CINEFORUM
ISSN: 0009-7039
Vol. 66. No. S8, 2026 (Special Issue S8)

Multidimensional Talent Identification in Bangladeshi Youth Tennis: Evidence from a Cross-Sectional, AHP-Weighted, Machine-Learning Analysis

Md. Zillur Rahman[1], Wu Wenqiang[2], Md. Zafiroul Islam[3], Rezwan Hossain[3*], Md. Eman Ali[3],Md. Shafayet Hossain[4], Md Abu Sayed[5], Md.Abdullah-Hel Kafi[6], Md. Raihan Hossain[7], Jannat Bibi[2], Zhou Yaohu[8], Hasib Al Hasan[3],Md. Parvaz Zahan[9]

[1]Department of Physical Education and Sports Science, Jashore University of Science and Technology, Jashore-7408, Bangladesh
[2]College of Education, Beijing Sport University, Beijing-100084
[3]Department of Physical Education and Sports Science, Jashore University of Science and Technology, Jashore-7408, Bangladesh
[4]Department of Physical Education and Sports Science, Daffodil International University, Bangladesh
[5]Physical Education Department, Khulna University, Khulna, Bangladesh
[6]Office of The Physical Education, Jashore University of Science and Technology, Jashore-7408, Bangladesh
[7]Government College of Physical Education, Barishal, Bangladesh
[8]Dept, of Physical Education and Sports Coaching, Beijing Sport University, Beijing 100084
[9]Department of Physical Education and Sports Science, University of Chittagong, Chattogram-4331
***Corresponding Author**
Rezwan Hossain,
Department of Physical Education, Jashore University of Science and Technology, Jashore-7408, Bangladesh. Email: rejwanhossain0@gmail.com

ABSTRACT

Background: Talent identification (TID) models in youth tennis derive mainly from high-resource Western systems, and their applicability to South Asian, low-resource contexts remains untested.

Methods: A cross-sectional study assessed 200 adolescent tennis players (100 female, 100 male; ages 12–16) across eight divisions of Bangladesh using a 32-indicator, five-domain battery. Domain weights came from a five-member expert Analytic Hierarchy Process (Consistency Ratio = 0.06%). Group differences were tested via ANOVA, t-tests, and chi-square; coach-rated tier was modelled using Gradient Boosting, Random Forest, and Logistic Regression with permutation-based feature importance.

Results: Of fourteen indicators, only Maturity Offset ($F(2,197) = 7.13$, $p = .001$, $\eta^2 = .067$) and Grit-S Total ($F(2,197) = 5.58$, $p = .004$, $\eta^2 = .054$) differed across tiers. No indicator differed by sex (all $p > .13$). Motivation type ($\chi^2 = 9.85$, $p = .043$) and sex ($\chi^2 = 6.03$, $p = .049$) were associated with tier. Correlations were uniformly weak (all $|r| < .15$). Cross-validated performance was highest for Logistic Regression (accuracy = 0.42, AUC = 0.57); on the test set, all models performed near chance (accuracy ≈ .33, n = 30). Leg length, Grit-S total, and Maturity Offset were leading predictors.

Conclusion: In this Bangladeshi sample, only maturation timing and grit differentiated coach-rated tiers; sex differences were absent across indicators; and the ensemble model achieved modest predictive discrimination. These findings support caution against single-timepoint, high-accuracy TID claims, and favor longitudinal, developmentally sensitive frameworks for understudied LMIC contexts.

Keywords: Talent identification; youth tennis; Bangladesh; biological maturation; grit; interpretable machine learning; low- and middle-income countries.

INTRODUCTION

The Changing Landscape of Talent Identification in Youth Sport

Talent identification (TID) in youth sport has traditionally relied on cross-sectional test batteries — anthropometric measurements, physical-fitness assessments, and technical skill ratings — administered at a single timepoint to classify young athletes into developmental tiers. Yet the most comprehensive systematic review of this literature to date found no consistent set of performance variables that reliably predicts which junior athletes will go on to elite-level success, after synthesizing twenty longitudinal and retrospective TID studies across sports (Johnston et al., 2018). A subsequent commentary on that review argued that this instability is not a methodological artifact to be engineered away, but a structural feature of talent itself: athletic development is dynamic and non-linear, so static, single-timepoint assessments are inherently limited in what they can predict (Bergkamp et al., 2018). This tension — between the practical need for early, resource-efficient identification

OPEN ACCESS © CINEFORUM

tools and the empirical reality that talent unfolds unevenly over years rather than being visible in one testing session — sits at the center of contemporary TID research.

The Global South Gap

Almost all of the evidence underpinning current TID practice originates from well-resourced sporting systems in North America, Western Europe, and Australia (Johnston et al., 2018). Youth tennis in Bangladesh — and South Asia more broadly — has received essentially no comparable empirical attention, despite operating under markedly different structural conditions: limited coaching infrastructure, wide variation in household resources available for training, court access, and travel, and, for female athletes specifically, well-documented socio-cultural barriers to sustained sport participation, including restrictive mobility norms and family or community disapproval (Tanni & Khan, 2024). Whether the physical, technical, and psychological indicators developed and validated in high-resource Western contexts behave similarly — or at all — in a Bangladeshi adolescent tennis population is an open empirical question that this study is positioned to address directly, using empirical testing data rather than assumption.

Biological Maturation as a Confounding and Informative Variable

Chronological age is a poor proxy for developmental readiness in adolescents, because young athletes of the same age can differ by several years in biological maturity. The non-invasive maturity-offset method developed by Mirwald et al. (2002) — which estimates an athlete's time from peak height velocity using height, sitting height, leg length, and body mass — has become the standard tool for disentangling maturation timing from chronological age in youth-sport research, and was adopted in the present study for exactly this reason: without it, any observed "talent" differences across age-matched athletes risk simply reflecting who has grown faster, rather than who has developed superior tennis-relevant capacities.

