



Observations of Player (de)Selection Within a Professional UK Soccer Academy

Rich J. Kite^{1,2} · Mark R. Noon^{2,3} · Rhys Morris³ · Peter Mundy^{2,3} · Neil D. Clarke^{2,3}

Received: 14 April 2022 / Accepted: 7 February 2023
© Beijing Sport University 2023

Abstract

The present study engaged in an ethnographical observation of the processes used to determine player (de)selections within a professional academy. English category-2 youth academy players ($n=96$) from U10–U16 age groups undertook anthropometric profiling (height, mass and somatic maturation) and fitness assessments (10 m, 20 m & 30 m linear sprints, 505-agility test, countermovement and squat jumps). Each player's lead coach ($n=4$) subjectively graded players utilising a red, amber and green (RAG) rating system on a weekly (current performance) and quarterly (perceived potential) basis, across 25 weeks. A MANCOVA, controlling for maturation, was applied to determine differences in (de)selection by physical performance. Mann Whitney-*U* tests were used to distinguish difference in (de)selection by subjective grading (weekly and quarterly). The key finding was that quarterly subjective gradings established a higher cumulative score of green ratings in selected players and a low cumulative score of red ratings, and vice versa for deselected players ($P \leq 0.001$ to 0.03). However, whilst these findings suggest that quarterly subjective grades of potential were able to provide the best predictors for player (de)selection, the findings should be viewed with caution due to high potential for confirmatory bias.

Keywords Coach intuition · Physiological · Performance grading · Maturation · Subjective assessment

Introduction

In an attempt to develop home-grown professional soccer players, the English Premier League implemented the 'Elite Player Performance Plan' (EPPP) within all English professional soccer academies [35]. The EPPP was created to provide a long-term model of development, encompassing the holistic development of players (technical/tactical, physical, psychological and social) [35]. Within the EPPP, national benchmarking of sports science and medical assessments are a mandatory criterion of data collection [35], ensuring that each player's physical profiles are monitored and measured throughout their academy journey. Additional variables include anthropometric measures in order to determine

somatic maturation [21, 26, 27]. Physical assessments typically measure components of fitness; speed, power, stamina, to list a few [13, 41, 45].

Previous research [10, 13, 14, 23, 36] in academy soccer investigating the discriminative effects of fitness (speed, power, endurance) and anthropometrics on player status (elite to non-elite, academy to non-academy, and selected to deselected players) has provided inconsistent and conflicting findings. An example of such conflicting findings has been observed with reports from le Gall et al. [23] acknowledging anthropometric differences in playing status yet no differences in speed performances. In contrast, Deprez et al. [10] established speed as a defining factor for playing status, yet anthropometry to be an insignificant variable. Therefore, the use of physical attributes to define (de)selection status remains somewhat inconclusive.

Within the EPPP, coaches are required to provide subjective feedback and reflections on player performance, providing a timeline of evidence for player development [35]. The collation of subjective feedback can later be used to inform coaches of a player's performance developments, especially when considering player (de)selection. This is particularly important when considering a player's 'potential', which

✉ Rich J. Kite
kiter@uni.coventry.ac.uk

¹ Academy Sports Science and Medical Department, Coventry City Football Club, Coventry, UK

² Centre for Sport, Exercise and Life Sciences, Coventry University, Coventry, UK

³ Faculty of Health and Life Sciences, Coventry University, Coventry, UK

may be predictive of their future abilities. Typically, subjective feedback is attained through coach intuition, defined as a coach's subjective beliefs based on experience and acquired knowledge, perceived contextual performance and self-belief in a coach's capacity to develop the individual [38]. Whilst research has demonstrated confidence in using coach intuition [22, 37, 42], concerns are raised in a coaches ability to distinguish performance within homogenous groups. Dugdale et al. [13] established that coaches could not determine performance differences within the top or bottom-performing groups of players. Likewise, Jokuschies et al. [19] reported that coaches vary in perceived areas of importance, reducing the objectivity within coach intuition. Consequently, these findings highlight limitations in the talent development process, based solely on coach intuition. However, when coach intuition is supplemented with objective assessment data, greater accuracy is attained within the player selection processes [12, 13, 42]. Therefore, in line with previous reports [13, 42], holistic assessments employing both subjective (i.e., coach intuition) and objective outputs (i.e., components of fitness, maturation, etc.) are likely to provide greater precision in player development and (de) selection outcomes.

Maturation has been established to be a confounding variable within player selection processes. In the absence of informed data, coaches have reportedly demonstrated an unconscious bias in perceived levels of potential, with late-maturing players deemed as having low-potential to succeed and a bias towards early-maturing players [7, 28, 30]. Maturation variations will likely influence selection outcomes, particularly around the years of the peak height velocity (PHV) [4, 11]. Given that players of the same chronological age may vary in maturation by two years or potentially greater [29], physical performance variations are likely to be present. Moreover, several reports have demonstrated that players exhibiting advanced maturational timing are more likely to possess superior performances when testing components of fitness [8, 11, 24, 47]. However, beyond PHV, when all players have transitioned into early adulthood, physical advantage from early maturational timing is often attenuated and/or reversed [8, 9]. Moreover, the 'Underdog Hypothesis' has provided evidence that late-maturing players can possess a greater long-term potential in performance, due to higher reported levels of self-regulation evolving from extensive periods of overstimulation [9]. Therefore, informing coaches about each players' maturational status may reduce coach (sub)conscious selection bias, enhancing development opportunities and realising player potential. Likewise, caution should be taken when (de)selecting players based on physical markers alone, particularly within adolescent age bands [8, 14, 24].

