Construction and standardization of a physical and physiological fitness test battery for students of colleges of Physical Education and Sport Sciences

Construcción y estandarización de una batería de pruebas de aptitud física y fisiológica para estudiantes de las facultades de Educación Física y Ciencias Del Deporte

Authors

Alhassan Emad Aldin[1]
Shakir Mahmood Abdullah[2]

[1, 2] University of Anbar (Iraq)

Corresponding author:
Othman Abed
ofcbqwer@gmail.com

How to cite in APA

Emad Aldin, A., & Mahmoud Abdullah, S. (2025). Construction and standardization of a physical and physiological fitness test battery for students of colleges of Physical Education and Sport Sciences. *Retos*, *70*, 1481–1495. https://doi.org/10.47197/retos.v70.116885

Abstract

Objective: This study aimed to construct and standardize a physical and physiological fitness test battery specifically designed for students of Colleges of Physical Education and Sport Sciences in the Arab region, thereby addressing a gap in local assessment tools.
Methodology: A multi-stage research methodology was employed, beginning with the nomination of a set of preliminary tests and the validation of their content by 16 experts. Subsequently, the psychometric properties of the approved tests (discriminant validity, reliability, and objectivity) were verified using a construction sample of 30 students. Finally, the definitive battery was administered to a standardization sample of 490 students to establish normative performance standards using Z-scores and T-scores.
Results: The results showed that the selected tests possessed high psychometric properties. They revealed significant variability in the students' physical and physiological fitness levels, confirming the need for a precise measurement tool. Performance norms were successfully established for components such as explosive power, flexibility, muscular endurance, and speed, in addition to vital physiological indicators.
Discussion and Conclusions: This test battery serves as a powerful diagnostic tool that contributes to an objective understanding of students' physical capabilities. It provides educators with a practical framework to identify strengths and weaknesses and to design more effective, individualized training programs. It is recommended that these norms be used as a primary reference for student assessment and guidance, thereby enhancing the quality of sports education and professional readiness in the region.

Keywords

Physical fitness; physiological fitness; test battery; standardization; Physical Education students.

Resumen

Objetivo: Este estudio tuvo como objetivo construir y estandarizar una batería de pruebas de aptitud física y fisiológica diseñada específicamente para estudiantes de Facultades de Educación Física y Ciencias del Deporte en la región árabe, con el fin de abordar una brecha en las herramientas de evaluación locales.
Metodología: Se empleó una metodología de investigación de múltiples etapas, comenzando con la nominación de un conjunto de pruebas preliminares y la validación de su contenido por 16 expertos. Posteriormente, se verificaron las propiedades psicométricas de las pruebas aprobadas (validez discriminante, fiabilidad y objetividad) utilizando una muestra de construcción de 30 estudiantes. Finalmente, la batería definitiva se administró a una muestra de estandarización de 490 estudiantes para establecer estándares de rendimiento normativos utilizando las puntuaciones Z y T.
Resultados: Los resultados mostraron que las pruebas seleccionadas poseían altas propiedades psicométricas. Revelaron una variabilidad significativa en los niveles de aptitud física y fisiológica de los estudiantes, lo que confirma la necesidad de una herramienta de medición precisa. Se establecieron con éxito normas de rendimiento para componentes como la potencia explosiva, la flexibilidad, la resistencia muscular y la velocidad, además de indicadores fisiológicos vitales.
Discusión y Conclusiones: Esta batería de pruebas sirve como una poderosa herramienta de diagnóstico que contribuye a una comprensión objetiva de las capacidades físicas de los estudiantes. Proporciona a los educadores un marco práctico para identificar fortalezas y debilidades y para diseñar programas de entrenamiento individualizados más efectivos. Se recomienda que estas normas se utilicen como referencia principal para la evaluación y orientación de los estudiantes, mejorando así la calidad de la educación deportiva y la preparación profesional en la región.

Palabras clave

Aptitud física; aptitud fisiológica; batería de prueba; normalización; estudiantes de Educación Física.

Introduction

Physical and physiological fitness are fundamental pillars of individual performance and quality of life. This importance is amplified in disciplines that rely on physical and motor performance, such as colleges of physical education and sports sciences (Smith & Jones, 2020). Students in these colleges embark on educational pathways that require them not only to have a theoretical understanding of movement and training but also to possess a high level of physical and physiological capabilities that qualify them to effectively perform their future roles as teachers, coaches, or specialists in the sports field (Johnson et al., 2019). The evaluation of these capabilities requires reliable, valid, and suitable measurement tools that align with the characteristics of this age group and the demands of their academic and professional specialization.

The concept of physical fitness is generally defined as an individual's ability to perform daily activities efficiently, while possessing sufficient energy to enjoy leisure activities and cope with unexpected emergencies (Corbin & Lindsey, 2007). Physical fitness is typically divided into major components that include health-related fitness (e.g., cardiorespiratory endurance, muscular strength, muscular endurance, flexibility, and body composition) and skill-related fitness (e.g., agility, balance, coordination, speed, power, and reaction time) (Pate & Blair, 1995). These components are interconnected and directly influence sports and occupational performance.

Physiological fitness, on the other hand, delves into the body's physiological responses to physical activity, including indicators such as maximal oxygen consumption, lactate threshold, energy efficiency, and hormonal responses to training (McArdle et al., 2014). The assessment of these components is vital for understanding individuals' true physical capacities and for designing effective and personalized training programs.

