

Reliability and Construct Validity of the Persian Version of the ALPHA-Fitness Test Battery in Iranian Children

Roya Hosseinzaeh Peyghan

PhD Student, Department of Motor Behavior, Urmia university, Iran. Email: Ro.hosseinzaeh@urmia.ac.ir

Article Info	Abstract
Article type: Research Article Review Article Article history: Received 18 February 2026 Received in revised form 20 May 2026 Accepted 05 June 2026 Available online 29 June 2026 Keywords: Physical Fitness, Child, Exercise Test, Psychometrics, Schools.	Introduction: School-based physical fitness (PF) assessments are widely used to monitor health-related fitness in children. However, evidence regarding the measurement properties of internationally recognized fitness batteries remains limited in some cultural contexts, including Iran. Objective: To evaluate the reliability and construct validity of the Iranian adaptation of the ALPHA-Fitness Test Battery among Iranian children. Method: A methodological study was conducted with 458 Iranian schoolchildren (241 boys, 217 girls; mean age=10.45 years). Participants completed the Iranian adaptation of the ALPHA-Fitness Test Battery. Test-retest reliability was evaluated over a 7-day interval using ICC, SEM, MDC95, CV, and Bland-Altman analyses. Construct validity was assessed through EFA and CFA, while measurement invariance across sex was examined using multi-group CFA. Results: All fitness tests demonstrated good-to-excellent reliability (ICC=0.84-0.95), with low measurement error and acceptable agreement between testing sessions. EFA identified a three-factor structure representing muscular fitness, functional/cardiorespiratory fitness, and flexibility, explaining 71.8% of the total variance. CFA supported the proposed model ($\chi^2/df=2.14$, CFI=0.96, TLI=0.95, RMSEA=0.051, SRMR=0.043). Multi-group CFA confirmed configural, metric, and scalar invariance across sex. Conclusions: The Iranian adaptation of the ALPHA-Fitness Test Battery demonstrated satisfactory reliability and strong construct validity, supporting its use for school-based assessment of PF in Iranian children. Cite this Article: Hosseinzaeh Peyghan, R. (2026). Reliability and Construct Validity of the Persian Version of the ALPHA-Fitness Test Battery in Iranian Children. <i>Physical Activity in Children</i>, 3(1), 107-115. https://doi.org/10.22034/pach.2026.588287.1096

1. Introduction

Physical fitness (PF) is a multifaceted concept that comprises several health-related components, including cardiovascular stamina, muscular capacity, flexibility, power, speed, agility, and body fat percentage (1,2). For pediatric and adolescent populations, these PF metrics serve as vital indicators of both present and long-term health, as they are significantly linked to cardiometabolic stability, cognitive performance, psychological health, and overall quality of life (3,4). Research indicates that youth who possess superior PF levels exhibit a reduced predisposition toward chronic health issues, such as hypertension, obesity, and metabolic syndrome, underscoring the critical necessity of monitoring PF throughout developmental years (5-7).

Schools provide an ideal environment for evaluating PF because they enable standardized assessment of large numbers of children in a practical and cost-effective manner (8). Consequently, school-based fitness testing has become an integral

component of physical education programs and public health initiatives in many countries (9). Reliable and valid PF assessments allow educators, clinicians, and researchers to monitor children's physical development, identify individuals requiring additional support, evaluate intervention effectiveness, and describe population-level trends in youth PF over time (10). However, the usefulness of such assessments depends fundamentally on the quality of their measurement properties within the population in which they are applied.

Several standardized field-based PF batteries have been developed for use in youth. Among the most widely adopted are the EUROFIT Test Battery, FITNESSGRAM, and the ALPHA-Fitness Test Battery (11,12). These batteries assess multiple components of PF using inexpensive, easily administered field tests suitable for educational settings. Among them, the ALPHA-Fitness Test Battery has received considerable international attention because its test components were selected on the basis of their associations with present and future health outcomes and their

feasibility for large-scale school-based assessment. Consequently, the ALPHA battery has been widely used in epidemiological studies and school-based PF surveillance across numerous countries.

The implementation of the ALPHA-Fitness Test Battery across diverse cultural and linguistic landscapes requires rigorous validation to ensure test integrity. Because variations in educational environments, pedagogical instructions, and cultural nuances can impact measurement accuracy, the Persian adaptation of this battery requires systematic psychometric verification before widespread application in Iranian schools (13).

Following established cross-cultural adaptation guidelines, the transition from an original instrument to an adapted version demands more than mere translation; it requires a structured process including expert review, pilot testing, and formal psychometric assessment to maintain conceptual equivalence and standardized administration (14). Central to this validation are reliability, which ensures consistency across repeated trials, and validity, which confirms that test scores accurately reflect their intended constructs (15, 16). For school-based fitness monitoring, test-retest reliability is particularly vital, often assessed through Intraclass Correlation Coefficients (ICC) for relative reliability, alongside indices like the Standard Error of Measurement (SEM) and Minimal Detectable Change (MDC) for absolute reliability (17).

Given the multidimensional nature of PF, construct validity is typically evaluated through both Exploratory and Confirmatory Factor Analysis (EFA and CFA) to confirm that the observed test performance aligns with the latent structure of established PF models (18). Furthermore, modern validation protocols increasingly prioritize measurement invariance to ensure that a test functions identically across demographic groups, such as boys and girls, thereby preventing measurement bias (19–21).