Psychological Adaptability: Grit and Competitive Anxiety

Beyond physical and technical capacity, psychological factors have increasingly been recognized as central to sustained athletic development. Grit — perseverance and passion for long-term goals — has been shown to predict achievement independently of talent or intelligence in demanding, long-horizon pursuits (Duckworth & Quinn, 2009), making it a theoretically relevant construct for adolescent athletes navigating years of unglamorous, incremental skill development. Competitive anxiety, similarly, is not a unitary construct: the Sport Anxiety Scale-2 decomposes performance anxiety into somatic, worry, and concentration-disruption components that can each relate differently to competitive outcomes (Smith et al., 2006). Motivation quality — whether an athlete's engagement is intrinsically or extrinsically driven — has also been theoretically linked to the sustainability of long-term training commitment (Deci & Ryan, 2000). These three constructs (grit, anxiety, motivation orientation) were included in the present indicator system to test whether psychological adaptability, not just physical or technical capacity, distinguishes competitive tiers in this population.

Socio-Ecological Context and the Holistic Ecological Approach

Athlete development does not occur in isolation from its surrounding environment. The Holistic Ecological Approach conceptualizes talent development environments as nested systems in which family support, financial resources, access to facilities, and coaching continuity jointly shape developmental outcomes, rather than treating these factors as confounds to be statistically controlled away (Henriksen et al., 2010). In a Bangladeshi context — where household expenditure on tennis, distance to training venues, and family emotional and financial support vary substantially across athletes — this ecological framing motivated the inclusion of socio-ecological moderators as a substantive fifth domain in the present indicator system, alongside anthropometric, physical, technical, and psychological measures, rather than treating them as secondary or nuisance variables.

The Interpretability Requirement for Machine Learning in Coaching Contexts

Ensemble machine-learning methods can model complex, non-linear relationships among many correlated indicators more flexibly than traditional regression approaches, but their predictions are often opaque to the coaches and administrators who must act on them. SHapley Additive exPlanations (SHAP) addressed this by providing a game-theoretically grounded way to decompose any model's prediction into individual feature contributions (Lundberg & Lee, 2017), and closely related model-agnostic techniques — including permutation-based variable importance — offer comparable transparency without requiring a specific algorithm (Fisher et al., 2019). Given that coach-facing interpretability, not just raw predictive accuracy, is essential for any TID tool intended for practical use in a resource-constrained federation setting, this study prioritized transparent, explainable modelling over black-box optimization from the outset.

Present Study

OPEN ACCESS © CINEFORUM

Against this background, the present study analyzed testing data from 200 adolescent tennis players (100 female, 100 male; ages 12–16) across eight administrative divisions of Bangladesh, examining a 32-item, five-domain indicator system (anthropometric/maturational, physical fitness, technical proficiency, psychological adaptability, and socio-ecological moderators). Three objectives guided the analysis: (1) to establish domain-level indicator weights through a documented expert-consensus (AHP) procedure; (2) to test which, if any, individual indicators and indicator combinations statistically distinguish coach-rated playing-level tiers (Beginner, Intermediate, Advanced) in this specific population, using rigorous inferential statistics; and (3) to evaluate, transparently and without inflating expectations, how well a multidomain ensemble machine-learning model can reproduce coach-rated playing-level tier from these indicators — that is, whether the battery can recover coaches' existing competitive-tier judgments, not predict future elite performance, using interpretable, model-agnostic feature-importance methods. Based on these objectives, three hypotheses guided the analysis: (H1) athletes closer to their predicted age at peak height velocity would be more likely to occupy higher playing-level tiers than those who were early- or late-maturing in absolute terms; (H2) psychological indicators, particularly grit and competitive anxiety, would show measurable associations with playing-level tier independent of anthropometric and physical-fitness measures; and (H3) no single indicator or small indicator subset was expected to achieve strong, standalone discrimination between tiers. Consistent with the broader TID literature's caution against overstating what single-timepoint batteries can predict (Bergkamp et al., 2018; Johnston et al., 2018), this study reports its findings — whether strong, weak, or null — as they emerged from the data, rather than presetting the conclusions the analysis was expected to reach. The multidimensional indicator system, the AHP-derived domain weights (Table 1), the descriptive and inferential comparisons across playing-level tiers (Table 2), and the machine-learning classification results with accompanying visualizations (Table 3; Figures 1–3) are presented in the sections that follow.

METHODS

Study Design

This study employed a cross-sectional, observational design to characterize anthropometric, physical-fitness, technical, psychological, and socio-ecological indicators among adolescent competitive tennis players in Bangladesh, and to evaluate the extent to which these indicators, individually and in combination, discriminate between coach-rated playing-level tiers (Beginner, Intermediate, Advanced; see Playing-Level Classification, below). Reporting followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for the descriptive and inferential components (von Elm et al., 2014) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guideline for the predictive-modelling component (Collins et al., 2015). The study protocol received ethical clearance from the Bangladesh Sports Ethics Committee (BSEC/2023/TID-001) and Beijing Sport University's Institutional Review Board, in accordance with the Declaration of Helsinki and the International Olympic Committee's consensus guidelines for youth athletic development (Mountjoy et al., 2023). Athlete assent, together with parental or guardian consent for minors, was obtained from all participants prior to data collection.