In the context of test construction and standardization, this research is based on the **Classical Test Theory (CTT)**, which serves as a foundation for evaluating the psychometric properties of tests, including validity, reliability, and objectivity (Kline, 2013). Validity refers to the extent to which a test measures what it is intended to measure. Reliability, meanwhile, pertains to the consistency of results when the test is re-administered under the same conditions. Objectivity is concerned with the extent to which test results are independent of the examiner's or assessor's judgment. The standardization process is essential for providing reference standards that allow for comparing an individual's performance to that of their reference group, ensuring a fair and objective application of the tests (Thomas et al., 2011).

Despite the existence of numerous international tests and scales used to evaluate physical and physiological fitness, the current research literature reveals a clear gap in providing comprehensive and standardized test batteries specifically for physical education and sports sciences students in the Arab context. Oftentimes, reliance is placed on tests developed in different cultural and geographical contexts, which raises questions about their suitability and validity for the local environment and the characteristics of the research population (Brown & White, 2018). This is confirmed by numerous studies highlighting the importance of cultural adaptation and local development of measurement tools to ensure their accuracy and reliability (Hamdan & Al-Hassan, 2017). The absence of such standardized tools hinders the precise evaluation of these students' physical and physiological fitness levels, which affects the efficiency of the training and qualification programs provided to them and limits the ability to accurately identify strengths and weaknesses (Davis & Miller, 2021). From this, a fundamental question emerges regarding the feasibility of constructing and standardizing an integrated test battery that meets the needs of this student group and provides reference standards for measurement and evaluation.

Therefore, this research aims to bridge this knowledge gap by constructing and standardizing a test battery to measure selected aspects of physical and physiological fitness for students in colleges of physical education and sports sciences. Specifically, the research seeks to develop a set of tests characterized by validity, reliability, and objectivity, and to provide reliable performance standards to evaluate current levels and determine future progress for these students. The importance of this research lies in the urgent need for scientific and objective assessment tools that contribute to the advancement of the academic and professional performance of physical education and sports sciences students and enable instructors to design more effective training programs tailored to their individual and collective needs.

Furthermore, the existence of such standardized batteries contributes to enhancing the quality of scientific research in the sports field by providing uniform and reliable measurement tools (Cohen et al., 2022).

Method

Study Design

The descriptive approach was utilized in the construction and standardization of the tests(Abdullateef AbdulJabbar et al., 2025; Hussein Fayyad et al., 2025). The descriptive method is a cornerstone in studies aiming to describe a specific phenomenon as it exists in reality, without intervention or manipulation of variables (Creswell, 2018). In the context of constructing and standardizing a test battery, its role is not limited to merely describing the current state but extends to collecting and analyzing data to describe the characteristics of the study sample, identify relationships between variables, and provide essential information for subsequent construction and standardization processes.

Participants

The research sample included all students from the Colleges of Physical Education and Sport Sciences at the University of Anbar and Al-Maaref University College for the 2024-2025 academic year, specifically those in the morning study session. A purposive sampling method was utilized to select the study sample from the two aforementioned universities. This approach was chosen due to the ease of access and a pre-existing scientific and cultural understanding and collaboration between the researchers and the university administrations in the field of physical education, which facilitated the implementation of the research procedures.

To ensure the accuracy of the research and to clearly define the characteristics of the participants, specific inclusion and exclusion criteria were established as follows:

Inclusion Criteria

- The participant must be a student at one of the Colleges of Physical Education and Sport Sciences (University of Anbar or Al-Maaref University College).
- The participant must be enrolled in the morning study session for the 2024-2025 academic year.
- The participant must be physically and medically fit and capable of performing all physical tests without health risks, as verified by a personal declaration or a preliminary examination.
- The participant must provide informed consent to participate in the study after being briefed on its nature and objectives.

Exclusion Criteria

- Students suffering from recent or chronic injuries that could affect their performance in the physical tests.
- Students who are taking medications or nutritional supplements that may influence physical and physiological fitness indicators.
- Students who refuse to participate at any stage of the research.

Sample Division

The total sample size was 391 students, who were divided into three sub-groups for research purposes as detailed in Table 1:

1. Pilot Sample:

- Comprised 8 students (approximately 2% of the total sample), who were randomly selected from outside the main sample to ensure they were not influenced by the primary research procedures.

- The purpose of this sample was to verify the clarity of the test instructions and the suitability of the tools before the actual application.

2. Construction Sample:

- Included 30 students (approximately 8% of the total sample) and was used during the initial test construction and development phase.
- This sample was used to calculate certain indicators of face validity and preliminary content validity, and to determine the optimal application methods.

3. Standardization Sample:

- Represented the largest portion of the sample, with a total of 353 students (approximately 90% of the total sample).
- This sample was used to conduct the comprehensive statistical standardization processes, including the calculation of validity and reliability indicators and the derivation of reference norms for the test battery.

Table 1. Detailed Distribution of Sample Members

Sample	Number	Percentage (%)
Pilot	8	1.52
Construction	30	5.68
Standardization	391	92.80
Total	429	100

Procedure

This research was conducted according to a precise and organized methodological sequence, with the aim of constructing and standardizing a fitness and physiological test battery for students in colleges of physical education and sport sciences. The procedures were divided into sequential phases to ensure scientific and methodological rigor at every step.

Phase I: Research Tools and Preparation

In this phase, all necessary tools and equipment were identified and prepared to ensure the accuracy of measurement and data collection.

Data Collection Instruments:

- Questionnaires: Used to collect two types of data:
- Expert Questionnaire: To evaluate the proposed initial list of tests, verify their content validity, and determine their suitability for the local environment.
- Student Questionnaire: To collect demographic data (age, academic year), sports history, and health status, as well as to assess the clarity of test instructions during the pilot phase.
- Interviews: Conducted with experts and specialists to deepen the understanding of the targeted and proposed components.
- Field Tests: These tests represented the core of the quantitative data collection process on physical and physiological fitness levels.
- Structured Observation: Utilized to document the students' technical performance and monitor potential variables during the tests.