Despite the widespread use of the ALPHA-Fitness battery globally, comprehensive psychometric evidence, specifically regarding the factor structure and measurement invariance of the Persian version in Iranian youth, remains insufficient (22). To address this research gap, the current study evaluated the test-retest reliability of the battery in Iranian children, hypothesizing ICC values ≥ 0.75 with acceptable measurement error, and examined its construct validity through EFA and CFA, while further testing for measurement invariance across sex to ensure the instrument aligns with the multidimensional theoretical models of physical fitness.

2. Methods

2.1. Study Design

This study used a cross-sectional design with repeated assessments among students from public primary schools in Iran during the 2024–2025 academic year. The methodology and reporting procedures followed COSMIN guidelines for assessing the reliability and construct validity of measurement instruments.

2.2. Participants

The sample included 458 children (241 boys and 217 girls) aged 8–12 years (mean age = 10.45 ± 1.47 years). Students were recruited from grades 2 to 6 using a multistage school-based sampling approach (Figure 1). A three-stage sampling procedure was applied. Schools were randomly selected from different educational districts, followed by random selection of eligible classes. All students who met the inclusion criteria within the selected classes were then invited to participate. The sample included children from both urban and suburban public schools, representing a broad range of socioeconomic backgrounds typically encountered within the participating educational districts. Although the sample was not intended to provide a nationally representative cohort, the multistage recruitment strategy was designed to enhance external validity by including schools from multiple geographic areas. Eligibility criteria included: (a) enrollment in a participating public school; (b) age between 8 and 12 years; and (c) absence of any known medical condition preventing participation in PF testing. Students with diagnosed physical disabilities, musculoskeletal disorders, neurological impairments, or other health conditions limiting safe participation in physical activity were excluded. The overall sample size exceeded commonly recommended requirements for psychometric investigations. Recommendations generally suggest a minimum of 5–10 participants per observed variable and at least 200 participants for robust factor analysis. Given the five observed PF variables included in the measurement model, the available sample was considered more than adequate for both EFA and CFA. Furthermore, the sample size provided sufficient precision for estimating test-retest reliability coefficients and measurement-error indices. For construct validity analyses, participants were randomly allocated using a computerized randomization procedure into two independent subsamples of approximately equal size. One subsample was used for EFA, whereas the second independent subsample was reserved for CFA.

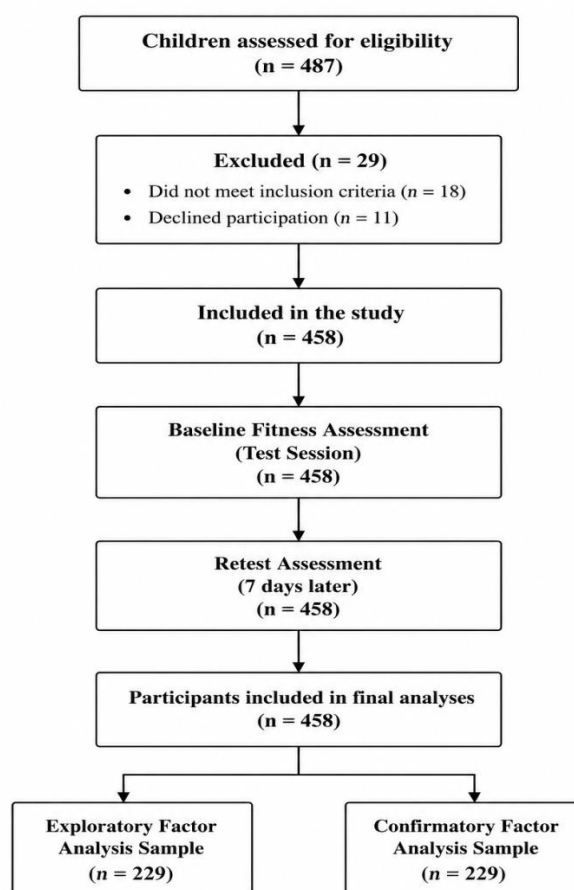


Figure 1. Flowchart of Participant Recruitment, Testing, and Data Analysis Procedures.

2.3. Physical Fitness Test Battery

The Persian adaptation of the school-based PF battery was derived from the ALPHA-Fitness Test Battery (12) and included field-based assessments of major components of PF.

2.3.1. Cross-Cultural Adaptation of the ALPHA-Fitness Test Battery

Because the ALPHA-Fitness Test Battery consists primarily of performance-based field tests rather than questionnaire items, the adaptation process focused on linguistic and procedural standardization rather than modification of the PF assessments themselves. Consequently, the term “Persian adaptation” is used throughout this manuscript to reflect the nature of the adaptation process. The adaptation followed internationally recognized recommendations for cross-cultural adaptation of performance-based assessment protocols. Initially, all administration instructions and participant guidance materials were translated independently from English into Persian by two bilingual experts in physical education and sport sciences. The translated versions were subsequently reviewed by an expert panel consisting of five specialists in pediatric exercise science, physical education, measurement and evaluation, and school health. Panel members evaluated linguistic clarity, cultural appropriateness, instructional equivalence, and practical feasibility for implementation in Iranian schools. Overall agreement among experts exceeded 90% across evaluated domains, indicating excellent content agreement regarding the adapted administration procedures. Because no questionnaire items or psychometric scales requiring semantic

equivalence were adapted, formal back-translation was not performed. This decision represents a limitation of the adaptation process and is acknowledged in the Discussion. A pilot study involving children within the target age range was subsequently conducted to assess comprehension of the instructions and feasibility of test administration. Minor modifications were made to improve clarity of wording and logistical organization during school implementation. Importantly, no modifications were made to the original ALPHA-Fitness testing protocols, scoring criteria, or performance standards. Only the language of administration and organizational procedures were culturally adapted for Persian-speaking schoolchildren. The finalized Persian administration manual was used throughout the study. Before data collection, all examiners completed standardized training sessions covering participant instruction, equipment calibration, demonstration procedures, scoring rules, and quality-control procedures to ensure uniform administration across testing sites.