The present study serves as a pragmatic exploration to observe the current processes utilised to measure and

identify player standards and the protocols used for player (de)selection within a professional soccer academy. Moreover, an ethnographic research approach is employed to understand the suitability of the current processes applied within professional practice. Therefore, the data analysed emerged from the current methods and instruments used within a singular academy. Whilst this provides feedback unique to one academy, it is of the firm belief that similar processes are currently employed within other professional academies and clubs. Such assumptions are derived from previous research, whereby similar player grading instruments have been utilised [7, 12]. Therefore, upon the completion of data collection, data interrogation will look at the difference in (de)selection and performance grading, with the hypothesis that i) subjective grading will align with (de) selection outcomes and ii) objective measures will highlight attributes aligned with (de)selection outcomes (whilst controlling for maturation). In answering these questions, further clarity is provided on academy decisions and processes applied, whilst further offering the research findings and practical applications for other academies' interpretations.

Methods

Ninety-six male academy players of under-10 to under-16 age groups (U10–U16) (age: 13.3 ± 2.0 years) and 4 coaches (age: 32.3 ± 2.3 years) were recruited from a category-2 English professional soccer academy during the 2020–2021 season. The distribution of participation across each age was; U10 ($n = 10$), U11 ($n = 11$), U12 ($n = 10$), U13 ($n = 14$), U14 ($n = 15$), U15 ($n = 18$) and U16 ($n = 16$). The four coaches consisted of lead phase coaches covering the U9–U16 age groups with each coach covering two consecutive age groups (e.g., U16 & U15, U14 & U13, U12 & U11, and U10), all holding UEFA A-Licence qualifications and had been coaching within their respective age groups for a minimum of one season (Mean = 1.6 ± 0.5 years). Institutional ethical approval was obtained prior to data collection.

Objective Assessments

Objective assessments consisted of anthropometry and components of fitness tests and were measured using a singular time point in November (with additional planned assessment dates cancelled due to COVID-19 protocols). Given the ethnographical approach to this research, the assessments were already applied in practice. Two age groups undertook all of the assessments per day, with all age groups completing testing within one-week. Schedules were planned to ensure that each age group was provided a minimum of 48 h rest from previous training or games. Of the two age groups tested per day, a rotation was performed whereby one group would

perform speed and agility tests first, whilst the other underwent anthropometric and jumping assessments. The players were familiar with the testing protocols and had undertaken the assessments previously (as part of the academy quarterly testing battery). A standardised warm-up preceded testing, consisting of a pulse raiser and muscle activation and mobility, to ensure players were suitably prepared.

Physical Profile

Anthropometric measures were taken using a stadiometer and scales (Seca, UK) with the removal of footwear. All measures were taken abiding by the guidelines provided by the International Society for the Advancement of Kinanthropometry (ISAK), taken by the same practitioner throughout. Somatic maturation was determined by calculating the percentage of adult height [21, 26], which required a player's decimal age, current height and mass, and mid-parent heights. Parental heights were attained before the start of the season, and where self-reports were used, necessary adjustments were applied to handle typical over-estimation [15].

Components of Fitness

Linear sprints and 505 change of direction (COD) tests were complete using single-beam light gates (Smartspeed, USA) on an indoor 3G pitch. Tests were initiated with a falling start, whereby the player starts 0.5 m before the first gate, with the feet in line. In both linear and COD assessments, players were informed to run beyond the final gate to prevent early deceleration. In the COD, light gates were placed at 10 m, with test markings produced at 15 m for players to change direction on. The assessed foot was required to be placed beyond the 15 m marking before returning through the light gate. Three trials were used for the linear sprints, and two trials per turning leg in the 505, using the best trials for further analysis..

Jumping tasks were performed using two force plates (Pasco, USA) set to 1000 Hz, and a compatible analysis software package (Capstone, USA). For both the countermovement- and squat jumps (CMJ and SQJ), players placed one foot on each force plate and were asked to remain stationary to capture bodyweight. In both jumps, players used a self-prescribed jump depth with the arms on the hips until the completion of each jump and were asked to jump 'as high as you can'. During the SQJ, players were asked to hold the dipped position for at least 2 s before initiating the jump. Players were required to land back on the plates, whilst absorbing the landing forces. Three trials of each jump were collected, with the best trial being used for further analysis.

Whilst this research undertakes an ethnographic approach, previous research has reported good test retest reliability (10 m, $ICC=0.91$, $CV=2.3\%$; 20 m, $ICC=0.91$,

$CV=2.9\%$; 30 m, $ICC=0.99$, $CV=0.9\%$) in linear sprints [40], likewise good test–retest reliability ($ICC>0.899$) [3] and validity ($ICC=0.77$, $CV=2.80\%$) [43] in the 505 agility test. Furthermore, to determine the reliability of the data, the coefficient of variation (CV) was calculated for each objective variable. Participants who demonstrated $CV>10\%$ were deemed poor for reliability and, therefore, excluded from further analysis within each variable [6, 18].

Subjective Assessments

Coach subjective assessments were used to measure the technical and tactical abilities and overall potential of players, based on coach beliefs and perceptions. Subjective measures were taken on a weekly basis, identifying current performance, and a quarterly basis identifying a player's perceived future potential. The academy employed a Red, Amber and Green (RAG) rating system for all subjective gradings as standard practice. The RAG rating system is commonly used within academy infrastructures, whereby common definitions state that red is 'performing below the expected standard', amber is 'performing at the expected standard' and green is 'performing above the expected standard' (or similar). This system is further integrated into the EPPP online audit system, the PMA, arguably explaining the original use of this method. Tallies of both weekly subjective performance grading and the quarterly subjective potential grading were used for further statistical analysis. Previous studies have looked at the test–retest reliability of coach subjective grading of players [19] establishing partly acceptable and partly unacceptable ($-0.57 \leq r \leq -0.81$) reliability outcomes.