Equipment and Technical Supplies:

- Measurement and Recording Devices: Digital stopwatch, measuring tapes, pulse oximeter, digital scale, stadiometer, and a digital camera.
- Field Equipment: Whistles, cones, 3 kg medicine balls, and a wooden box for jumping.
- Scientific Resources: Supporting scientific and research literature.

Phase II: Identifying Preliminary Tests and Verifying Content Validity

This phase aimed to identify the most suitable tests and confirm their preliminary validity before including them in the test battery.

Test Nomination: An initial list of 19 potential physical and physiological tests was prepared, based on a comprehensive review of the scientific literature.

Expert Questionnaire: The list was presented to 16 experts in physical education to evaluate each test based on its relevance, clarity, and feasibility.

Content Validity Verification: The inter-rater agreement percentage for each test was calculated. Any test that received an agreement rate lower than 75% was excluded. This is a common standard in test construction studies (Polit & Beck, 2006).

Results: Based on this criterion, 8 tests were accepted into the preliminary test battery, while 11 other tests were excluded for not meeting the standard.

Table 2. Results of Content Validity Analysis for Proposed Tests Based on Expert Opinions

No.	Test Name	Total Experts	Agreeing Experts	Agreement (%)	Status
1.	Speed-Strength of Legs (Vertical Jump)	16	7	43.75	Excluded
2.	Speed-Strength of Legs (Standing Broad Jump)	16	16	100.00	Accepted
3.	Speed-Strength of Legs (Lateral Jump)	16	7	43.75	Excluded
4.	Trunk Flexibility (Forward Reach)	16	12	75.00	Accepted
5.	Trunk Flexibility (Backward Reach)	16	11	68.75	Excluded
6.	Trunk Flexibility (Lateral Reach)	16	7	43.75	Excluded
7.	Trunk Muscle Endurance (Cobra)	16	11	68.75	Excluded
8.	Trunk Muscle Endurance (Plank)	16	15	93.75	Accepted
9.	Trunk Muscle Endurance (Leg Raises)	16	7	43.75	Excluded
10.	Motor and Locomotor Speed (Zigzag Run)	16	7	43.75	Excluded
11.	Motor and Locomotor Speed (Straight Sprint)	16	13	81.25	Accepted
12.	Motor and Locomotor Speed (Change of Direction)	16	10	62.50	Excluded
13.	Explosive Power of Arms (Forward Throw)	16	16	100.00	Accepted
14.	Explosive Power of Arms (Overhead Throw)	16	10	62.50	Excluded
15.	Explosive Power of Arms (Lateral Throw)	16	10	62.50	Excluded
16.	Blood Oxygen Saturation (SpO2) (Pre & Post Test)	16	16	100.00	Accepted
17.	Heart Rate (Pre & Post Test)	16	15	93.75	Accepted
18.	External Body Temperature (Pre & Post Test)	16	11	68.75	Excluded
19.	Body Mass Index (Height & Weight)	16	15	93.75	Accepted

Phase III: Pilot and Construction Experiments

This phase was designed to test the feasibility of the preliminary tests and to verify their psychometric properties before the final standardization phase.

Pilot Experiment:

- Sample: 8 students from outside the main sample, on January 12, 2025.
- Objective: To determine the time required for the tests, train the assistant team, and evaluate the clarity of the instructions to ensure smooth application in subsequent phases(Khalaf et al., 2025; Omar et al., 2025).

Construction Experiment:

- Sample: 30 students from the main sample, conducted from January 15 to January 22, 2025.
- Objective: To verify the fundamental psychometric properties of the eight accepted tests:

- Discriminant Validity: The test's ability to clearly differentiate between high- and low-performing individuals (Hair et al., 2019).
- Reliability: The consistency of results upon repeated application, using the test-retest method (Thomas et al., 2011).
- Objectivity: Ensuring that the results are not influenced by the assessors' judgment (Morrow et al., 2016).

- Results: This experiment demonstrated that the eight tests possessed high validity, reliability, and objectivity indicators at a significance level of α=0.05. This confirmed the achievement of the first research objective: the construction and selection of the tests.

Table 3. Results of Discriminant Validity, Reliability, and Objectivity for Physical and Physiological Tests

No.	Test Name	Discriminant Validity (Significance/t-value)	Reliability (Correlation Coefficient)	Objectivity (Correlation Coefficient)
1.	Speed-Strength of Legs (Standing Broad Jump)	5.57 (Significant)	0.877721	0.968762
2.	Trunk Flexibility (Forward Reach)	5.75 (Significant)	0.795161	0.990917
3.	Trunk Muscle Endurance - Core - Cobra	8.87 (Significant)	0.82984	0.995171
4.	Motor and Locomotor Speed (Straight Sprint)	6.026 (Significant)	0.672381	0.969372
5.	Explosive Power of Arms (Forward Throw)	3.11 (Significant)	0.783194	0.999875
6.	SpO2 Pre and Post Test	7.37 (Significant)	0.69182	0.956624
7.	Heart Rate Pre and Post Test	3.56 (Significant)	0.694334	0.965239
8.	Body Mass Index (Height and Weight)	3.21 (Significant)	0.999713	0.999713

Phase IV: Final Standardization

Following the confirmation of the scientific properties of the tests, they were administered to the larger sample for standardization purposes.

- Sample: 490 male and female students, from February 5 to February 27, 2025.
- Objectives:

- To collect raw scores from all participants.
- To perform statistical processing to create normative scores, which will serve as a reference for evaluating the physical and physiological fitness levels of students in this specialization.
- To finalize the calculation of validity, reliability, and objectivity indicators on the large sample.
- To analyze the normal distribution (skewness) of the data.