2.3.2. Anthropometric Measurements

Height and body mass were measured using a portable stadiometer and calibrated digital scale, respectively, and BMI was calculated as kg/m^2 .

2.3.3. Handgrip Strength Test

Maximal isometric handgrip strength was assessed with a dynamometer. Two trials were performed for each hand, and the best score was used for analysis.

2.3.4. Standing Long Jump Test

Lower-limb power was evaluated by the standing long jump test, with the longest distance from two

maximal attempts recorded in centimeters.

2.3.5. 20-Meter Shuttle Run Test

Cardiorespiratory fitness was assessed using the progressive 20-m shuttle run test, with the highest completed stage recorded based on audio-paced running intervals.

2.3.6. 4 × 10-Meter Shuttle Run Test

Speed and agility were evaluated through the 4 × 10-m shuttle run test, and the fastest completion time (0.01 s accuracy) was recorded.

2.3.7. Sit-and-Reach Test

Lower-body flexibility was measured using the sit-and-reach test, with the maximum distance achieved across two attempts recorded.

2.4. Testing Procedures

Data collection was conducted during regular school hours by four trained examiners. Two male examiners assessed boys, whereas two female examiners assessed girls. All examiners completed standardized training before the commencement of testing. Training included standardized participant instruction, demonstration of each PF test, scoring procedures, equipment calibration, and practice sessions designed to minimize examiner-related variability. Participants completed the PF battery on two separate occasions to evaluate test-retest reliability. The retest session was conducted seven days after the initial assessment. A seven-day interval was selected because it is widely recommended in methodological reliability studies as sufficiently long to minimize recall or learning effects while being short enough to avoid meaningful physiological changes in PF among healthy children. Both testing sessions were performed under similar environmental conditions and at approximately the same time of day. Participants completed a standardized warm-up consisting of light aerobic activity and dynamic stretching before testing. PF assessments were administered in the same order during both sessions. Because the primary objective was to evaluate temporal stability of the battery, only test-retest

reliability was examined. Inter-rater and intra-rater reliability were not evaluated because the same trained examiner assessed each participant during both testing sessions using standardized administration procedures. This limitation is acknowledged in the Discussion.

2.5. Statistical Analysis

Data were analyzed using IBM SPSS Statistics and AMOS. Descriptive statistics (mean, standard deviation, and 95% confidence intervals) were calculated, and distributional assumptions were examined using skewness, kurtosis, and visual inspection. Test-retest reliability was assessed using ICC (2,1) with absolute agreement, while absolute measurement error was evaluated through SEM, MDC95, and CV%. Bland-Altman plots and limits of agreement were used to examine session agreement and possible heteroscedasticity. Construct validity was evaluated through sequential EFA and CFA. EFA was performed on a randomly selected subsample using principal axis factoring with Promax rotation, with factor retention determined by eigenvalues, scree plot evaluation, and theoretical relevance. KMO and Bartlett's test were used to confirm sampling adequacy. The identified factor structure was subsequently tested using CFA in an independent subsample, with model fit assessed using χ^2/df , CFI, TLI, RMSEA, and SRMR. Multi-group CFA was conducted to examine measurement invariance across sex by testing configural, metric, and scalar models, with changes in CFI and RMSEA used for model comparisons. Statistical significance was set at $p < 0.05$.

3. Results

3.1. Participant Characteristics

A total of 458 children participated in the study, including 241 boys (52.6%) and 217 girls (47.4%). Participants ranged in age from 8 to 12 years, with a mean age of 10.45 ± 1.47 years. No injuries or adverse events were reported during the testing procedures. All participants completed both the test and retest assessments and were included in the final analyses (Table 1).

Table 1. Participant Characteristics.

Variable	Total Sample (n=458)
Age (years)	10.45 ± 1.47
Boys, n (%)	241 (52.6)
Girls, n (%)	217 (47.4)
Height (cm)	141.8 ± 10.9
Weight (kg)	39.6 ± 11.2
BMI (kg/m ²)	19.3 ± 3.7

3.2. Test-Retest Reliability

All PF tests demonstrated good-to-excellent test-retest reliability. ICC ranged from 0.84 to 0.95, with handgrip strength demonstrating the highest relative reliability (ICC = 0.95) and the sit-and-reach test showing the lowest, although still within the good

reliability range (ICC = 0.84). Absolute reliability indices were also favorable across all tests. The SEM and MDC at the 95% confidence level (MDC95) are presented in the original units of each PF test to facilitate practical interpretation. Coefficients of variation ranged from 3.8% to 7.1%, indicating acceptable measurement precision for all assessments (Table 2).

Table 2. Relative and Absolute Reliability of the Persian ALPHA-Fitness Test Battery.

Test	ICC (95% CI)	SEM*	MDC95*	CV (%)
Handgrip Strength	0.95 (0.94–0.96)	1.21 kg	3.35 kg	3.8
Standing Long Jump	0.92 (0.90–0.94)	5.44 cm	15.07 cm	4.9
20-m Shuttle Run	0.89 (0.86–0.92)	0.78 stages	2.16 stages	6.4
4 × 10-m Shuttle Run	0.87 (0.84–0.90)	0.41 s	1.14 s	5.8
Sit-and-Reach	0.84 (0.80–0.88)	1.87 cm	5.18 cm	7.1

*SEM and MDC95 are expressed in the original measurement units of each fitness test.