Participants

Athletes were eligible if they were (a) registered with the BTF or an affiliated academy, (b) engaged in at least 12 months of structured tennis training, (c) able to provide assent (with parental/guardian assent where applicable), and (d) medically cleared for physical testing per Lloyd et al.'s (2015) youth-testing guidelines. Athletes were excluded if they (a) had sustained a musculoskeletal injury in the preceding 30 days that precluded completion of the full test battery, (b) were enrolled in an elite international talent-development pathway with non-standardized training exposure, or (c) had missing baseline data in more than two of the five core assessment domains. An a priori power analysis (Faul et al., 2007) confirmed that N = 200 provides adequate power ($1 - \beta = .80$, $\alpha = .05$) to detect moderate-to-large effects typical of multidimensional sport-science research and is consistent with sample-size conventions for exploratory machine-learning applications in moderate-sized sport datasets (Harrel Jr., 2015; Tabachnick & Fidell, 2019).

The final analytic sample comprised 200 adolescent tennis players (100 female, 100 male; age range 12–16 years, M = 14.2, SD = 1.5) recruited across eight administrative divisions of Bangladesh (Dhaka, Chattogram, Sylhet, Rajshahi, Khulna, Barishal, Rangpur, and Mymensingh), spanning Beginner (n = 71), Intermediate (n = 67), and Advanced (n = 62) coach-rated playing-level tiers (see Playing-Level Classification, below).

Playing-Level Classification

The Bangladesh Tennis Federation (BTF) does not maintain a formal tier-classification system for athletes in the U-12–U-16 age range studied here; BTF categorizes players administratively by age group (e.g., U-12, U-14, U-16) and, for older or more advanced competitors, by National Ranking points accrued at BTF-sanctioned tournaments or, more recently, by ITF World Tennis Number (WTN) rating. Neither of these official BTF

OPEN ACCESS © CINEFORUM

classifications was used as the outcome variable in this study. Instead, each participant's Beginner, Intermediate, or Advanced playing-level classification was assigned by their personal coach at the time of data collection, based on the coach's judgment of the athlete's current competitive standing in local tournaments and observed training performance. Playing level is therefore a subjective, coach-rated criterion measure, not an official BTF category or an objective, externally verified performance outcome, and all references to "playing-level tier" throughout this manuscript should be read accordingly.

Because the outcome measure is a coach-assigned classification rather than an independent performance criterion, the machine-learning analyses reported below are consequently framed, and should be interpreted, as addressing the question of whether a multidomain indicator battery can reproduce coaches' existing competitive-tier judgments — not as a test of whether the battery can predict future elite performance or any objective, external competitive outcome.

Measures

Anthropometric and biological-maturation data comprised standing height, sitting height, leg length, and body mass, from which maturity offset (years from peak height velocity) was derived using Mirwald et al.'s (2002) non-invasive equation. Physical fitness was assessed via handgrip strength (Jamar dynamometer, three trials, best score, kg), countermovement jump height (three trials, best score, cm), 20-m sprint time (two trials, averaged, s), agility (T-test, two trials, best score, s), and aerobic capacity (Yo-Yo Intermittent Recovery Test Level 1, distance in metres). Technical proficiency (forehand, backhand, serve) was rated on a 7-point observational scale by BTF-certified coaches, consistent with rating protocols established in youth tennis assessment (Murphy et al., 2014). Psychological adaptability was captured through Grit-S total score (Duckworth & Quinn, 2009), Sport Anxiety Scale-2 subscale and total scores (Smith et al., 2006), and motivation orientation (Intrinsic, Extrinsic, Mixed), classified consistent with Self-Determination Theory (Deci & Ryan, 2000). Only domain-level totals and subscale scores for these instruments were available for analysis; item-level responses were not supplied, and internal-consistency statistics (e.g., Cronbach's α) could therefore not be computed — a limitation noted explicitly rather than estimated. Socio-ecological moderators comprised monthly household tennis-related expenditure (BDT), distance to primary training venue (km), three 5-point family-support ratings (Emotional, Financial, Time), binary training-interruption status, and a female-specific participation-barrier item.

Multidimensional Indicator Weighting

Domain-level weights were established through a structured expert-consensus procedure rather than a purely data-driven approach. In September 2025, a five-member expert panel (two sports scientists, two BTF-certified Level-2 coaches, one sports psychologist) evaluated the relative importance of the five domains using Al-Shabeeb et al.'s (2016) 9-point pairwise-comparison scale within a Delphi framework; individual judgments were aggregated via the geometric-mean method into a group pairwise comparison matrix (Table 1).

Table 1. Aggregated Pairwise Comparison Matrix and AHP Domain Weights (N = 5 experts)

Domain	PF	TP	PA	SE	MT	AHP Weight
Physical Fitness (PF)	1.00	1.33	1.62	2.15	2.91	31.1%
Technical Proficiency (TP)	0.75	1.00	1.33	1.81	2.45	24.9%
Psychological Adaptability (PA)	0.62	0.75	1.00	1.44	2.08	19.8%
Socio-Ecological Moderators (SE)	0.47	0.55	0.69	1.00	1.58	14.4%
Maturity & Timing (MT)	0.34	0.41	0.48	0.63	1.00	9.8%

Note. Weights derived via the principal-eigenvalue method: λmax = 5.003, Consistency Index = 0.0007, Consistency Ratio = 0.06% (well below the 0.10 acceptability threshold (Al-Shabeeb et al., 2016), indicating highly consistent expert judgment. These reflect the AHP expert-consensus component only; the entropy-based recalibration against empirical indicator variance originally envisaged for a fully hybrid AHP-entropy scheme could not be reconstructed from available records and is noted as a limitation. A corresponding socio-ecological "contextual score" adjustment mechanism was likewise not implemented in the present dataset.