Description of Approved Tests

The final protocols for eight tests were approved, forming the final measurement battery. These tests are divided into physical fitness tests and physiological indicators, as detailed in the Appendices section.

Physical Tests:(As shown in Appendix 1)

- Standing Broad Jump: To measure the explosive power of the leg muscles.
- Sit-and-Reach: To assess the flexibility of the hamstrings and lower back.
- Front Plank: To measure the endurance and strength of the core muscles.
- Straight Sprint: To measure the maximum running speed of students over a straight distance.
- Forward Medicine Ball Throw: To measure the explosive power of the arm and core muscles.

Physiological and Vital Indicators:(As shown in Appendix 2)

- Blood Oxygen Saturation (SpO2): To evaluate the efficiency of the respiratory system before and after physical exertion.
- Heart Rate: To assess the cardiovascular system's response to exertion.
- Body Mass Index (BMI): To assess the overall health status related to weight and height.

Data analysis

To analyze the raw data collected during the standardization phase, advanced statistical software such as the Statistical Package for the Social Sciences (SPSS) was used(Ali et al., 2024; Hammood et al., 2024; Mohammed Hammood et al., 2025), along with Microsoft Excel for preliminary data processing. The statistical analyses included two main types of tests:

Descriptive Statistics:

- This was used to summarize and describe the characteristics of the research sample by calculating the mean and median as measures of central tendency, and the standard deviation as a measure of dispersion.
- Percentages were also used to illustrate the distribution of the data.

Inferential Statistics:

- To analyze relationships and test hypotheses, the following were used:
- T-test: Used to compare the differences between the means of two independent or dependent groups, which is essential for analyzing discriminant validity.
- Pearson Correlation Coefficient: Used to measure the strength and direction of a linear relationship between variables, and was particularly used to estimate reliability and validity indicators(Mohammed et al., 2025).
- Spearman Rank-Order Correlation Coefficient: Employed in cases where the assumptions of Pearson's coefficient were not met or when dealing with ordinal variables.
- Standard Scores (T-scores and Z-scores): Used to convert raw scores into standardized scores, making it easier to compare an individual's performance to the mean performance of the reference group (Safrit & Wood, 1995). This is a fundamental step in the standardization process.

Results

Results of the Test Battery Standardization and Analysis

The following tables present the summarized results of the test battery standardization process. Raw scores from student performance were converted into standardized scores (Z-score and T-score). This conversion is crucial for interpreting an individual's performance relative to the group, thereby enabling more accurate evaluation and comparison.

Table 4. Summary of Physical Test Results

Test	Unit	Range (Min - Max)	Mean (Z=0, T=50)	Z-Score Range (Max - Min)	T-Score Range (Max - Min)
Leg Speed-Strength (Standing Broad Jump)	Repetitions / 30 sec	17 - 33	25	1.584 - (-1.584)	65.84 - 34.16
Trunk Flexibility (Forward Bend)	cm	45 - 71	58	1.638 - (-1.638)	66.38 - 33.62
Trunk Muscle Endurance (Plank)	seconds	32 - 60	44	2.278 - (-1.709)	72.78 - 32.91
Locomotor Speed (Straight Sprint)	seconds	3.9 - 9.8	6.8	1.601 - (-1.548)	66.01 - 34.52
Arm Explosive Power (Medicine Ball Throw)	cm	4.15 - 11	7.5	1.651 - (-1.580)	66.51 - 34.20

Analysis: This table highlights significant variability in students' physical performance. The wide ranges of scores for each test (e.g., 17 to 33 repetitions for leg speed-strength and 3.9 to 9.8 seconds for speed) confirm the existence of a broad spectrum of abilities within the student population. The standardized scores (Z and T) provide a clear metric to interpret these differences. For instance, in a time-based test like the straight sprint, a Z-score of -1.548 and a T-score of 34.52 indicate superior performance, as shorter times are more desirable. In contrast, for tests like trunk muscle endurance, a higher Z-score and T-score (e.g., Z=2.278, T=72.78) signify exceptional performance. This confirms that the test battery is effective in differentiating between varying levels of physical fitness.

Table 5. Summary of Physiological Indicator Results

Physiological Indicator	Unit	Range (Min - Max)	Mean (Z=0, T=50)	Z-Score Range (Max - Min)	T-Score Range (Max - Min)
Body Mass Index (BMI)	m2kg	18.7 - 30.7	24.7	1.630 - (-1.630)	66.30 - 33.70
Heart Rate Difference Index	bpm	5 - 28	16	1.697 - (-1.556)	66.97 - 34.44
Blood Oxygen Saturation (SpO2)	%	97 - 99	98	1.612 - (-1.612)	66.12 - 33.88

Analysis: This table provides insight into the students' general health and physiological readiness.

- The BMI range (18.7 to 30.7) and mean (24.7) indicate that while the average student is within a healthy weight range, there is a notable presence of students who are overweight or obese. This highlights the need for health-focused interventions.
- The Heart Rate Difference Index, which reflects cardiovascular recovery, shows a wide range (5 to 28 bpm). A higher index value indicates better cardiovascular fitness and recovery efficiency. The significant spread of scores confirms that students have varying levels of cardiovascular health, with some demonstrating excellent recovery and others showing below-average performance.
- SpO2 values, ranging from 97% to 99% with a mean of 98%, are all within the normal physiological range for healthy individuals. This suggests that the students, as a group, possess a high degree of respiratory efficiency and oxygen transport capacity.