3.3. Agreement Analysis

Bland-Altman analyses demonstrated acceptable agreement between the initial and repeated assessments for all PF tests. Mean differences between testing occasions were close to zero, indicating no evidence of systematic measurement bias. More than 95% of observations fell within the corresponding limits of agreement. Visual inspection of the Bland-

Altman plots further revealed no apparent heteroscedasticity, as the dispersion of measurement differences remained relatively constant across the range of observed performance values. These findings suggest that measurement error did not increase systematically among children with higher or lower PF levels. Overall, the agreement analyses support the stability and reproducibility of the Persian ALPHA-Fitness Test Battery (Figure 2).

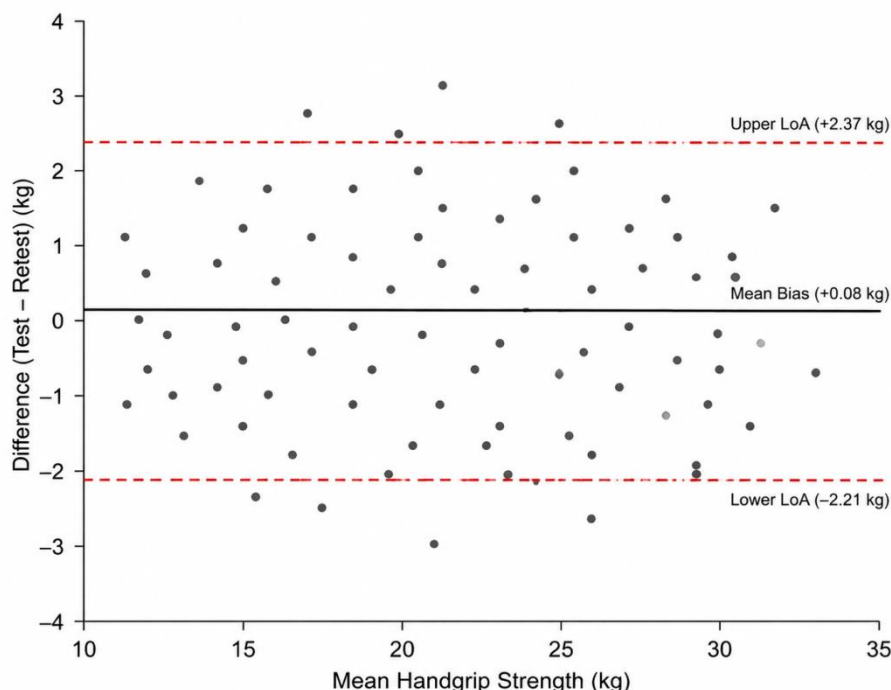


Figure 2. Bland-Altman Plots for Test-Retest Agreement.

3.4. Exploratory Factor Analysis

Participants were randomly allocated into two independent subsamples. The first subsample ($n = 229$) was used for EFA. Sampling adequacy was confirmed by a KMO value of 0.82, while Bartlett's test of sphericity was statistically significant ($\chi^2 = 894.36$, $p < 0.001$) (Table 3). Three factors with eigenvalues greater than 1.0 were

retained, accounting for a cumulative 71.8% of the total variance. Factor 1 (Muscular Fitness) explained 31.6% of the variance (eigenvalue = 2.38), Factor 2 (Functional/Cardiorespiratory Fitness) explained 24.3% (eigenvalue = 1.61), and Factor 3 (Flexibility) explained 15.9% (eigenvalue = 1.12). Communalities ranged from 0.58 to 0.84. No substantial cross-loadings (>0.30) were observed (Table 4).

Table 3. Exploratory Factor Analysis of the Persian ALPHA-Fitness Test Battery..

Variable	Factor 1 Muscular Fitness	Factor 2 Functional/Cardiorespiratory Fitness	Factor 3 Flexibility	Communality (h^2)
Handgrip Strength	0.86	0.14	0.09	0.77
Standing Long Jump	0.81	0.19	0.21	0.73
20-m Shuttle Run	0.18	0.88	0.12	0.82
4 × 10-m Shuttle Run	0.22	0.76	0.25	0.68
Sit-and-Reach	0.11	0.08	0.84	0.71

Table 4. Eigenvalues and Explained Variance

Factor	Eigenvalue	Variance Explained (%)	Cumulative (%)
Muscular Fitness	2.38	31.6	31.6
Functional/Cardiorespiratory Fitness	1.61	24.3	55.9
Flexibility	1.12	15.9	71.8

3.5. Confirmatory Factor Analysis

The second independent subsample ($n = 229$) was used to evaluate the hypothesized three-factor measurement model. The model demonstrated satisfactory goodness-of-fit across all indices (Figure 3 and Table 5). All standardized factor loadings were statistically significant ($p < 0.001$), ranging from 0.69 to 0.91. Correlations among the latent factors were

moderate. Specifically, Muscular Fitness correlated moderately with Functional/Cardiorespiratory Fitness ($r = 0.48$), whereas weaker correlations were observed between Flexibility and Muscular Fitness ($r = 0.24$) and between Flexibility and Functional/Cardiorespiratory Fitness ($r = 0.27$).