Weekly Subjective Grading

Weekly subjective player grading of technical and tactical abilities were determined by lead coaches of their respective age groups. Coaches would provide a weekly score (red, amber or green) per player. For the entirety of the 2020–2021 season, performance RAG ratings were collected and tallied into a total quantity of red, amber and green scores. However, due to COVID-19 restrictions during the season, only 25-weeks (of the traditional ~36-weeks) across all ages (excluding U16s) were recorded for assessment. Due to the U16 (de)selection process being earlier than other ages, due to scholarship transitions, only 16-weeks (of the traditional ~26-weeks) of coach gradings were collected before selection decisions.

Quarterly Subjective Grading

Further assessments included perceived measures of potential. Coaches underwent quarterly (Q1 = September,

Q2 = December, Q3 = April) subjective assessments of player future potential, further utilising the RAG rating. Coaches would assign each player either a red, amber or green score during each quarter of the season. Therefore, following the end of the season, each player would have three RAG scores for potential. The quarterly subjective potential grading is determined across a player's holistic ability (psychology, technical, tactical, social, physical), using available objective outcomes (provided by the sports science and medical department), and the coach's belief in the continual rate of progression in development and performance.

Player Selection

The process of player selection was undertaken as per normal academy procedures. Coaches were provided with all objective data prior to meetings between the lead coach, academy manager and head of coaching to discuss player selections. The U16 age group was the only age group where more staff was present within the selection process, including sports science and medical staff, U18 coaching staff, head of recruitment and head of education. Consequently, the outcome of this process results in 29 deselected players and 67 selected players (Table 1).

Statistical Analysis

Data distribution was repeatedly assessed by age group, using the Shapiro–Wilk test of normality. The U11 and U12 age groups combined (21 players) featured only one deselected player, and therefore was ineligible for further analysis and removed from the dataset. A multivariate analysis of covariance (MANCOVA) was applied for the objective assessments, utilising maturation as a covariant and linear speed (10, 20 m and 30 m sprints), COD (505 left and right), jumping tasks (CMJ and SQJ) and anthropometry (height and mass) as dependent variables, to investigate differences between age groups and selection status (independent variables). Maturation was controlled for given its potentially confounding influence within the analysis of physical performance [25, 28, 33]. An alpha level of <0.05 was applied and

follow-up univariate analysis (with Bonferroni adjustments) [1] were used where appropriate.

Where violations of normal distributions were observed, such as within the tally of RAG ratings, non-parametric tests were used. A Mann Whitney *U*-test was applied to identify the difference in player (de)selection and RAG tallies. Tallies were determined using cumulative RAG scores awarded by coaches across a season, per player. Therefore, a player can only be awarded either a red, amber or green per week or quarter. Tallies of each red, amber and green were then compared for differences by (de)selection outcomes. Due to the multiple comparisons of data, a Bonferroni correction was applied to reduce type 1 error [1]. A Bonferroni correction level was determined as the quantity of independent variables multiplied by the quantity of dependant variables, with the outcome providing the division of alpha set at <0.05. Therefore, a new alpha was set at <0.017. For the Mann Whitney *U* tests, *r* values were determined from Z-scores [16, 17], with outcomes ≥ 0.1 –0.29 = small, 0.3–0.49 = medium and ≥ 0.5 = large effect size. Subsequently, Eta-squared was calculated from *r* value outcomes. Eta-squared effect sizes were interpreted as >0.01 = small effect, >0.06 = medium effect and >0.14 = large effect. Outcomes from the Mann Whitney *U* test were reported as medians and interquartile range (IQR).

Given the high subjectivity of the selection process, further statistics were applied to explore the potential for subjective bias. Recent player performance may influence decisions on selection outcome (over the utility of the full season report of performance). Time course influence was investigated by comparing differences in group performances by selection status between weeks 1–20 and weeks 21–25, using a MANOVA with observations of status by weeks interaction. Additionally, a Cramer's *V* was used to interrogate the associations of within quarterly subjective potential gradings, and between quarterly subjective potential gradings and selection outcomes. Effect size for Cramer's *V* were interpreted based upon degrees of freedom (Table 2), as outlined by Cohen [5]. All data were analysed using SPSS Statistics for Windows, Version 26.0 (Armonk, NY: IBM Corp.).

Table 1 Distribution of selection outcomes by academy age group

Age group	Selected	Deselected
U10	<i>n</i> = 8	<i>n</i> = 4
U11	<i>n</i> = 10	<i>n</i> = 1
U12	<i>n</i> = 10	<i>n</i> = 0
U13	<i>n</i> = 10	<i>n</i> = 4
U14	<i>n</i> = 9	<i>n</i> = 6
U15	<i>n</i> = 12	<i>n</i> = 6
U16	<i>n</i> = 8	<i>n</i> = 8

Table 2 Effect size for Cramer's *V* based on degrees of freedom, adapted from Cohen [5]

Degrees of freedom	Small effect	Medium effect	Large effect
1	0.1 to <0.3	0.3 to <0.5	≥ 0.5
2	0.07 to <0.21	0.21 to <0.35	≥ 0.35
3	0.06 to <0.17	0.17 to <0.29	≥ 0.29

Results

The findings of the MANCOVA determined significant variations in performance within only the age group [Wilks Lambda = 0.11, $F(54, 254) = 2.61, P = 0.00$] and the covariant of maturation [Wilks Lambda = 0.55, $F(9, 49) = 4.39, P = 0.00$]. No significant variances were observed within player selection status [Wilks Lambda = 0.79, $F(9, 49) = 1.49, P = 0.18$] or player selection status \times age interaction [Wilks Lambda = 0.50, $F(36, 185) = 1.05, P = 0.16$].

