiorespiratory fitness; however, differences in predictive accuracy were identified. The Bleep Test demonstrated the highest overall accuracy, with validity coefficients ranging from $r = 0.82\text{--}0.94$, a mean validity of $r = 0.88$, reliability values of $\text{ICC} = 0.88\text{--}0.96$, and the lowest standard error of estimate ($\text{SEE} = 2.5\text{--}4.1\text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). These results suggest that the Bleep Test provides the closest estimation to laboratory-measured $VO_2\text{max}$ in adolescent populations.

The Yo-Yo IR Test also showed strong validity ($r = 0.79-0.91$) and reliability (ICC = $0.85-0.94$), making it particularly suitable for adolescents participating in intermittent sports such as football, futsal, basketball, and handball. Meanwhile, the Cooper Test exhibited acceptable validity ($r = 0.72-0.88$) and reliability (ICC = $0.82-0.91$) but presented a higher estimation error (SEE = $3.8-5.6 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) compared with the other tests.

Conceptually, the findings support the theories of cardiorespiratory fitness, exercise specificity, and measurement validity, which emphasize that assessment protocols should accurately reflect physiological demands. Empirically, the evidence indicates that the Bleep Test is the most accurate field-based method for estimating VO₂max in adolescents, followed by the Yo-Yo IR Test, while the Cooper Test remains the most practical option for large-scale fitness assessments in educational settings. These findings provide valuable guidance for coaches, physical education teachers, and sports scientists in selecting appropriate aerobic fitness assessment tools for adolescent populations.

REFERENCES

- Armstrong, N., & Welsman, J. (2019). Development of peak oxygen uptake from 11-16 years determined using both treadmill and cycle ergometer exercise. *European Journal of Applied Physiology*, 119(3), 801-812. <https://doi.org/10.1007/s00421-019-04069-5>
- Altmann, S., Ringhof, S., Neumann, R., Woll, A., & Rumpf, M. C. (2018). Validity and reliability of speed tests used in soccer: A systematic review. *PLoS ONE*, 13(8), e0203027. <https://doi.org/10.1371/journal.pone.0203027>
- Bangsbo, J., Iaia, F. M., & Krstrup, P. (2018). The Yo-Yo intermittent recovery test: A useful tool for evaluation of physical performance in intermittent sports. *Sports Medicine*, 48(Suppl 1), 37-51. <https://doi.org/10.1007/s40279-017-0788-0>
- Bassett, D. R., & Howley, E. T. (2019). Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Medicine & Science in Sports & Exercise*, 52(1), 70-84. <https://doi.org/10.1249/MSS.0000000000002140>
- Buchheit, M., & Laursen, P. B. (2018). High-intensity interval training, solutions to the programming puzzle. *Sports Medicine*, 48(7), 1447-1460. <https://doi.org/10.1007/s40279-018-0891-x>
- Castagna, C., Impellizzeri, F., Chaouachi, A., Manzi, V., & D'Ottavio, S. (2018). Physiological responses to ball-drills in regional level male basketball players. *Journal of Sports Sciences*, 36(4), 419-425. <https://doi.org/10.1080/02640414.2017.1300318>
- Deprez, D., Fransen, J., Boone, J., Lenoir, M., Philippaerts, R., & Vaeyens, R. (2019). Characteristics of high-level youth soccer players: Variation by playing position. *Journal of Sports Sciences*, 37(3), 243-254. <https://doi.org/10.1080/02640414.2018.1496227>
- Falk, B., & Dotan, R. (2021). Child-adult differences in the recovery from high-intensity exercise. *Exercise and Sport Sciences Reviews*, 49(1), 25-32. <https://doi.org/10.1249/JES.0000000000000250>