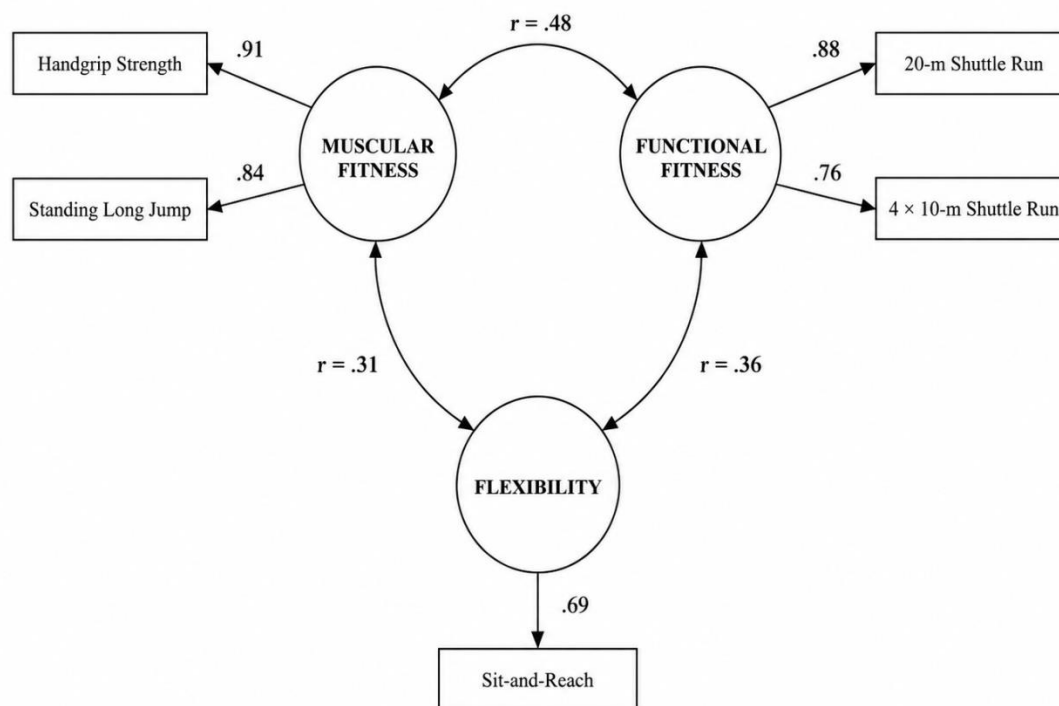


Figure 3. Confirmatory Factor Analysis Model.

Table 5. Confirmatory Factor Analysis Fit Indices

Fit Index	Value	Recommended
χ^2/df	2.14	< 3.00
CFI	0.96	> 0.90
TLI	0.95	> 0.90
RMSEA	0.051	< 0.08
SRMR	0.043	< 0.08

3.6. Measurement Invariance Across Sex

Multi-group CFA demonstrated support for configural, metric, and scalar invariance across boys and girls. Changes in ΔCFI and $\Delta RMSEA$ remained below

recommended thresholds ($\Delta CFI < 0.010$ and $\Delta RMSEA < 0.015$) for each successive model, supporting equivalence of the measurement structure across sex.

Table 6. Measurement Invariance Across Sex

Model	χ^2/df	CFI	RMSEA	ΔCFI	$\Delta RMSEA$	Interpretation
Configural	2.16	0.962	0.052	—	—	Supported
Metric	2.18	0.960	0.053	−0.002	+0.001	Supported
Scalar	2.22	0.957	0.054	−0.003	+0.001	Supported

4. Discussion

The present study examined the psychometric properties of the Persian adaptation of the ALPHA-Fitness Test Battery among Iranian children. The findings demonstrated that the battery possesses good-to-excellent test-retest reliability, acceptable absolute measurement error, and strong evidence of construct validity based on EFA and CFA. Furthermore, multi-group CFA supported configural, metric, and scalar invariance across boys and girls, indicating that the latent structure of the battery was interpreted similarly across sexes. Collectively, these findings provide robust evidence supporting the reliability and construct validity of the Persian adaptation of the ALPHA-Fitness Test Battery for use in school-aged Iranian children.

A major finding of this study was the high level of reliability observed across all PF tests. Handgrip strength demonstrated the highest reliability among all assessments. This finding is consistent with

previous research showing that handgrip dynamometry is one of the most stable field-based measures of muscular fitness in youth (22–24). The relatively simple administration procedure and limited influence of environmental conditions may explain its high reproducibility. The standing long jump and shuttle-run tests also demonstrated strong reliability coefficients, consistent with previous investigations of other batteries (11). Moreover, the low SEM and MDC values indicate that relatively small changes in performance can be interpreted as true changes rather than measurement error, enhancing the practical utility of the battery for repeated assessments in educational settings. The Bland-Altman analyses further strengthened the reliability findings by demonstrating good agreement between testing sessions without evidence of systematic bias or heteroscedasticity. These results indicate that measurement error remained relatively constant across the range of observed PF performances, suggesting that the precision of the battery is

maintained among children with different PF levels.

Construct validity represented another major strength of the present investigation. EFA identified three latent dimensions representing muscular fitness, functional/cardiorespiratory fitness, and flexibility, while CFA supported this structure using an independent validation sample. The emergence of these three factors is theoretically meaningful and aligns with contemporary multidimensional models of youth PF, which recognize that PF comprises several related but distinct physiological domains rather than a single global construct. Muscular fitness reflects neuromuscular capacity and force production, functional/cardiorespiratory fitness encompasses aerobic performance and movement efficiency, whereas flexibility represents an independent physical capacity associated with musculoskeletal function and movement quality. The moderate correlations observed among these latent factors further support the concept that these PF domains are interconnected while maintaining sufficient independence to justify separate assessment. Consequently, the identified factor structure provides meaningful evidence that the Persian adaptation adequately captures the multidimensional nature of PF in children. These findings are comparable to previous validation studies conducted in school-aged populations (25–27). However, relatively few previous studies have combined EFA, CFA, and measurement invariance testing within the same validation framework. Therefore, the present study extends the existing literature by providing a more comprehensive psychometric evaluation of the ALPHA-Fitness Test Battery in an Iranian population.