The outcomes of the weekly subjective grade tallies identified significantly (with applied Bonferroni

correction) higher tallies of green ratings in selected players ($U = 0.0, P < 0.001, r = -0.83, \eta^2 = 0.68$) and red ratings in deselected players ($U = 4.5, P = 0.01, r = -0.82, \eta^2 = 0.67$), within the U14 age group only (Table 3). Within the quarterly grading tallies, a consistent finding was observed with the low quantity of red ratings for selected players, and high quantity for deselected players in the U13, U14, U15 and U16 age groups (Table 3). Only in the U15 age group were significant differences identified in the tally of green ratings ($U = 2.5, P < 0.001, r = -0.74, \eta^2 = 0.55$) between the selected (median = 2, IQR = 2–3) and deselected (median = 0, IQR = 0–0) players.

In considerations of subjective bias at different time courses, the results from the MANOVA found no significant

Table 3 Results of the Mann Whitney U test reporting differences and median (and interquartile range) performances between select/deselect groups in weekly and quarterly subjective gradings

Age group	RAG	<i>U</i>	Sig	Selected median (IQR)	Deselected median (IQR)	γ	η^2
Weekly							
U10	Red	7.0	0.15	4 (2–4)	6 (4–7)	-0.45	0.20
	Amber	14.5	0.81	13 (12–14)	13 (11–14)	-0.07	0.01
	Green	3.0	0.03	6 (4–6)	1 (0–2)	-0.65	0.43
U13	Red	9.0	0.14	0 (0–1)	1 (0–2)	-0.17	0.03
	Amber	7.5	0.08	9 (8–12)	12 (11–12)	-0.36	0.13
	Green	4.5	0.02	2 (2–3)	2 (1–2)	-0.49	0.24
U14	Red	4.5	0.01*	0 (0–0)	1 (1–2)	-0.82	0.67
	Amber	13.0	0.11	10 (10–11)	14 (13–15)	-0.43	0.19
	Green	0.0	<0.001*	7 (7–9)	2 (1–3)	-0.83	0.68
U15	Red	30.0	0.62	0 (0–0)	0 (0–0)	-0.12	0.01
	Amber	12.0	0.02	7 (3–9)	10 (9–12)	-0.53	0.28
	Green	11.0	0.02	16 (13–18)	10 (4–13)	-0.39	0.15
U16	Red	31.5	0.96	0 (0–0)	0 (0–0)	-0.02	0.00
	Amber	18.0	0.16	2 (1–3)	4 (3–6)	-0.37	0.14
	Green	23.0	0.38	8 (8–9)	8 (5–9)	-0.24	0.06
Quarterly							
U10	Red	4.5	0.18	0 (0–0)	1 (1–2)	-0.53	0.29
	Amber	8.0	0.67	1 (1–2)	1 (1–1)	-0.20	0.04
	Green	7.0	0.52	1 (0–2)	0 (0–1)	-0.27	0.07
U13	Red	1.0	<0.001*	0 (0–0)	2 (2–2)	-0.80	0.65
	Amber	19.0	0.95	2 (0–2)	1 (1–1)	-0.04	0.00
	Green	6.0	0.05	1 (0–3)	0 (0–0)	-0.57	0.33
U14	Red	0.0	<0.001*	0 (0–0)	2 (1–2)	-0.92	0.85
	Amber	18.0	0.49	0 (0–2)	1 (0–2)	-0.23	0.05
	Green	6.0	0.02	3 (1–3)	0 (0–1)	-0.66	0.43
U15	Red	0.0	<0.001*	0 (0–0)	1 (1–2)	-0.95	0.91
	Amber	12.5	0.09	1 (0–1)	2 (1–2)	-0.45	0.20
	Green	2.5	<0.001*	2 (2–3)	0 (0–0)	-0.74	0.55
U16	Red	2.5	<0.001*	0 (0–0)	1 (1–1)	-0.83	0.69
	Amber	17.5	0.41	1 (0–1)	1 (0–1)	-0.24	0.06
	Green	7.0	0.03	2 (2–3)	1 (0–1)	-0.61	0.37

IQR interquartile range (25%–75%)

* = significant outcome < 0.017 (with Bonferroni correction)

($P > 0.05$) differences in all age groups when considering interactions between selection status (selected vs. deselected) and weeks (weeks 1–20 vs. weeks 21–25). When investigating the associations of quarterly potential gradings, Cramer's V observed various associations between different quarters and selection outcome (Table 4). In comparison of gradings across quarters, whilst all age groups across each quarter demonstrated large effect sizes, only few significant

associations were reported within Q1 and Q2, and Q1 and Q3. Likewise, when looking at associations with quarterly grading and selection outcome, only a few large significant findings were established in Q1 and Q2, whereas Q3 demonstrate large to perfect significant associations across all age groups.

Discussion

This study looked to report on the current processes employed within a professional soccer academy undertaking player (de)selection, via an ethnographical approach, with a focus on identifying the difference in performance between (de)selection status. The key findings of this study were that coach subjective measures of player abilities were inconclusive in determining selection status. Whilst coach perceptions of player potential were capable of determining selection status in the final quarter only, this was potentially indicative of confirmative bias given the close proximity to when selection decisions are finalised. Additionally, weekly subjective gradings (current performance) were unable to consistently distinguish selection status.