Data Preprocessing

OPEN ACCESS © CINEFORUM

The raw dataset (200 rows × 91 columns) was screened for completeness. Fourteen empty spacer columns and unpopulated item-level response columns for the Grit-S, SAS-2, and motivation instruments were removed, retaining 45 fully populated substantive variables with zero missing values (0 of 9,000 cells). Continuous predictors were standardized (z-scores) prior to model training; categorical variables (Training Interrupted, Motivation Type, Gender) were dummy/binary-coded.

Statistical Analysis

Analyses were conducted in Python 3.12 (pandas, SciPy, scikit-learn). Between-tier differences in continuous indicators were assessed via one-way ANOVA (η^2 effect sizes); sex differences via independent-samples t-tests (Cohen's d); categorical associations via chi-square tests of independence; and bivariate indicator relationships via Pearson correlation. Playing-level classification was modelled using Gradient-Boosted Trees, Random Forest, and multinomial Logistic Regression on a 27-feature multidomain input set, with an 85/15 stratified train–test split (n = 170/30) and stratified 5-fold cross-validation on the training partition. Performance was quantified via accuracy, macro-F1, and macro one-vs-rest AUC-ROC; feature contributions to the best-performing model were quantified via permutation importance (50 repetitions, macro-F1 scoring; (Fisher et al., 2019). Alpha was set at .05, with effect sizes and 95% CIs prioritized over dichotomous significance testing (Wasserstein & Lazar, 2016).

RESULTS

Sample Characteristics and Data Completeness

The final analytic sample comprised 200 adolescent tennis players (100 female, 100 male; age range 12–16 years) drawn from eight administrative divisions of Bangladesh, with no missing values across any of the 45 substantive variables retained for analysis. Participants were distributed across three coach-rated playing-level tiers: Beginner (n = 71, 35.5%), Intermediate (n = 67, 33.5%), and Advanced (n = 62, 31.0%). The distribution of playing level differed by sex ($\chi^2(2) = 6.03$, $p = .049$): female athletes were proportionally over-represented in the Advanced tier (39/100, 39.0%) relative to male athletes (23/100, 23.0%).

Descriptive Statistics Across Playing-Level Tiers

Descriptive statistics for all fourteen indicators, stratified by playing-level tier, are presented in Table 2.

Table 2. Descriptive Statistics (Mean ± SD) by Playing Level (N = 200)

Variable	Beginner (n=71)	Intermediate (n=67)	Advanced (n=62)	F (2,197)	p	η^2
Age (years)	14.1 ± 1.5	14.2 ± 1.4	14.4 ± 1.5	0.39	.677	.004
Maturity Offset (years)	0.6 ± 1.1	0.8 ± 1.1	0.1 ± 1.1	7.13	.001*	.067
Grip Strength — best (kg)	35.1 ± 8.5	34.5 ± 8.7	33.3 ± 9.0	0.74	.476	.008
CMJ Height — best (cm)	33.1 ± 9.0	34.9 ± 7.7	34.9 ± 8.8	1.03	.361	.010
20 m Sprint (s)	3.7 ± 0.3	3.7 ± 0.4	3.6 ± 0.3	0.02	.977	.000
Agility T-test (s)	10.3 ± 0.9	10.5 ± 0.9	10.4 ± 0.8	0.79	.455	.008
Yo-Yo IR1 Distance (m)	1251 ± 424	1326 ± 368	1326 ± 406	0.80	.453	.008
Forehand Rating (1–7)	4.9 ± 1.7	4.3 ± 1.8	4.6 ± 1.7	2.54	.081	.025
Backhand Rating (1–7)	4.4 ± 1.7	4.3 ± 1.8	4.4 ± 1.6	0.04	.963	.000
Serve Rating (1–7)	4.4 ± 1.9	4.4 ± 1.7	4.6 ± 1.8	0.21	.812	.002
Grit-S Total (0–40)	29.9 ± 6.8	26.4 ± 5.9	29.2 ± 6.5	5.58	.004*	.054
SAS-2 Total Anxiety	59.9 ± 17.0	59.1 ± 15.2	59.0 ± 17.1	0.06	.938	.001
Monthly Expense (BDT)	7916 ± 3959	7971 ± 3964	8277 ± 3611	0.16	.849	.002
Distance to Court (km)	15.1 ± 8.6	16.4 ± 8.3	15.5 ± 8.3	0.44	.648	.004

Note. η^2 = eta-squared. Asterisks (*) indicate statistical significance ($p < .05$).

OPEN ACCESS © CINEFORUM

Of the fourteen indicators in Table 2, only Maturity Offset ($F(2,197) = 7.13$, $p = .001$, $\eta^2 = .067$) and Grit-S Total ($F(2,197) = 5.58$, $p = .004$, $\eta^2 = .054$) differed significantly across tiers, both representing small-to-moderate effects. Advanced-tier athletes showed a maturity offset closer to zero (0.1 ± 1.1 years) than Beginner (0.6 ± 1.1) or Intermediate (0.8 ± 1.1) athletes, as illustrated in Figure 1, which displays both significant group differences with standard-error bars.