An important contribution of the present study was the demonstration of measurement invariance across sex. Configural, metric, and scalar invariance were all supported, indicating that the relationships between observed PF tests and the underlying latent constructs were equivalent in boys and girls. This finding suggests that sex-related differences in observed scores are more likely to reflect genuine differences in PF rather than measurement bias or differences in construct interpretation. Similar findings have been reported in previous psychometric evaluations of youth PF batteries, supporting the robustness of PF constructs across sexes. Establishing measurement invariance is particularly important for research and educational applications because it permits meaningful comparisons between boys and girls while strengthening confidence in the fairness of the assessment battery.

From a practical perspective, the Persian adaptation of the ALPHA-Fitness Test Battery offers several advantages for school-based assessment. The required equipment is inexpensive, administration procedures are relatively simple, and testing can be completed efficiently during regular physical education classes. These characteristics make the battery suitable for routine assessment of PF within school environments. Nevertheless, although the present findings support its use for PF assessment, additional evidence regarding criterion validity, predictive validity, and responsiveness to longitudinal changes is required before recommending its application for large-scale surveillance or long-term monitoring programs.

The age range of 8–12 years represents an important developmental period characterized by substantial variability in growth, biological maturation, and neuromuscular development. Such developmental differences may influence PF performance independently of measurement quality. Although the present analyses demonstrated satisfactory psychometric properties across the overall sample, future investigations should examine whether the factorial structure and reliability of the battery remain stable across narrower age groups or biological maturation stages.

The present study has several strengths. These include the relatively large sample size, balanced representation of boys and girls, standardized examiner training, rigorous evaluation of both relative and absolute reliability, and the combined use of EFA, CFA, and multi-group measurement invariance analyses. Together, these methodological features provide a comprehensive evaluation of the construct validity and reliability of the battery.

Several limitations should also be acknowledged. First, although the adaptation process followed accepted recommendations for performance-based assessment tools, formal back-translation was not performed because only the administration instructions and testing materials were linguistically adapted. This should be considered when interpreting the cross-cultural adaptation process. Second, participants were recruited from public schools within selected educational districts, and therefore the findings may not be fully generalizable to all regions of Iran or to private-school populations. Third, only test-retest reliability was evaluated; inter-rater and intra-rater reliability were not examined because the same trained examiners conducted repeated assessments. Fourth, the present study evaluated construct validity but did not investigate criterion validity against laboratory-based PF measures, predictive validity, or responsiveness to intervention-related changes. Finally, the cross-sectional design precludes conclusions regarding the ability of the battery to detect longitudinal changes in PF over time.

Future studies should therefore examine additional aspects of validity, including criterion validity relative to laboratory-based assessments, predictive validity for health outcomes, responsiveness to exercise interventions, and longitudinal stability across different developmental stages. Broader validation studies involving children from diverse geographic, socioeconomic, and educational backgrounds would further strengthen the generalizability of the Persian adaptation.

4.1. Conclusions

The Persian adaptation of the ALPHA-Fitness Test Battery demonstrated good-to-excellent test-retest reliability, low measurement error, and strong evidence of construct validity among Iranian children aged 8–12 years. The identified three-factor structure representing muscular fitness, functional/cardiorespiratory fitness, and flexibility was supported by both EFA and CFA and demonstrated measurement invariance across boys and girls. These findings support the use of the battery as a reliable and psychometrically sound instrument for assessing PF in school-aged Iranian children. Although the present

results provide encouraging evidence for its application in educational and research settings, further studies evaluating additional sources of validity evidence including criterion validity, predictive validity, and responsiveness, are warranted before broader implementation for nationwide surveillance or longitudinal monitoring is recommended.

Authors' Contribution: The author was solely responsible for the study's conception and design, participant recruitment and assessment, data collection, statistical analysis, interpretation of the results, and the preparation and writing of the entire manuscript.

Funding Support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process: The author declares that no generative artificial intelligence (AI) or AI-assisted technologies were used in the writing, editing, proofreading, or generation of any content for this manuscript.

Ethical Considerations: All research procedures were carried out with strict adherence to established ethical principles and in accordance with the Declaration of Helsinki. Written informed consent was obtained from the parents or legal guardians before the children took part in the study. The children were informed about the study procedures, and participation was entirely voluntary. The privacy and confidentiality of all participants were strictly protected throughout the study.

Data Availability: The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: The author gratefully acknowledges the children and their parents or legal guardians for their participation, as well as the school administrators and physical education teachers for their vital cooperation and support during the data collection process.

Conflicts of Interest: The author confirms that no financial, personal, or professional conflict of interest influenced the conduct or reporting of this research.