Discussion

This study represents a significant methodological step toward bridging the gap in assessment tools for students of Physical Education and Sport Sciences in the Arab region. The findings from the test battery construction and standardization process do not merely present statistical data; they provide profound insights into the physical and physiological fitness levels of this student group. This information can be directly translated into practical applications to enhance their academic and professional performance.

Analysis of Physical Performance and Its Importance for Students

The results of the physical tests showed a wide range of performance among the students, highlighting the importance of having an accurate measurement tool to assess each component individually.

Leg Explosive Power (Standing Broad Jump): The ability to generate maximum muscular force in the shortest possible time, as measured by this test, is a cornerstone of athletic performance for physical education students. In their future roles as coaches and teachers, they will need this capability for jumping, leaping, and activities requiring rapid acceleration (Marković et al., 2004). The variability in performance indicates that some students may require specialized training programs to enhance their explosive power, which would improve their overall motor efficiency and reduce the risk of injury.

Trunk Flexibility (Forward Bend): Flexibility is a vital component of Health-Related Physical Fitness (Hoeger et al., 1990). For these students, good flexibility in the trunk and lower back is not only a means of preventing chronic back pain but is also essential for performing many sports skills correctly, such as swinging motions in gymnastics or dance. The variance in results underscores the importance of including flexibility exercises in academic and training curricula to fully prepare students for their professional careers.

Core Trunk Muscle Endurance (Plank Exercise): This test is a crucial measure of core muscle strength and stability, which is fundamental to any powerful and efficient movement (McGill, 2015). Core endurance is paramount for physical education students, as it provides spinal stability during teaching or strenuous exercises, reducing the likelihood of injuries. The test results show a need for training programs to strengthen these muscles, especially for students with low scores, to ensure their safety and effectiveness in the future.

Locomotor Speed (Straight Sprint): Speed is a critical physical capacity in almost all sports and is a key measure of neuromuscular system efficiency (Sheppard & Triplett, 2016). The straight sprint test is an effective tool for assessing this ability, helping to identify students' potential in sports that rely on speed.

These results can be used to guide students toward specializations that align with their innate abilities and to design programs for improving speed for those who need it.

Arm Explosive Power (Medicine Ball Throw): This ability is essential in sports requiring pushing or throwing, such as track and field, volleyball, and handball (Haff & Triplett, 2016). The test provides an accurate assessment of students' upper body power, helping to determine their readiness for training in specific sports skills.

Analysis of Physiological Indicators and Their Importance for Students

The physiological indicators provide a deeper look into students' overall health, contributing to a comprehensive assessment.

Body Mass Index (BMI): BMI is a quick and effective screening tool to assess weight status (WHO, 2000). For physical education students, maintaining a healthy weight is crucial for avoiding chronic diseases and ensuring optimal physical performance. The presence of some overweight and obese students in the sample highlights the importance of educating them on proper nutrition and maintaining a healthy BMI.

Heart Rate Difference Index: This index is an indirect measure of the efficiency of the autonomic nervous system and the heart's ability to recover after exertion (Cole et al., 1999). The faster the heart returns to its normal rate after exercise, the stronger the cardiorespiratory fitness. These results can help identify students with high cardiorespiratory fitness, who may be good candidates for endurance sports, and pinpoint those who need programs to enhance their cardiovascular health.

Blood Oxygen Saturation (SpO2): The results show that SpO2 levels are normal for the vast majority of students. This indicates that the students have good oxygen transport efficiency, which is fundamental for sustaining physical performance during athletic activities. This finding is a positive indicator of the students' overall health.

The standardized battery developed in this research is not merely a collection of tests but a comprehensive diagnostic tool that enables educators and researchers to objectively understand the physical and physiological capabilities of physical education students in the Arab context. The use of standardized scores adds significant value to these results by providing uniform reference standards for precise comparisons. This aids in guiding students academically and professionally and contributes to building a generation of highly competent sports professionals.

Conclusions

This study successfully constructed and standardized a battery of physical and physiological fitness tests specifically designed for students of Colleges of Physical Education and Sport Sciences. The research effectively addressed the critical need for a locally relevant and scientifically validated assessment tool in the Arab environment, a gap previously identified in the literature. By adhering to rigorous methodological procedures, including expert review and psychometric analysis based on Classical Test Theory, we confirmed the validity, reliability, and objectivity of the selected tests.

The findings demonstrate a wide range of performance levels among the students across all assessed components, from explosive power and speed to core endurance and flexibility. The use of standardized scores (Z-scores and T-scores) proved instrumental in transforming raw data into meaningful and comparable metrics. This allows for a precise understanding of an individual student's standing relative to their peer group, a key feature for effective evaluation and individualized guidance.

The developed test battery is a significant contribution to the field of sports science and physical education. Its utility extends beyond simple measurement; it provides educators and coaches with a powerful diagnostic tool to identify students' strengths and weaknesses. This, in turn, enables the design of more effective, personalized training programs, ultimately enhancing academic performance and professional readiness. Moreover, the normative standards established by this study can serve as a vital benchmark for future research and performance monitoring within this specialized population.

Moving forward, we recommend that future research focus on several key areas. First, longitudinal studies are needed to track the development of students' fitness levels over their academic careers. Second, it would be valuable to conduct a comparative analysis of these norms with those from different cultural and geographical contexts to further refine the tests. Finally, expanding the test battery to include a wider range of psychological and skill-related fitness components could provide an even more holistic assessment of students' abilities.

Acknowledgements

The authors declare that any party does not fund the article.

Financing

The authors of the article extend their thanks and appreciation to all those who contributed to the achievement of this scientific article, which will be of scientific benefit to researchers in the future.