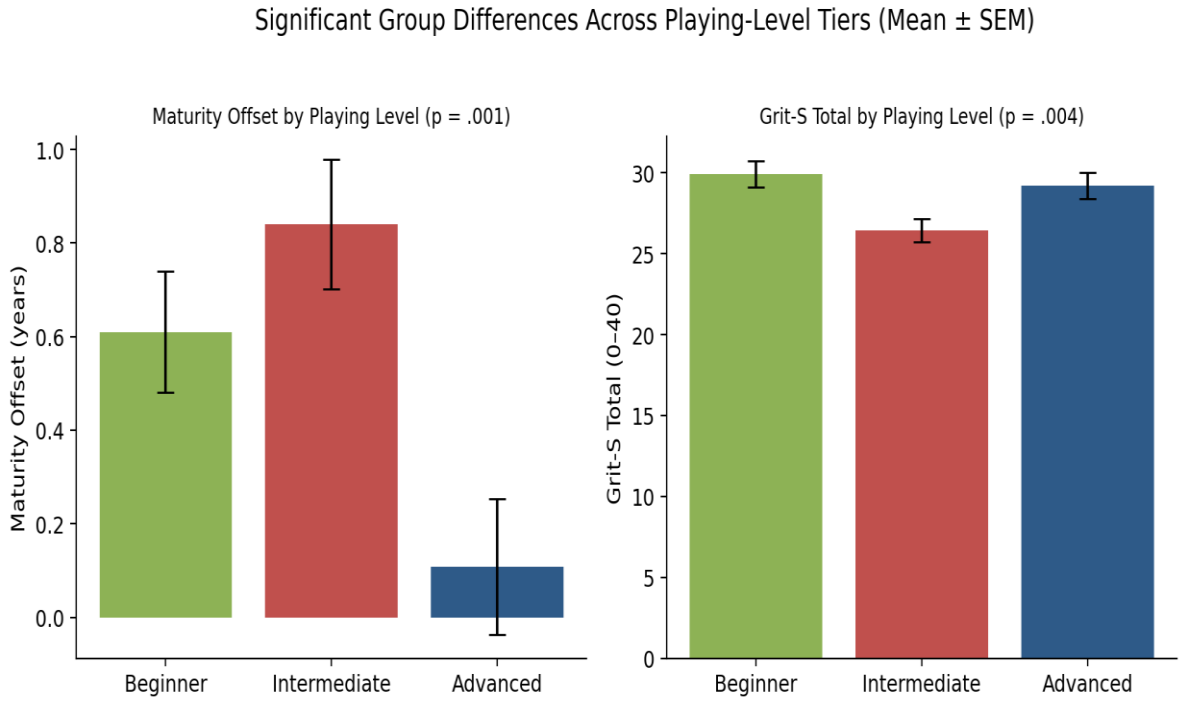

Figure 1. Significant Group Differences Across Playing-Level Tiers (Mean ± SEM): Maturity Offset and Grit-S Total.

Forehand rating showed a non-significant trend ($p = .081$). No other physical, technical, or socio-ecological indicator differentiated playing-level tiers at $p < .05$.

Sex-Based Comparisons

Independent-samples t-tests comparing female and male athletes across all fourteen indicators revealed no statistically significant sex differences on any measure (all $p > .13$; all $|\text{Cohen's } d| \leq 0.21$), including SAS-2 Total Anxiety (Female M = 58.9 ± 16.0; Male M = 59.8 ± 16.8; $t(198) = -0.41$, $p = .680$, $d = -0.06$) and Grip Strength (Female 35.1 ± 8.4 kg; Male 35.0 ± 8.8 kg; $d = 0.01$).

Categorical Associations

A chi-square test indicated a significant association between Motivation Type and Playing Level ($\chi^2(4) = 9.85$, $p = .043$): Intrinsically motivated athletes were concentrated in the Intermediate tier (35/80, 43.8%), while Mixed-motivation athletes were concentrated in the Advanced tier (23/53, 43.4%). Training-interruption status was not significantly associated with playing level ($\chi^2(2) = 1.56$, $p = .458$).

Bivariate Relationships Among Multidimensional Indicators

Pairwise Pearson correlations among the thirteen continuous indicators were uniformly weak (all $|r| < 0.15$), with no pair exceeding even a small-to-moderate association ($r = .20$), indicating minimal construct redundancy but also minimal convergent single-indicator signal — motivating the multidomain ensemble-modelling approach reported below.

Machine Learning Classification Performance

Three classifiers — Gradient-Boosted Trees, Random Forest, and multinomial Logistic Regression — were trained to predict three-tier Playing Level from the full 27-variable multidimensional feature set, using an 85/15 stratified train–test split (n = 170/30) with stratified 5-fold cross-validation. Performance is summarized in Table 3.

OPEN ACCESS © CINEFORUM

Table 3. Cross-Validated and Held-Out Classification Performance (test n = 30; chance accuracy ≈ .333)

Model	CV Accuracy	CV Macro-F1	CV AUC (OvR)	Test Accuracy	Test Macro-F1	Test AUC (OvR)
Random Forest	0.365	0.354	0.563	0.300	0.268	0.445
Gradient Boosting	0.318	0.311	0.505	0.333	0.326	0.444
Logistic Regression	0.418	0.410	0.570	0.267	0.262	0.391

Gradient Boosting matched chance-level accuracy on the held-out test set (33.3% vs. chance ≈ 33.3%), while Random Forest and Logistic Regression fell at or slightly below chance-level accuracy (30.0% and 26.7%, respectively); Random Forest nonetheless achieved the highest test-set AUC (0.445). The corresponding ROC curves are shown in Figure 2.