References

- Ding Y. Physical fitness characteristics and comprehensive physical fitness evaluation model of basketball players based on association rule algorithm. *PLoS One*. 2025;20(7):e0325925. [PubMed ID: 40729331]. [PubMed Central ID: PMC12306734] doi: 10.1371/journal.pone.0325925.
- Wierzbicka-Rucińska A, Wrona A, Szalecki M, Mazur J, Podogrodzki J. Associations Between Physical Fitness and Health-Related Quality of Life in Children with Obesity. *Diagnostics (Basel)*. 2026;16(3):371. [PubMed ID: 41681689]. [PubMed Central ID: PMC12897360] doi: 10.3390/diagnostics16030371.
- Rosi A, Scazzino F, Touriz Bonifaz MA, Giampieri F, Ammar A, Trabelsi K, Abdelkarim O, Aly M, Frias-Toral E, Pons J, Vázquez-Araújo L, Alemany-Iturriaga J, Monasta L, Decembrino N, Mata A, Chacón A, Busó P, Grosso G. Children's and adolescents' lifestyle factors associated with physical activity in five Mediterranean countries: the DELICIOUS project. *Front Public Health*. 2025;13:1654645. [PubMed ID: 41069811]. [PubMed Central ID: PMC12506677] doi: 10.3389/fpubh.2025.1654645.
- Demchenko I, Prince SA, Merucci K, Cadenas-Sanchez C, Chaput JP, Fraser BJ, Manyanga T, McGrath R, Ortega FB, Singh B, Tomkinson GR, Lang JJ. Cardiorespiratory fitness and health in children and adolescents: an overview of systematic reviews with meta-analyses representing over 125 000 observations covering 33 health-related outcomes. *Br J Sports Med*. 2025;59(12):856-865. [PubMed ID: 40101938]. [PubMed Central ID: PMC12229061] doi: 10.1136/bjsports-2024-109184.
- Vandoni M, Gatti A, Carnevale Pellino V, Marin L, Cavallo C, Rossi V, Lascialfari G, Zuccotti G, Calcaterra V. Exploring the link between metabolic syndrome risk and physical fitness in children with obesity: a cross-sectional study. *Eur J Pediatr*. 2025;184(8):497. [PubMed ID: 40705115]. [PubMed Central ID: PMC12289717] doi: 10.1007/s00431-025-06339-7.
- Gesek M, Fornal AD, Zarzycka D. Promoting Health in Pediatric Obesity: A Decade's Research of Physical Activity's Influence on Cardiometabolic Parameters. *Med Sci Monit*. 2023;29:e940742. [PubMed ID: 37771141]. [PubMed Central ID: PMC10546901] doi: 10.12659/MSM.940742.
- Men J, Yu Z, An W, Wang P, Hou X, Zhang Y, Wu S, Zhu G, Wang P, Cui C, Zhang Y, Wang J, Ding J, Wang Y. Effects of exercise on cardiorespiratory fitness in children and adolescents with overweight and obesity: a systematic review and meta-analysis of 72 randomized controlled trials. *BMC Public Health*. 2025;25(1):3899. [PubMed ID: 41219861]. [PubMed Central ID: PMC12607143] doi: 10.1186/s12889-025-25254-y.
- Viskari T, Koivumäki T, Appelqvist-Schmidlechner K, Ståhl T, Ruiz-Ariza A, Fröjd S. The Effectiveness of School-Based Physical Activity Promotion on Mental Health Among Children and Adolescents: A Systematic Review. *Scand J Med Sci Sports*. 2025;35(10):e70150. [PubMed ID: 41111160]. [PubMed Central ID: PMC12536061] doi: 10.1111/sms.70150.
- Krochmal P, Cooper DM, Radom-Aizik S, Lu KD. US School-Based Physical Fitness Assessments and Data Dissemination. *J Sch Health*. 2021;91(9):722-729. [PubMed ID: 34235722]. [PubMed Central ID: PMC9291210] doi: 10.1111/josh.13067.
- Hamdani SMZH, Zhuang J, Hadier SG, Khurram H, Hamdani SDH, Danish SS, Fatima SU, Tian W. Establishment of health-related physical fitness evaluation system for school adolescents aged 12-16 in Pakistan: a cross-sectional study. *Front Public Health*. 2023;11:1212396. [PubMed ID: 37829094]. [PubMed Central ID: PMC10564982] doi: 10.3389/fpubh.2023.1212396.
- Silventoinen K, Maia J, Sillanpää E, Sund R, Gouveia ER, Antunes A, Marques G, Thomis M, Kaprio J, Freitas D. Genetic Regulation of Physical Fitness in Children: A Twin Study of 15 Tests from Eurofit and Fitnessgram Test Batteries. *Med Sci Sports Exerc*. 2024;56(10):2000-2006. [PubMed ID: 38949118]. [PubMed Central ID: PMC12376799] doi: 10.1249/MSS.0000000000003496.
- Ruiz JR, Castro-Piñero J, España-Romero V, Artero EG, Ortega FB, Cuenca MM, Jimenez-Pavón D, Chillón P, Girela-Rejón MJ, Mora J, Gutiérrez A, Suni J, Sjöström M, Castillo MJ. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med*. 2011;45(6):518-24. [PubMed ID: 20961915]. doi: 10.1136/bjsm.2010.075341.
- Jukes MCH, Ahmed I, Baker S, Draper CE, Howard SJ, McCoy DC, Obradović J, Wolf S. Principles for Adapting Assessments of Executive Function across Cultural Contexts. *Brain Sci*. 2024;14(4):318. [PubMed ID: 38671970]. [PubMed Central ID: PMC11047958] doi: 10.3390/brainsci14040318.
- Swan K, Speyer R, Scharitzer M, Farneti D, Brown T, Woisard V, Cordier R. Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. *Front Psychol*. 2023;14:1225850. [PubMed ID: 37790221]. [PubMed Central ID: PMC10543275] doi: 10.3389/fpsyg.2023.1225850.
- Ahmed I, Ishtiaq S. Reliability and validity: Importance in Medical Research. *J Pak Med Assoc*. 2021;71(10):2401-2406. [PubMed ID: 34974579]. doi: 10.47391/JPMA.06-861.
- Smits-Engelsman BCM, Smit E, Doe-Asinyo RX, Lawerteh SE, Aertssen W, Ferguson G, Jelsma DL. Inter-rater reliability and test-retest reliability of the Performance and Fitness (PERF-FIT) test battery for children: a test for motor skill related fitness. *BMC Pediatr*. 2021;21(1):119. [PubMed ID: 33706742]. [PubMed Central ID: PMC7948346] doi: 10.1186/s12887-021-02589-0.
- Dishman RK, Saunders RP, McIver KL, Dowda M, Pate RR. Construct validity of selected measures of physical activity beliefs and motives in fifth and sixth grade boys and girls. *J Pediatr Psychol*. 2013;38(5):563-76. [PubMed ID: 23459310]. [PubMed Central ID: PMC3716273] doi: 10.1093/jpepsy/jt013.
- Ke D, Wang D, Huang H, Hu X, Sasaki J, Liu H, Wang X, Lu D, Wang J, He G. Study of the Reliability of Field Test Methods for Physical Fitness in Children Aged 2-3 Years. *Int J Environ Res Public Health*. 2022;19(12):7522. [PubMed ID: 35742771]. [PubMed Central ID: PMC924297] doi: 10.3390/ijerph19127522.
- Brazo-Sayavera J, Silva DR, Lang JJ, Tomkinson GR, Agostinis-Sobrinho C, Andersen LB, García-Hermoso A, Gaya AR, Jurak G, Lee EY, Liu Y, Lubans DR, Okely AD, Ortega FB, Ruiz JR, Tremblay MS, Dos Santos L. Physical Fitness Surveillance and Monitoring Systems Inventory for Children and Adolescents: A Scoping Review with a Global Perspective. *Sports Med*. 2024;54(7):1755-1769. [PubMed ID: 39242971] doi: 10.1080/17513758.2024.2342971.