The outcomes of quarterly RAG gradings reported a greater number of red grades associated with deselection, with observations of higher green frequencies associated to selection. This implies that coaches are capable in distinguishing players at either extremity of performance (i.e., top- or bottom-performing players). Such findings align with previous reports [13, 42] that identify coach intuition as a capable tool in determining (de)selected players. However, these findings also demonstrated that coaches may display indecision concerning the players of a moderate standard (amber), similar to previous research [13]. Whilst this middle ground is expected to cater to players with greater uncertainty towards their future, the range of abilities is vast and features players that were opted for both selection and deselection. Therefore, to reduce this range of abilities, it may be of greater benefit to further sub-divide grades to offer greater clarity in perceptions of player abilities and potential, similar to the nine by nine grid of performance and potential proposed by Baker et al. [2].

Conversely, the quarterly subjective grade results fail to control for confirmation bias, given that the same coaches who graded the players were also the coaches who (de)selected them. Considering the process of deselection observed evolved from weekly performance feedback with managerial staff, it is fair to elude that the lead phase coaches may (sub)consciously inform and influence their viewpoints on player performances. Therefore, it is impossible to dismiss the potential for confirmation bias. Additionally, this study established large-to-perfect associations across all quarters of potential grading and selection outcomes, with

Table 4 The associations between quarters of subjective gradings of potential, and the associations of quarterly gradings and selection outcomes

Age	DF	Cramer's V	Sig	Effect size
Q1 and Q2				
U10	–	–	–	–
U13	2	0.89	0.03	Large
U14	2	0.98	0.01*	Large
U15	2	0.66	0.13	Large
U16	2	0.48	0.19	Large
Q1 and Q3				
U10	2	0.38	0.57	Large
U13	2	0.51	0.15	Large
U14	2	0.57	0.12	Large
U15	2	0.73	0.07	Large
U16	2	0.73	0.02	Large
Q2 and Q3				
U10	–	–	–	–
U13	2	0.45	0.27	Large
U14	2	0.62	0.08	Large
U15	2	0.66	0.14	Large
U16	2	0.31	0.86	Large
Q1 and selection status				
U10	1	0.22	0.79	Small
U13	1	0.57	0.12	Large
U14	1	0.57	0.12	Large
U15	1	0.71	0.02	Large
U16	1	0.53	<0.001*	Large
Q2 and selection status				
U10	–	–	–	–
U13	1	0.57	0.12	Large
U14	1	0.62	0.08	Large
U15	1	0.65	0.04	Large
U16	1	0.13	0.90	Small
Q3 and selection status				
U10	1	0.79	0.05	Large
U13	1	0.84	0.01*	Large
U14	1	1.00	<0.001*	Perfect
U15	1	1.00	<0.001*	Perfect
U16	1	1.00	<0.001*	Perfect

Q1 = quarter 1, Q2 = quarter 2, Q3 = quarter 3

* = significant outcome < 0.017 (with Bonferroni correction)

the greatest associations observed within the third quarter. Given the proximity to selection decisions, it is probable that the Q3 potential scores are more representative of selection decisions. Furthermore, the gradings are gathered by the lead coach alone. Therefore, to both mitigate selection bias and improve clarity for player developments and selection, further objectivity towards measures of potential should be explored, potentially incorporating the use of multiple inputs from additional coaching staff.

Weekly subjective grades demonstrated a low utility in identifying selected from deselected players, with only the U14 age group reporting significant findings of red and green grades. Whilst the weekly subjective grades found low ability in identifying (de)selection outcomes, one concern of this instrument surrounded the bias of more recent performance influencing selection decisions, over the full season collection of grades. However, when assessing associations between the final five weeks of performance to the complete season, fair to very strong associations were observed. Therefore, providing greater reassurance that the weekly subjective grades provided a fair reflection of performance across the season. Furthermore, associations between weekly and quarterly subjective grades established a consistent correlation within green scores primarily. This suggests that top performing players maintain a more consistent playing performance than mid to lower performing players.

One explanation as to why weekly subjective grades failed to distinguish player selection outcomes is due to coaches providing grades to their individual age groups only, with no additional inputs. Each coach will maintain an expected performance standard based upon experience, knowledge, and beliefs, as outlined within the definition of coach intuition [22, 37]. This will undoubtedly vary between coaches. Therefore, whilst one coach may perceive a player to be a 'amber' performance grade, another may perceive them to be a 'green'. This lack of uniformity in player grading highlights the need for academies to provide comprehensive anchor points within the grading tool, so to enhance clarity in player grading, like in the earlier proposed instrument by Baker et al. [2]. Likewise, academies may benefit from the adoption of a validated and reliable instrument for the subjective measure of player performance.

In consideration of the differing outcomes between weekly and quarterly subjective gradings, it may be postulated that the focus of each instrument (weekly vs. quarterly) provides the difference in selection outcome. Whilst weekly subjective grading focuses on current performance, quarterly subjective grading emphasizes perceived future potential. It is therefore plausible that coaches may measure current performance against a player's perceived future potential, i.e., a player who is believed to have high potential may be graded as currently underperforming, based on the coach's perceptions of the players' future abilities yet to be realised.

Likewise, a player with low potential may be scored higher in their current performance, given that the coach perceives them as playing to the highest standard expected of them to achieve (or in some cases, beyond this). Therefore, current performance ratings may instead provide some context in the development and attainment of (perceived) potential and may be best employed by utilising frameworks that account for both 'current performance' vs. 'perceived potential' [2, 44].

In the present study, physical performance was not capable of distinguishing selected from deselected players. Given that the study undertook an ethnographic approach, whereby procedures tested are standard practice, and that objective assessments failed to distinguish selection outcome, potentially highlights the need for wider metrics to be applied to aid player selection. For example, a previous report has demonstrated the high interaction of acceleration within the 505 COD task design, resulting in only 31% of time spent changing direction [32]. Therefore, those with greater linear speed abilities are more likely to demonstrate greater overall COD outcomes [31, 32]. Nimphius et al. [31] suggested using a change of direction deficit (COD_{def}) calculation to enhance the measurement of turning ability, consequently mitigating linear speed bias and providing a more reflective measure of task assessment. It may be suggested that the use of more task-specific and holistic (physical, psychological, social, technical and tactical) assessments, or contrary, assessments with higher ecological validity, may better serve the player (de)selection process.