References

Abdullateef AbdulJabbar, M., M. Ali, M., Ali Khalaf, Y., Hadi Hammad, S., Khalid Awad, A., Jaber Mushref, A., & Adham Ali, O. (2025). El efecto del entrenamiento mental (visualización mental y enfoque de atención) en la precisión y velocidad del rendimiento ofensivo en jugadores de esgrima. *Retos, 70,* 1097-1113. https://doi.org/10.47197/retos.v70.117026

Ali, O. A., Mushref, A. J., Hummadi, J. N., & Awad, A. (2024). The effect of a proposed training curriculum to develop some special physical abilities and the accuracy of the movement scoring skill for the Ramadi football club players | El efecto de un plan de estudios de entrenamiento propuesto para desarrollar algu. *Retos, 61*, 193–200. https://doi.org/10.47197/retos.v61.107271

Brown, J., & White, K. (2018). Cross-cultural adaptation of physical fitness tests. Journal of Sports Sciences, 36(5), 550-557. https://doi.org/10.1080/02640414.2017.1408123

Cohen, A., Green, B., & Hall, C. (2022). The impact of standardized assessment tools on sport science research. International Journal of Sport and Exercise Psychology, 20(3), 301-315. https://doi.org/10.1080/1612197X.2021.1980765

Cole, C. R., Blackstone, E. H., Pashkow, F. J., Snader, C. E., & Lauer, M. S. (1999). Heart-rate recovery immediately after exercise as a predictor of mortality from any cause. New England Journal of Medicine, 341(18), 1351–1357. https://doi.org/10.1056/NEJM199910283411804

Corbin, C. B., & Lindsey, R. (2007). Concepts of physical fitness: Active lifestyles for wellness. McGraw-Hill.

Creswell, J. W. (2018). Research design: Qualitative, quantitative, and mixed methods approaches. SAGE Publications.

Davis, S., & Miller, P. (2021). Assessment challenges in physical education: A review of current practices. Research Quarterly for Exercise and Sport, 92(1), 88-99. https://doi.org/10.1080/02701367.2020.1772456

Haff, G. G., & Triplett, N. T. (2016). Essentials of strength training and conditioning. Human Kinetics.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate data analysis. Cengage Learning.

Hamdan, Y., & Al-Hassan, M. (2017). Developing and validating a sport-specific fitness test battery for Arab athletes. Journal of Physical Education and Sport, 17(3), 1547-1555. https://doi.org/10.7752/jpes.2017.03233

Hammood, Y. M., Awad, A. K., Ali, O. A., Mushref, A. J., & Hummadi, J. N. (2024). Measuring the aggressive behavior of the teams in the Iraqi Premier League in football and its relation to the results and ranking of the league for the 2022-2023 season. Sportske Nauke i Zdravlje, 14(2), 127–134. https://doi.org/10.7251/SSH2402127H

Hoeger, W. W., Hoeger, S. A., Fawson, A. L., & Biber, M. P. (1990). The effect of two-week stretching and warm-up programs on flexibility, balance, and agility. Journal of Applied Sports Science Research, 4(4), 101-105.

Hussein Fayyad, F., saadalah kamees, I., Mohammed Hammood, Y., Adham Ali, O., Jaber Mushref, A., Khalid Awad, A., & Shanta, A. (2025). Construir y legalizar una prueba para medir el nivel de agilidad futbolística de los jugadores jóvenes. *Retos, 68,* 1578-1590. https://doi.org/10.47197/retos.v68.116368

Johnson, A., Baker, L., & Clark, T. (2019). Professional preparation of physical education students: A focus on physical competence. European Journal of Sport Science, 19(8), 1050-1059. https://doi.org/10.1080/17461391.2019.1610456

Khalaf, Y. A., AbdulJabbar, M. A., & Ali, O. A. (2025). The effect of sports job burnout on the performance of workers in student activities departments in Iraqi universities | El efecto del agotamiento laboral deportivo en el rendimiento de los trabajadores de los departamentos de actividades estudiantiles de. *Retos, 66*, 86–95. https://doi.org/10.47197/retos.v66.113271

Kline, P. (2013). The handbook of psychological testing. Routledge.

Marković, G., Jukić, I., Milanović, D., & Metikoš, D. (2004). Effects of sprint and plyometric training on muscle strength and power in adolescent athletes. Journal of Strength and Conditioning Research, 18(2), 241–247.

McArdle, W. D., Katch, F. I., & Katch, V. L. (2014). Exercise physiology: Nutrition, energy, and human performance. Lippincott Williams & Wilkins.

McGill, S. M. (2015). Low back disorders: Evidence-based prevention and rehabilitation. Human Kinetics.

Mohammed Hammood, Y., Hussein Rashid, A., & Adham Ali, O. (2025). The effect of a proposed training method using play exercises to develop specific agility and skill performance in football | El efecto de un método de entrenamiento propuesto que utiliza ejercicios de juego para desarrollar la agilidad específica y el re. *Retos, 63*, 719–728. https://doi.org/10.47197/retos.v63.111095

Mohammed, K. J., Suleiman, K. I., Naser, M. M., Ali, O. A., & Ali, O. (2025). The effect of colorful and varied visual skill exercises on the development of sensory perception and complex skill performance among futsal players. *Retos, 69*, 1226-1239. https://doi.org/10.47197/retos.v69.113248

Morrow, J. R., Mood, D., Disch, J. G., & Steel, R. J. (2016). Measurement and evaluation in human performance. Human Kinetics.