- 38710913]. [PubMed Central ID: [PMC11258155](#)] doi: [10.1007/s40279-024-02038-9](#).
20. Soto-Sánchez J, Leyton-Dinamarca B, Kain J, Fernández-Valero P, Castro-Cisterna S, Arias-Tellez MJ, Ortega FB. Reliability of the PREFIT fitness-test battery in Chilean preschoolers. *Front Pediatr*. 2026;13:1654731. [PubMed ID: [41602872](#)]. [PubMed Central ID: [PMC12833291](#)] doi: [10.3389/fped.2025.1654731](#).
 21. Zamani MH, Hashemi A, Siavashi E, Khanmohamadi R, Saeidi H. Validity and Reliability of the Fundamental Motor Skills in Sports (FUS) Test for Iranian Children and Adolescents. *Int J School Health*. 2024;11(3):157-169. [PubMed ID: [42518322](#)]. [PubMed Central ID: [PMC13342030](#)] doi: [10.30476/INTJSH.2024.101089.1368](#).
 22. Zárate-Osuna F, Zapico AG, González-Gross M. Handgrip Strength in Children and Adolescents Aged 3 to 16 Years and Residing in Spain: New Reference Values. *Children (Basel)*. 2025;12(4):471. [PubMed ID: [40310105](#)]. [PubMed Central ID: [PMC12026169](#)] doi: [10.3390/children12040471](#).
 23. Karagianni E, Antoniou V, Dimitriadis Z, Panagiotakos D, Cordovil R, Pepera G. Assessment of Overall Muscle Strength in Children and Adolescents Using Handheld Dynamometry: A Systematic Review of Reference Values and Quality of Data. *J Clin Med*. 2025;14(23):8454. [PubMed ID: [41375762](#)]. [PubMed Central ID: [PMC12693157](#)] doi: [10.3390/jcm14238454](#).
 24. García-Hermoso A, Yáñez-Sepúlveda R, Hormazábal-Aguayo I, Muñoz-Pardeza J, Martínez-Vizcaino V, Hurtado-Almonacid J, Ezzatvar Y. Handgrip Strength Thresholds to Detect Cardiometabolic Risk in Youth: Cross-Sectional Study and Meta-Analysis. *J Cachexia Sarcopenia Muscle*. 2025;16(5):e70091. [PubMed ID: [42207172](#)]. [PubMed Central ID: [PMC12536247](#)] doi: [10.1002/jcsm.70091](#).
 25. Bird M, Barnett TA, Fuller D, Chinerman D, Mathieu ME, Datta GD. Multidimensional school features associated with physical activity among youth at risk of obesity: an exploratory principal component and generalized estimating equation analysis. *BMC Public Health*. 2023;23(1):2010. [PubMed ID: [37845705](#)]. [PubMed Central ID: [PMC10577956](#)] doi: [10.1186/s12889-023-16889-w](#).
 26. Cuberek R, Janíková M, Dygrýn J. Adaptation and validation of the Physical Activity Questionnaire for Older Children (PAQ-C) among Czech children. *PLoS One*. 2021;16(1):e0245256. [PubMed ID: [33434223](#)]. [PubMed Central ID: [PMC7802927](#)] doi: [10.1371/journal.pone.0245256](#).
 27. Aly M, Ammar A, Trabelsi K, Masmoudi L, El-Gyar N, Shalaby AM, Abdelkarim O. School-based physical activity and health-related fitness in Mediterranean students: findings from the DELICIOUS project. *Front Public Health*. 2025;13:1603043. [PubMed ID: [40529715](#)]. [PubMed Central ID: [PMC12170614](#)] doi: [10.3389/fpubh.2025.1603043](#).