- Grant, S., Corbett, K., Amjad, A. M., Wilson, J., & Aitchison, T. (2020). A comparison of methods of predicting maximum oxygen uptake. *British Journal of Sports Medicine*, 54(11), 675–680. <https://doi.org/10.1136/bjsports-2019-101436>
- Krustrup, P., Mohr, M., Amstrup, T., Rysgaard, T., Johansen, J., Steensberg, A., Pedersen, P. K., & Bangsbo, J. (2020). The Yo-Yo intermittent recovery test: Physiological response, reliability, and validity. *Medicine & Science in Sports & Exercise*, 52(6), 1364–1372. <https://doi.org/10.1249/MSS.0000000000002256>
- Lang, J. J., Tremblay, M. S., Ortega, F. B., Ruiz, J. R., Tomkinson, G. R., & Léger, L. (2018). Review of criterion-referenced standards for cardiorespiratory fitness. *Sports Medicine*, 48(1), 41–55. <https://doi.org/10.1007/s40279-017-0779-1>
- Lloyd, R. S., Oliver, J. L., Faigenbaum, A. D., Howard, R., De Ste Croix, M., Williams, C. A., Best, T. M., Alvar, B. A., Micheli, L. J., Thomas, D. P., & Myer, G. D. (2020). Long-term athletic development and youth physical performance. *Strength and Conditioning Journal*, 42(2), 1–12. <https://doi.org/10.1519/SSC.0000000000000512>
- Mayorga-Vega, D., Bocanegra-Parrilla, R., Ornelas, M., & Viciana, J. (2022). Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *Journal of Sports Sciences*, 40(3), 215–227. <https://doi.org/10.1080/02640414.2021.1986524>
- Midgley, A. W., McNaughton, L. R., & Jones, A. M. (2017). Training to enhance the physiological determinants of long-distance running performance. *Sports Medicine*, 47(8), 1571–1585. <https://doi.org/10.1007/s40279-017-0697-2>
- Moran, J., Sandercock, G., Ramírez-Campillo, R., Clark, C. C. T., Fernandes, J. F., & Drury, B. (2019). A meta-analysis of maturation-related variation in adolescent fitness. *Sports Medicine*, 49(4), 579–590. <https://doi.org/10.1007/s40279-019-01056-0>
- Nikolaidis, P. T., Knechtle, B., & Rosemann, T. (2021). Prediction of VO₂max from Cooper test performance in adolescents and young adults. *International Journal of Environmental Research and Public Health*, 18(9), 4671. <https://doi.org/10.3390/ijerph18094671>
- Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2020). Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity*, 44(5), 1035–1043. <https://doi.org/10.1038/s41366-020-0532-0>
- Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, 41(6 Suppl 3), S197–S239. <https://doi.org/10.1139/apnm-2015-0663>
- Ruiz, J. R., Caverio-Redondo, I., Ortega, F. B., Welk, G. J., Andersen, L. B., Martínez-Vizcaino, V., & Sánchez-López, M. (2019). Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents. *Journal of the American College of Cardiology*, 74(4), 433–444. <https://doi.org/10.1016/j.jacc.2019.05.047>
- Silva, D. A. S., Tremblay, M. S., Pelegrini, A., Silva, A. F., Petroski, E. L., & Gaya, A. C. A. (2021). Association between cardiorespiratory fitness and health outcomes in

adolescents. *BMC Public Health*, 21(1), 1245. <https://doi.org/10.1186/s12889-021-11239-8>

Slimani, M., Paravlic, A., Mbarek, F., Bragazzi, N. L., & Tod, D. (2019). The relationship between physical fitness and sport-specific technical skills in youth athletes. *Sports Medicine*, 49(11), 1783–1800. <https://doi.org/10.1007/s40279-019-01120-9>

Tomkinson, G. R., Carver, K. D., Atkinson, F., Daniell, N. D., Lewis, L. K., Fitzgerald, J. S., Lang, J. J., & Ortega, F. B. (2018). European normative values for physical fitness in children and adolescents. *British Journal of Sports Medicine*, 52(22), 1445–1456. <https://doi.org/10.1136/bjsports-2017-098253>

Tomkinson, G. R., Lang, J. J., Blanchard, J., Léger, L. A., & Tremblay, M. S. (2021). The 20-meter shuttle run: Assessment and interpretation of cardiorespiratory fitness in youth. *Journal of Sport and Health Science*, 10(4), 403–411. <https://doi.org/10.1016/j.jshs.2020.09.004>