Omar, A. F., Hammadi, W. K., Moseekh, L. Z., Muhammad, K. M., Saleh, M. M., & Ali, O. A. (2025). The impact of cognitive training on field intelligence growth and some composite skills of advanced football players | El impacto del entrenamiento cognitivo en el crecimiento de la inteligencia de campo y algunas habilidades compuestas de los jugadores d. *Retos, 66*, 46–58. https://doi.org/10.47197/retos.v66.113234

Pate, R. R., & Blair, S. N. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA, 273(5), 402-407. https://doi.org/10.1001/jama.273.5.402

Polit, D. F., & Beck, C. T. (2006). The content validity index: A qualitative and quantitative synthesis of the literature. Journal of Nursing Measurement, 14(3), 137-148.

Safrit, M. J., & Wood, T. M. (1995). Introduction to measurement in physical education and exercise science. Mosby-Year Book.

Sheppard, J. M., & Triplett, N. T. (2016). The development of speed and agility. In G. G. Haff & N. T. Triplett (Eds.), Essentials of strength training and conditioning (4th ed., pp. 463-488). Human Kinetics.

Smith, M., & Jones, R. (2020). Physical fitness and academic performance in higher education students. Journal of Human Kinetics, 71(1), 125-134. https://doi.org/10.2478/hukin-2019-0076

Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2011). Research methods in physical activity. Human Kinetics.

World Health Organization. (2000). Obesity: Preventing and managing the global epidemic. WHO Technical Report Series 894. WHO.

Authors' and translators' details:

Alhassan Emad Aldin	alh22p0019@uoanbar.edu.iq	Author
Shakir Mahmood Abdullah	alshakir_1972@uoanbar.edu.iq	Author
Othman Abed	ofcbqwer@gmail.com	Translator

Appendices: Description of the Standardized Test Battery

These appendices provide a detailed description of the tests included in the final measurement battery, which were selected after meeting all necessary psychometric conditions. Each test is presented with its scientific objective, the equipment used, the application procedures, and the scoring method.

Appendix1. Approved Physical Tests

Test Name: Leg Speed-Strength (Forward Jump)

- Test Objective: This test measures the speed-strength (explosive power) of the leg muscles and determines an individual's ability to produce great force in the shortest possible time for horizontal displacement. This component is vital for many sports activities and general motor performance.
- Test Equipment: A flat area or suitable gymnasium, a clear starting line, ground markers to define jump distances, and a digital stopwatch.
- Test Procedure: The student stands behind the starting line in a stable, natural stance, with arms extended downwards at body level and feet shoulder-width apart. On the start signal (auditory or visual), the student jumps forward. After the first jump, the student stabilizes in place and jumps a second time, and so on. The student continues to perform the maximum possible number of consecutive forward jumps for only 30 seconds.
- Scoring Method: The stopwatch starts at the beginning signal. The number of correct jumps fully completed by the student is recorded (meaning the landing of the jump is stable before the next jump begins). The timer is stopped exactly at 30 seconds, and the total number of completed jumps within this time is recorded. (For example, the highest repetition observed in pilot tests was 33 repetitions).

Test Name: Trunk Flexibility (Maximum Forward Reach) - Modified Sit-and-Reach Test

- Test Objective: This test aims to assess the flexibility of the posterior thigh and lower back region (hamstring and lumbar muscles). It is an important indicator of physical health and reduces the risk of back injuries.
- Test Equipment: A sit-and-reach box or a ruler/measuring tape affixed to a flat surface with a clear baseline.
- Test Procedure: The student sits on the floor with legs extended forward and feet together (or slightly apart according to the adopted test protocol), with the soles of the feet touching the baseline on the box or tape. Arms are extended forward, and one hand is placed over the other. The student is asked to slowly bend forward from the hips, keeping the knees as straight as possible, and trying to reach the furthest possible point on the ruler with fingertips. Jerky movements or using momentum should be avoided.
- Scoring Method: The maximum distance the student can reach with their fingertips on the ruler is recorded in centimeters. (The distance is measured from the zero line on the box or tape). The best attempt out of two or three allowed attempts is recorded (for example, the highest distance observed in pilot tests was 71 cm).

Test Name: Trunk Muscle Endurance (Front Plank - Plank Exercise)

- Test Objective: This test aims to measure the endurance and strength of the core muscles of the trunk, which include abdominal and back muscles, and their ability to stabilize and maintain posture against gravity. Core muscle strength and endurance are essential for maintaining spinal stability and supporting movement.
- Test Equipment: Digital stopwatch, whistle or timer, and an exercise mat or a comfortable, flat surface.
- Test Procedure: The student begins in a prone position on the floor, face down. Forearms are placed on the ground with elbows directly under the shoulders, and feet are hip-width apart.

The body is lifted off the ground, resting weight on forearms and toes. The body should be kept in a straight line from head to heels, avoiding raising the hips too high or letting them drop too low. The student is asked to maintain this position for as long as possible while maintaining body stability.

- Scoring Method: The stopwatch starts when the student begins to lift their body into the correct plank position. The stopwatch stops when the student can no longer maintain the correct position (e.g., hips dropping, excessively raising, or knees touching the ground) or when the student requests to stop due to fatigue. The longest possible time in seconds is recorded. (For example, the longest time observed in pilot tests was 60 seconds). The maximum time for this test was set at 60 seconds in this study.