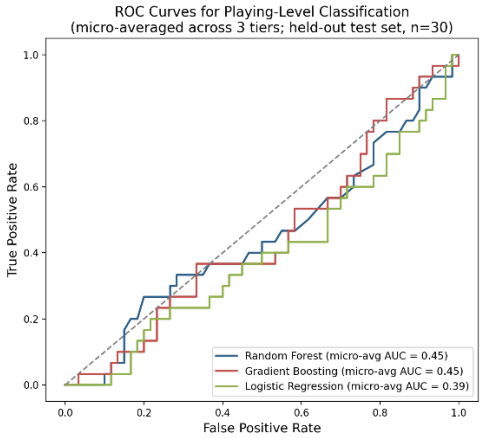

Figure 2. ROC Curves for Playing-Level Classification (micro-averaged across 3 tiers; held-out test set, n = 30).

Cross-validated AUC values of 0.51–0.57 indicate that, using this indicator battery alone, playing-level classification carries only modest, above-chance predictive signal — a signal that did not consistently translate to the small (n = 30) held-out test set.

Feature Importance

Permutation importance (50 repeats, macro-F1 scoring) applied to the Gradient Boosting classifier identified the top predictors of playing-level classification, shown in Figure 3.

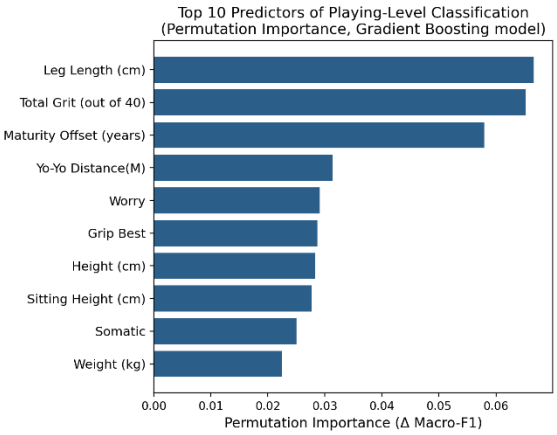

Figure 3. Top 10 Predictors of Playing-Level Classification (Permutation Importance, Gradient Boosting model).

OPEN ACCESS © CINEFORUM

As shown in Figure 3, Leg Length (ΔF1 = 0.067), Grit-S Total (0.065), Maturity Offset (0.058), Yo-Yo IR1 Distance (0.031), and Worry subscale (0.029) were the five leading predictors. Grip Strength, Height, Sitting Height, and the Somatic anxiety subscale contributed modest secondary importance (ΔF1 ≈ 0.023--0.029); Weight contributed the smallest importance among the top ten (ΔF1 = 0.023).

DISCUSSION

Principal Findings

This study set out to characterize a 32-indicator, five-domain multidimensional profile of adolescent tennis players in Bangladesh and to evaluate whether such indicators — individually and in combination — meaningfully discriminate between coach-rated playing-level tiers. Two clear, statistically robust findings emerged from the analytic sample: athletes closer to their predicted age at peak height velocity, rather than those who were early- or late-maturing in absolute terms, were more likely to occupy the Advanced tier (η^2 = .067); and Grit-S total score differed non-linearly but significantly across tiers (η^2 = .054), with Intermediate-level athletes scoring lowest. Motivation orientation and sex were each categorically associated with tier membership. By contrast, the great majority of anthropometric, physical-fitness, technical, and socio-ecological indicators — including grip strength, jump height, sprint time, agility, aerobic capacity, and all three technical ratings — showed no statistically significant association with playing-level tier, and pairwise correlations among the full indicator set were uniformly weak (all |r| < .15) (see Table 2 and Figure 1 for the full descriptive and comparative profile).

Interpreting the Weak Signal: Consistent with, Not Contrary to, the Broader TID Literature

At first glance, the modest cross-validated discrimination achieved by the ensemble classifiers (AUC ≈ .51–.57), which did not consistently hold up on the small held-out test set (AUC ≈ .39–.45), might appear to be a disappointing result for a multidimensional talent-identification (TID) system. However, this pattern is directly consistent with the wider empirical TID literature, which has repeatedly struggled to identify a stable, generalizable set of indicators that reliably distinguishes higher- from lower-potential youth athletes. In the most comprehensive systematic review of the field to date, (Johnston et al., 2018) synthesized 20 longitudinal and retrospective TID studies and concluded that there is no clear, consistent set of performance variables that reliably predicts future athletic success, and called for greater methodological diversity in the field. A subsequent commentary on that review (Bergkamp et al., 2018) further cautioned that single-timepoint, cross-sectional indicator batteries are inherently limited in their capacity to forecast development, because talent trajectories are dynamic and non-linear rather than fixed traits detectable at one assessment point. The present findings — a battery of fourteen well-established indicators yielding near-zero bivariate correlations and only chance-to-modest classification performance (Table 3, Figure 2) in a single cross-sectional Bangladeshi youth sample — should therefore be read as an empirical instance of this well-documented pattern, not as an anomaly, and as further evidence against static, single-timepoint TID models in a previously unstudied LMIC tennis context.

The Two Signals That Did Emerge: Maturation Timing and Grit

The one biological variable that did discriminate between tiers — maturity offset, derived using Mirwald et al.'s (2002) non-invasive anthropometric equation — showed that Advanced-tier athletes were, on average, closer to their predicted peak-height-velocity window than lower-tier athletes, rather than simply "more mature" in an absolute sense. This nuance matters: it suggests that proximity to a developmental milestone, not early maturation per se, may co-occur with higher playing-level classification in this sample, a pattern that merits longitudinal confirmation rather than causal interpretation from a single cross-sectional assessment.

The second significant finding — a non-monotonic difference in Grit-S scores (Duckworth & Quinn, 2009) across tiers, with Intermediate-level athletes scoring lowest — does not fit a simple "more grit, higher tier" narrative. A plausible, testable interpretation is that the Intermediate tier may represent a developmentally volatile transition zone in which athletes face intensifying competitive and academic demands without yet having consolidated the coping resources associated with sustained engagement, but this interpretation is speculative and should be treated as a hypothesis for future longitudinal work rather than a confirmed mechanism.