A consideration in the findings for physical performance is the control for maturational variation. The current study found maturation to be a significant factor within selection. Therefore, academies must be mindful towards the control for maturational influence within physical performance data. Previous research has highlighted the influence of maturation on physical performance [25, 28, 33], whereby players exhibiting an advanced maturation status will possess greater physical abilities to their biologically younger peers. As a consequence, players of an advanced maturation status are selected due to being perceived as beholding a higher potential for success [7]. Conversely, players of a late maturation status are released, due to a perception of low potential. However, research has highlighted that early maturation players are no more likely to achieve senior professional success to their biologically younger peers [20]. Without the affordance of time to achieve adult stature (therefore 'catching up' with their early and average maturing peers), late maturing players will not realise their full potential. Furthermore, late maturing players may only progress through academy selection processes by 'survival' [34], whereby they can tolerate over-stimulation without enduring burnout, injury or demotivation [34, 39, 46]. In summary, given the high variation of maturation identified within youths, controlling

for maturation in physical performance is essential for consideration within the player (de)selection process.

This study is not without its limitations. Whilst this was an observation of a single academy's practice, further research should include several academies to identify (de)selection trends. However, the outcomes of this study can inform academies utilising similar selection processes of the stated shortcomings and considerations for enhancement. Likewise, this study highlighted areas that are not currently accounted for during player selection process. Whilst there is an open understanding for holistic abilities, the lack of assessments for psychological abilities and social measures potentially implicates player developments and (de)selection. Further limitations of this study may consider the lack of reliability assessments undertaken through the use of the RAG rating method and coach scoring. Whilst previous studies have been undertaken that outline test–retest reliability in coach subjective scoring [19], further work should be undertaken assess the use of RAG ratings.

Additionally, uncontrollable challenges were associated with the present study; during the season of 2020–2021, the COVID-19 international pandemic brought closure to academy soccer within the UK at various time points e.g., delayed start and closures in December and January. Academy closure (therefore a cessation of training) and imposed restrictions negatively impacted training provisions and physical performance assessment collection, traditionally undergone quarterly (June, October, January and April). This resulted in a singular time point provided for the 2020–2021 season. Moreover, the outcome reflects the 2020–2021 academy season, where the best operating procedure was implemented to mitigate risks and superseded the need for further data collection.

Practical Implications

Based on the findings of this study, coach subjective perception of player future potential were inconclusive in determining player deselection. Coaches were only able to identify players perceived as high potential in the final quarter, and given the close proximity of selection decisions being made, suggests decisions are likely decided upon these outcomes. Moreover, this indicates a level of confirmative bias in the subjective process. Therefore, it may be suggested that further objectivity is required within the process for grading potential in order to enhance player development and selection processes.

When considering weekly subjective gradings, it has been postulated that performance grades may provide context and further inform the coach on the present status of the players' predicted journey. Moreover, the continual collection of weekly and quarterly subjective grading is logical, with further research looking to better understand

what further distinguishes player abilities. However, caution should be raised around the consistent pressure placed on players to perform. Given that measures of potential were unable to predict subsequent selection status, players remain under consistent pressure to perform throughout the season to retain selection status. Such issues are pertinent for players of moderate player abilities. Given it was clear that coaches were able to differentiate players at each extremity of performance, uncertainty remained around moderate ability players. Further research needs to be undertaken to investigate a coach's ability to perceive the development needs of these players, to ensure they are provided the appropriate and optimal provisions.

Given the confounding effect of maturation, bio-banding interventions may be convenient to ensure optimal developments are maintained. Banding players by biological age entails grouping players by growth status, regardless of chronological age. Such provisions will reduce any dependence on physical prowess and the higher demand for technical and tactical ability, typically exhibited by biologically advanced players. Likewise, such provisions can also be offered to biologically younger players, moving down and age group, which should afford the time to develop confidence and develop leadership skills; opportunities less likely to be presented within their own age groups.

Conclusion

The present study identified that a season of coach subjective perceptions (current performance and perceived potential) were inconclusive of determining subsequent selection status. Furthermore, the results indicate a high chance of confirmation bias associated with the current selection process. Additionally, it was apparent that whilst coaches were capable of distinguishing the extremities of player performance, coaches remain uncertain around the grading for moderate ability players. Further work is required to provide more objective assessments to mitigate the potential for bias within assessments, and enhance the accuracy in the player selection process.

Data availability The datasets generated and analysed during the current study are not publicly available due to maintaining confidentiality of the participants and the associated club, but are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors report no conflict of interest.

Consent to participate Consent to participation was obtained from all participants.

Consent for publication All authors have read and agreed to the published version of the manuscript.