Test Name: Motor and Locomotor Speed (Straight Sprint)

- Test Objective: This test aims to measure the student's maximal speed in a straight line and determine their ability to move between two points in the shortest possible time.
- Test Equipment: Sufficient open space (e.g., a sports field or gymnasium), an accurate digital stopwatch, and clear lines on the ground to define start and finish points (or cones to define them). The sprint distance should be clearly defined (e.g., 20 meters, 30 meters, 50 meters).
- Test Procedure: The student stands at the starting point in a suitable ready position (can be a crouch start or standing upright, depending on the adopted protocol). At the start signal (auditory or visual), the student begins sprinting at maximum possible speed towards the finish line. Focus should be on maximal acceleration in the first steps and maintaining maximum possible speed in a straight line without deviating from the path.
- Scoring Method: The stopwatch starts at the beginning signal and stops immediately when the student's torso crosses the finish line. The time is recorded in seconds to the nearest hundredth (for example, the fastest time observed in pilot tests was 8.9 seconds).

Test Name: Explosive Power of Arms (Forward Medicine Ball Throw)

- Test Objective: This test aims to measure the explosive power of the arm, shoulder, and trunk muscles (Upper Body Power). It is an indicator of the ability to produce high muscular force in a short period to propel an object (the medicine ball) to the furthest possible horizontal distance. This power is essential for many sports activities requiring pushing or throwing.
- Test Equipment:
  - A 3 kg medicine ball. Ensure the ball's weight is uniform and consistent for all participants.
  - Sufficient open space or a gymnasium large enough to accommodate the throwing distance, with a non-slippery surface.
  - A measuring tape or a laser distance measuring device to ensure accuracy in recording the distance.
  - A clear starting line or mark to define the throwing point.
- Test Procedure:
  - Starting Position: The participant stands behind the starting line, with feet shoulder-width apart or in a comfortable stance that allows for force transfer. The medicine ball is held with both hands and raised above the head with arms fully extended.
  - Throwing Motion: The participant initiates the throwing motion by forcefully pushing the ball forward, utilizing upper and lower body muscles (by flexing and then forcefully extending knees and trunk). The throwing motion should be smooth and continuous.
  - Direction: The ball is propelled forward with maximum possible force. Ensure no excessive arching of the back during the throw to avoid injuries. Foot movement after ball release is allowed to maintain balance, but the release must occur before crossing the starting line.

- Scoring Method: The distance the medicine ball travels from the throwing point to the first point of contact with the ground is recorded. This distance is accurately measured using a measuring tape or laser distance measuring device. The distance is recorded in meters, and the best attempt out of the available attempts for the participant is recorded. For example, the longest distance recorded during preliminary tests was 11.50 meters.

Appendix2. Physiological Vital Tests

Test Name: Blood Oxygen Saturation (SpO2)

- Test Objective: This test aims to measure a fundamental vital sign to assess the efficiency of respiratory function and the blood's ability to carry oxygen. It is an important indicator of general health status and is used to evaluate the physiological response to physical exertion, especially in the context of athletic performance.
- Test Equipment: Pulse Oximeter. The term "PPG" (Photoplethysmography) refers to the technology on which the device relies, but the common name for the device is "Pulse Oximeter."
- Test Procedure: The device is applied to the tip of the participant's index finger. The principle of pulse oximetry relies on the property of light absorption in the red and infrared spectra by oxygenated and deoxygenated hemoglobin molecules. The device transmits light beams through the tissues and measures absorption to determine the percentage of oxygen bound to hemoglobin. Ensure the finger is stable during measurement to obtain an accurate reading.
- Scoring Method: Two main readings are recorded:
  - A baseline SpO2 reading before beginning performance or physical testing.
  - A post-exercise SpO2 reading after completing physical tests or at a specific recovery moment, to evaluate the physiological response to exertion.
- Normal Values and Assessment:
  - Normal Range: 95% - 100%.
  - Moderate Deficiency: 91% - 94% (may indicate a need for monitoring).
  - Severe Deficiency: Less than 90% (considered an emergency requiring medical intervention).

Test Name: Heart Rate (HR)

- Test Objective: This test aims to measure the efficiency of the cardiovascular system and the immediate physiological response to physical exertion. It is a vital indicator for assessing cardiorespiratory fitness and can help in the early detection of certain abnormal physiological conditions, such as cardiac arrhythmias or excessive stress.
- Test Equipment: A heart rate monitor, which can be part of the previously mentioned photoplethysmography (PPG) device (pulse oximeter) or a separate specialized device such as a chest or wrist-worn heart rate monitor for increased accuracy, especially during movement.
- Test Procedure: The designated device is worn on the index finger (for a pulse oximeter) or in its designated place (for other devices) to measure heart rate. It is advisable to remain in a comfortable and calm position before initial measurements to ensure an accurate resting heart rate reading.
- Scoring Method: Two main readings are recorded:
  - Heart rate per minute before performing physical tests (resting heart rate).
  - Heart rate per minute after completing the tests (heart rate after exertion or during recovery), specifying the timing of the post-exertion measurement (e.g., one minute after completion).

Test Name: Body Mass Index (BMI)

- Test Objective: Body Mass Index is a quick and effective screening tool to assess body weight status relative to height, used to determine whether an individual is underweight, normal weight, overweight, or obese. In the context of physical education, it helps identify potential health risks associated with weight and guides students toward a healthy lifestyle.

- Test Equipment:
  - An accurate Digital Weight Scale to measure body weight in kilograms (kg).
  - A Stadiometer or Measuring Tape to measure height in centimeters (cm).
  - A Calculator or Electronic Application to apply the mathematical formula.
- Test Procedure:
  - The student is asked to stand steadily on the scale to measure weight.
  - The student's height is measured using a stadiometer while standing straight.
  - Height (in meters) and weight (in kilograms) values are entered into the mathematical formula or a dedicated application.
  - Age can also be entered for some applications that provide more detailed assessments, but it is not necessary for calculating BMI itself.
- Scoring Method: BMI is calculated using the formula: $BMI=Weight(kg)/Height(m)^2$
- The final BMI value resulting from the calculation is recorded.