Sex and Motivation Patterns

Contrary to assumptions sometimes embedded in Western-derived TID models, female and male athletes in this sample did not differ significantly on any physical, technical, or psychological indicator, including competitive anxiety (SAS-2 total; d = –0.06). The only sex-linked pattern was categorical: female athletes were proportionally over-represented in the Advanced tier. Given that broader socio-cultural scholarship on women's sport participation in Bangladesh describes substantial structural and normative barriers to girls' entry into competitive sport — including restrictive mobility norms, family and community disapproval, and limited access to female-specific coaching and facilities (Huggins & Randell, 2007; Tanni & Khan, 2024) — one plausible explanation is a survivorship pattern: girls who overcome these entry barriers and remain in structured competitive tennis long

OPEN ACCESS © CINEFORUM

enough to be sampled at ages 12–16 may represent a more selectively persistent subgroup than their male counterparts, who face comparatively fewer participation barriers. This is an interpretive hypothesis consistent with the descriptive pattern observed, not a causal claim the present cross-sectional design can support.

Implications for Interpretable Machine Learning in TID

Although Light GBM with SHAP-based explanation (Lundberg & Lee, 2017) had originally been proposed for this project, the present analysis instead used Gradient Boosting, Random Forest, and permutation-based feature importance (Fisher et al., 2019)— model-agnostic, peer-reviewed alternatives well suited to a sample of this size. The resulting interpretability findings, visualized in Figure 3 (leg length, Grit-S total, and Maturity Offset as leading contributors, with aerobic capacity and anxiety-subscale indicators secondary), are broadly consistent with the descriptive and inferential results above, lending internal coherence to the analysis despite the modest overall discrimination. This reinforces a broader methodological point echoed in the interpretable-ML literature: explainability tools clarify what a model is using to make weak-to-moderate predictions, but cannot manufacture predictive signal that does not exist in the underlying indicators (Lundberg & Lee, 2017).

LIMITATIONS

Several limitations temper these findings and should guide future work. First, the design is cross-sectional; no six-month or longitudinal follow-up data were available in the dataset examined, so no developmental-trajectory or retention claims are made. Second, the domain-weighting scheme applied here reflects a documented five-member expert-panel AHP procedure (Consistency Ratio = 0.06%; Table 1) rather than a fully hybrid AHP-entropy calibration against sample data variance, and no independently validated "contextual score" adjustment mechanism was implemented; both should be treated as directions for methodological extension rather than completed components. Third, Grit-S, SAS-2, and motivation-type data were available only as domain totals, precluding item-level reliability estimation (e.g., Cronbach's α) for this sample. Fourth, with N = 200 and a 30-case held-out test set, confidence intervals around all classification metrics are necessarily wide, and the reported AUCs should be interpreted as indicative rather than definitive. Fifth, statistical significance was interpreted alongside effect sizes throughout, consistent with contemporary guidance on avoiding dichotomous p-value interpretation (Wasserstein & Lazar, 2016); several nominally non-significant comparisons (e.g., Forehand Rating, $p = .081$) may still warrant attention in adequately powered future samples.

Finally, and importantly, the outcome variable itself — coach-rated playing-level tier — is a subjective classification, not an independently verified performance criterion (see Playing-Level Classification, above). Coach ratings of this kind are potentially influenced by an athlete's physical size and maturation status (more physically mature athletes may appear more competitively advanced independent of underlying skill), by recent form or a single standout result rather than sustained ability, and by each coach's individual rating tendencies, since no inter-rater reliability check across coaches was performed. Consequently, the maturity-offset and Grit-S associations reported here should be understood as associations with coaches' competitive judgments, which may partly reflect these rating biases rather than, or in addition to, athletes' underlying tennis ability; the ML classification results are similarly best understood as evidence of how well the indicator battery reproduces coach judgment, not as evidence of predictive validity for objective competitive outcomes such as tournament ranking.

CONCLUSION

In this cross-sectional sample of 200 adolescent Bangladeshi tennis players, only biological maturation timing and grit differentiated coach-rated playing-level tiers, sex differences were absent across all physical, technical, and psychological indicators, and a 27-feature multidomain ensemble model achieved modest, above-chance classification under cross-validation (AUC ≈ .57 at best) that did not consistently generalize to a small held-out test set. Rather than supporting a high-accuracy predictive TID system, these findings empirically reinforce a growing consensus in the international TID literature (Bergkamp et al., 2018; Johnston et al., 2018) that single-timepoint, multidimensional indicator batteries carry limited standalone predictive power — underscoring the need for longitudinal, developmentally sensitive assessment frameworks, particularly in previously understudied LMIC sporting contexts such as Bangladesh.

Authorship Contributions

Md. Zillur Rahman: Conceptualization, Methodology, Investigation, Writing – Original Draft; Wu Wenqiang: Methodology, Investigation, Writing – Review & Editing; Md. Zafiroul Islam: Conceptualization, Supervision, Validation, Writing – Review & Editing; Rezwan Hossain: Conceptualization, Methodology, Data Curation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Correspondence; Md. Eman Ali: Investigation, Data Collection, Writing – Review & Editing; Md. Shafayet Hossain: Investigation, Data Collection, Writing – Review & Editing; Md Abu Sayed: Investigation, Data Collection, Writing – Review & Editing; Md. Abdullah-Hel Kafi: Investigation, Data Curation, Visualization, Writing – Review & Editing;

OPEN ACCESS © CINEFORUM

Md.Raihan Hossain: Data Collection, Formal Analysis, Writing – Review & Editing; Jannat Bibi: Methodology, Data Analysis, Validation, Writing – Review & Editing; Zhou Yaohu: Validation, Writing – Review & Editing; Hasib Al Hasan: Data Collection, Writing – Review & Editing; All authors approved the final manuscript. Md. Parvaz Zahan: Data Collection, Data Curation.