References

1. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt.* 2014;34(5):502–8. <https://doi.org/10.1111/OPO.12131>.
2. Baker J, Schorer J, Wattie N. Compromising talent: issues in identifying and selecting talent in sport. *Quest.* 2018;70:48–63. <https://doi.org/10.1080/00336297.2017.1333438>.
3. Barber OR, Thomas C, Jones PA, McMahon JJ, Comfort P. Reliability of the 505 change-of-direction test in netball players. *Int J Sports Physiol Perform.* 2016;11(3):377–80. <https://doi.org/10.1123/ijsp.2015-0215>.
4. Bidaurrazaga-Letona I, Lekue JA, Amado M, Gil SM. Progression in youth soccer: selection and identification in youth soccer players aged 13–15 years. *J Strength Cond Res.* 2019;33(9):2548–58. <https://doi.org/10.1519/JSC.0000000000001924>.
5. Cohen J. *Statistical power analysis for the behavioral sciences*, second. Hillsdale: Lawrence Erlbaum Associates; 1988.
6. Cormack SJ, Newton RU, McGulgan MR, Doyle TLA. Reliability of measures obtained during single and repeated countermovement jumps. *Int J Sports Physiol Perform.* 2008;3(2):131–44. <https://doi.org/10.1123/ijsp.3.2.131>.
7. Cripps AJ, Hopper LS, Joyce C. Coaches' perceptions of long-term potential are biased by maturational variation. *Int J Sport Sci Coach.* 2016;11(4):478–81. <https://doi.org/10.1177/1747954116655054>.
8. Cumming SP, Lloyd RS, Oliver JL, Eisenmann JC, Malina RM. Bio-banding in sport: applications to competition, talent identification, and strength and conditioning of youth athletes. *Strength Cond J.* 2017;39(2):34–47. <https://doi.org/10.1519/SSC.0000000000000281>.
9. Cumming SP, Searle C, Hemsley JK, Haswell F, Edwards Hannah, Scott S, Gross A, Ryan D, LJ, White P, Cain A, Mitchell SB, Malina RM. Biological maturation, relative age and self-regulation in male professional academy soccer players: a test of the underdog hypothesis. *Psychol Sport Exerc.* 2018;39:147–53. <https://doi.org/10.1016/j.psychsport.2018.08.007>.
10. Deprez DN, Franssen J, Lenoir M, Philippaerts RM, Vaeyens R. A retrospective study on anthropometrical, physical fitness, and motor coordination characteristics that influence dropout, contract status, and first-team playing time in high-level soccer players aged eight to eighteen years. *J Strength Cond Res.* 2015;29(6):1692–704. <https://doi.org/10.1519/JSC.0000000000000806>.
11. Dodd KD, Newans TJ. Talent identification for soccer: physiological aspects. *J Sci Med Sport.* 2018;21(10):1073–8.
12. Dugdale JH, McRobert AP, Unnithan VB. Selected, deselected, and reselected: a case study analysis of attributes associated with player reselection following closure of a youth soccer academy. *Front Sport Act Living.* 2021;3:633124. <https://doi.org/10.3389/fspor.2021.633124>.
13. Dugdale JH, Sanders D, Myers T, Williams AM, Hunter AM. A case study comparison of objective and subjective evaluation methods of physical qualities in youth soccer players. *J Sports Sci.* 2020;38(11–12):1304–12. <https://doi.org/10.1080/02640414.2020.1766177>.
14. Emmonds S, Till K, Jones B, Mellis M, Pears M. Anthropometric, speed and endurance characteristics of English academy soccer players: do they influence obtaining a professional contract at 18 years of age? *Int J Sport Sci Coach.* 2016;11(2):212–8. <https://doi.org/10.1177/1747954116637154>.
15. Epstein LH, Valoski AM, Kalarchian MA, McCurley J. Do children lose and maintain weight easier than adults: a comparison of child and parent weight changes from six months to ten years. *Obes Res.* 1995;3:411–7. <https://doi.org/10.1002/J.1550-8528.1995.TB00170.X>.
16. Field A. *Effect Sizes Null Hypothesis Significance Testing (NHST)*. 2005. <http://www.discoveringstatistics.com/docs/effectsizes.pdf>. Accessed 7 July 2021.
17. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen.* 2012;141(1):2–18. <https://doi.org/10.1037/a0024338>.
18. Jennings D, Cormack S, Coutts AJ, Boyd L, Aughey RJ. The validity and reliability of GPS units for measuring distance in team sport specific running patterns. *Int J Sports Physiol Perform.* 2010;5(3):328–41. <https://doi.org/10.1123/ijsp.5.3.328>.
19. Jokuschies N, Gut V, Conzelmann A. Systematizing coaches' 'eye for talent': Player assessments based on expert coaches' subjective talent criteria in top-level youth soccer. *Int J Sport Sci Coach.* 2017;12(5):565–76. <https://doi.org/10.1177/1747954117727646>.
20. Kelly AL, Williams CA. Physical characteristics and the talent identification and development processes in male youth soccer: a narrative review. *Strength Cond J.* 2020;42(6):15–34.
21. Khamis HJ, Roche AF. Predicting adult stature without using skeletal age: the Khamis-Roche method. *Pediatrics.* 1994(4 Pt 1):94:504–7. <https://doi.org/10.1542/peds.94.4.504>.
22. Lath F, Koopmann T, Faber I, Baker J, Schorer J. Focusing on the coach's eye: towards a working model of coach decision-making in talent selection. *Psychol Sport Exerc.* 2021;56:102011. <https://doi.org/10.1016/j.psychsport.2021.102011>.
23. le Gall F, Carling C, Williams M, Reilly T. Anthropometric and fitness characteristics of international, professional and amateur male graduate soccer players from an elite youth academy. *J Sci Med Sport.* 2010;13:90–5. <https://doi.org/10.1016/j.jsams.2008.07.004>.
24. Lloyd RS, Oliver JL, Faigenbaum AD, Myer GD, De Ste Croix M. Chronological age vs. biological maturation. *J Strength Cond Res.* 2014;28(5):1454–64. <https://doi.org/10.1519/JSC.0000000000000391>.
25. Lloyd RS, Oliver JL, Radnor JM, Rhodes BC, Faigenbaum AD, Myer GD. Relationships between functional movement screen scores, maturation and physical performance in young soccer players. *J Sports Sci.* 2014. <https://doi.org/10.1080/02640414.2014.918642>.
26. Malina RM, Cumming SP, Morano PJ, Barron M, Miller SJ. Maturity status of youth football players: a noninvasive estimate. *Med Sci Sports Exerc.* 2005;37(6):1044–52. <https://doi.org/10.1249/01.mss.0000171622.45134.cb>.
27. Malina RM, Dompier TP, Powell JW, Barron MJ, Moore MT. Validation of a noninvasive maturity estimate relative to skeletal age in youth football players. *Clin J Sport Med.* 2007;17(5):362–8. <https://doi.org/10.1097/JSM.0b013e31815400f4>.
28. Meylan C, Cronin J, Oliver J, Hughes M. Talent identification in soccer: the role of maturity status on physical, physiological and technical characteristics. *Int J Sport Sci Coach.* 2010;5(Suppl-1):571–92. <https://doi.org/10.1260/1747-9541.5.4.571>.
29. Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. *Med Sci Sport Exerc.* 2002;34(4):689–94. <https://doi.org/10.1097/00005768-200204000-00020>.
30. Myburgh GK, Cumming SP, Silva CE, M, Cooke K, Malina RM. Growth and maturity status of elite British junior tennis players. *J Sports Sci.* 2016;34(20):1957–64. <https://doi.org/10.1080/02640414.2016.1149213>.