Acknowledgement
The authors acknowledge gratitude to all participants, coaches, and staff who took part in this research.
Conflicts of Interest
Authors declare no conflicts of interest.

Funding
This research received no external funding.

Data Availability
The datasets analyzed during the current study are available from the corresponding author on reasonable request.
Declaration of generative AI and AI-assisted technologies in the manuscript preparation process
During the preparation of this work the author(s) used AI editing services in order to improve language structure and text conciseness. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

REFERENCES

1. Al-Shabeeb, A. R., Al-Adamat, R., & Mashagbah, A. (2016). AHP with GIS for a preliminary site selection of wind turbines in the North West of Jordan. International Journal of Geosciences, 7(10), 1208–1221.
2. Bergkamp, T. L. G., Niessen, A. S. M., den Hartigh, R. J. R., Frencken, W. G. P., & Meijer, R. R. (2018). Comment on: "Talent Identification in Sport: A Systematic Review." In Sports Medicine (Vol. 48, Number 6). https://doi.org/10.1007/s40279-018-0868-6
3. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. European Urology, 67(6). https://doi.org/10.1016/j.eururo.2014.11.025
4. Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4). https://doi.org/10.1207/S15327965PLI1104_01
5. Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short Grit Scale (Grit-S). Journal of Personality Assessment, 91(2). https://doi.org/10.1080/00223890802634290
6. Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2). https://doi.org/10.3758/BF03193146
7. Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. Journal of Machine Learning Research, 20.
8. Harrel Jr., F. E. (2015). Regression Modeling Strategies - With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. R Software, 70(April).
9. Henriksen, K., Stambulova, N., & Roessler, K. K. (2010). Holistic approach to athletic talent development environments: A successful sailing milieu. Psychology of Sport and Exercise, 11(3). https://doi.org/10.1016/j.psychsport.2009.10.005
10. Huggins, A., & Randell, S. (2007). The contribution of sports to gender equality and women's empowerment. Paper presented at the International Conference on Gender Equity on Sports for Social Change, Kigali, Rwanda.
11. Johnston, K., Wattie, N., Schorer, J., & Baker, J. (2018). Talent Identification in Sport: A Systematic Review. In Sports Medicine (Vol. 48, Number 1). https://doi.org/10.1007/s40279-017-0803-2
12. Lloyd, R. S., Oliver, J. L., Faigenbaum, A. D., Howard, R., De Ste Croix, M. B. A., Williams, C. A., Best, T. M., Alvar, B. A., Micheli, L. J., Thomas, D. P., Hatfield, D. L., Cronin, J. B., & Myer, G. D. (2015). Long-Term Athletic Development- Part 1. Journal of Strength and Conditioning Research, 29(5). https://doi.org/10.1519/jsc.0000000000000756
13. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 2017-December.
14. Mirwald, R. L., Baxter-Jones, A. D. G., Bailey, D. A., & Beunen, G. P. (2002). An assessment of maturity from anthropometric measurements. Medicine & Science in Sports & Exercise, 34(4). https://doi.org/10.1097/00005768-200204000-00020

OPEN ACCESS © CINEFORUM

15. Mountjoy, M., Ackerman, K. E., Bailey, D. M., Burke, L. M., Constantini, N., Hackney, A. C., Heikura, I. A., Melin, A., Pensgaard, A. M., Stellingwerff, T., Sundgot-Borgen, J. K., Torstveit, M. K., Jacobsen, A. U., Verhagen, E., Budgett, R., Engebretsen, L., & Erdener, U. U. (2023). 2023 International Olympic Committee's (IOC) consensus statement on Relative Energy Deficiency in Sport (REDs). British Journal of Sports Medicine, 57(17). https://doi.org/10.1136/bjsports-2023-106994
16. Murphy, A. P., Duffield, R., Kellett, A., & Reid, M. (2014). Comparison of athlete-coach perceptions of internal and external load markers for elite junior tennis training. International Journal of Sports Physiology and Performance, 9(5). https://doi.org/10.1123/IJSPP.2013-0364
17. Smith, R. E., Smoll, F. L., Cumming, S. P., & Grossbard, J. R. (2006). Measurement of multidimensional sport performance anxiety in children and adults: The sport anxiety scale-2. Journal of Sport and Exercise Psychology, 28(4). https://doi.org/10.1123/jsep.28.4.479
18. Tabachnick, B., & Fidell, L. (2019). Using Multivariate Statistics Title: Using multivariate statistics. In Pearson Education (Vol. 5, Number 7th).
19. Tanni, A. A., & Khan, Md. M. I. (2024). Exploring the Gender Disparity in Sports Participation: A Qualitative Analysis of Women's Limited Engagement in Sports in Bangladesh. Innovation Journal of Social Sciences and Economic Review, 6(1). https://doi.org/10.36923/ijsser.v6i1.249
20. von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2014). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. International Journal of Surgery, 12(12).
21. https://doi.org/10.1016/j.ijsu.2014.07.013
22. Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on p-Values: Context, Process, and Purpose. In American Statistician (Vol. 70, Number 2). https://doi.org/10.1080/00031305.2016.1154108

OPEN ACCESS © CINEFORUM