31. Nimphius S, Callaghan SJ, Bezodis NE, Lockie RG. Change of direction and agility tests: challenging our current measures of performance. *Strength Cond J*. 2018;40(1):26–38. <https://doi.org/10.1519/SSC.000000000000309>.
32. Nimphius S, Callaghan SJ, Spiteri T, Lockie RG. Change of direction deficit: a more isolated measure of change of direction performance than total 505 time. *J Strength Cond Res*. 2016;30(11):3024–32. <https://doi.org/10.1519/JSC.000000000001421>.
33. Patel R, Nevill A, Smith T, Cloak R, Wyon M. The influence of birth quartile, maturation, anthropometry and physical performances on player retention: observations from an elite football academy. *Int J Sport Sci Coach*. 2020;15(2):121–34. <https://doi.org/10.1177/1747954120906507>.
34. Pion J, Lenoir M, Vandorpe B, Segers V. Talent in female gymnastics: a survival analysis based upon performance characteristics. *Int J Sports Med*. 2015;36(11):935–40. <https://doi.org/10.1055/s-0035-1548887>.
35. Premier League. Elite Player Performance Plan. 2011. <http://www.premierleague.com/content/premierleague/en-gb/youth/elite-player-performance-plan.html>. Accessed 15 Aug 2020.
36. Reilly T, Williams AM, Nevill A, Franks A. A multidisciplinary approach to talent identification in soccer. *J Sports Sci*. 2000;18:695–702. <https://doi.org/10.1080/02640410050120078>.
37. Roberts A, Greenwood D, Stanley M, Humberstone C, Iredale F, Raynor A. Understanding the “gut instinct” of expert coaches during talent identification. *J Sports Sci*. 2021;39(4):359–67. <https://doi.org/10.1080/02640414.2020.1823083>.
38. Roberts AH, Greenwood DA, Stanley M, Hurnberstone C, Iredale F, Raynor A. Coach knowledge in talent identification: a systematic review and meta-synthesis. *J Sci Med Sport*. 2019;22(10):1163–72.
39. Rothwell M, Rumbold JL, Stone JA. Exploring British adolescent rugby league players’ experiences of professional academies and dropout. *Int J Sport Exerc Psychol*. 2020;18(4):485–501. <https://doi.org/10.1080/1612197X.2018.1549579>.
40. Shalfawi SAI, Enoksen E, Tønnessen E, Ingebrigtsen J. Assessing test-retest reliability of the portable brower speed trap II testing system. *Assess TEST-RETEST Reliab*. 2012;44:24–30.
41. Sieghartsleitner R, Zuber C, Zibung M, Charbonnet B, Conzelmann. Talent selection in youth football: Specific rather than general motor performance predicts future player status of football talents. *Curr Issues Sport Sci*. 2019;4:1–14. https://doi.org/10.15203/CISS_2019.011.
42. Sieghartsleitner R, Zuber C, Zibung M, Conzelmann A. Science or coaches’ eye? both! beneficial collaboration of multidimensional measurements and coach assessments for efficient talent selection in elite youth football. *J Sport Sci Med*. 2019;18(1):32–43.
43. Stewart PF, Turner AN, Miller SC. Reliability, factorial validity, and interrelationships of five commonly used change of direction speed tests. *Scand J Med Sci Sport*. 2014;24(3):500–6. <https://doi.org/10.1111/sms.12019>.
44. Till K, Baker J. Challenges and [possible] solutions to optimizing talent identification and development in sport. *Front Psychol*. 2020;11:1–14. <https://doi.org/10.3389/fpsyg.2020.00664>.
45. Towson C, Cope E, Perry JL, Court D, Levett N. Practitioners’ multi-disciplinary perspectives of soccer talent according to phase of development and playing position. *Int J Sport Sci Coach*. 2019;14(4):528–40. <https://doi.org/10.1177/1747954119845061>.
46. Van Der Sluis A, Elferink-Gemser MT, Brink MS, Visscher C. Importance of peak height velocity timing in terms of injuries in talented soccer players. *Int J Sports Med*. 2015;36:327–32. <https://doi.org/10.1055/s-0034-1385879>.
47. Vandendriessche JB, Vaeyens R, Vandorpe B, Lenoir M, Lefevre J. Biological maturation, morphology, fitness, and motor coordination as part of a selection strategy in the search for international youth soccer players (age 15–16 years). *J Sports Sci*. 2012;30(15):1695–703. <https://doi.org/10.1080/02640414.2011.652654>.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.