European Journal of Sport Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tejs20

Coaches' Assessment of Players Physical Performance: Subjective and Objective Measures are needed when Profiling Players

Sam McCormack, Ben Jones, Dave Elliott, Dave Rotheram & Kevin Till

To cite this article: Sam McCormack, Ben Jones, Dave Elliott, Dave Rotheram & Kevin Till (2021): Coaches' Assessment of Players Physical Performance: Subjective and Objective Measures are needed when Profiling Players, European Journal of Sport Science, DOI: 10.1080/17461391.2021.1956600

To link to this article: https://doi.org/10.1080/17461391.2021.1956600

Published online: 01 Aug 2021.

EUROPEAN JOURNAL OF SPORT SCIENCE
https://doi.org/10.1080/17461391.2021.1956600

Routledge
Taylor & Francis Group

Coaches' Assessment of Players Physical Performance: Subjective and Objective Measures are needed when Profiling Players

Sam McCormack [a,b], Ben Jones [a,b,c,d,e], Dave Elliott[b], Dave Rotheram[b] and Kevin Till [a,c]

[a]Carnegie Applied Rugby Research (CARR) Centre, Carnegie School of Sport, Leeds Beckett University, Leeds, UK; [b]England Performance Unit, Rugby Football League, Leeds, UK; [c]Leeds Rhinos Rugby League Club, Leeds, UK; [d]School of Science and Technology, University of New England, Armidale, Australia; [e]Division of Exercise Science and Sports Medicine, Department of Human Biology, Faculty of Health Sciences, the University of Cape Town and the Sports Science Institute of South Africa, Cape Town, South Africa

ABSTRACT

This mixed methods study aimed to assess the agreement between coaches ranking of youth rugby league players compared against objective physical performance data and gather coaches' subjective descriptions of their players performance. Five hundred and eight male rugby league players (U16 $n = 255$, U18 $n = 253$) completed a fitness testing battery of anthropometric and physical performance measures. Subsequently, 22 rugby ($n = 11$) and strength and conditioning (S&C) coaches ($n = 11$) ranked each player's physical qualities using a 4-point Likert scale (1 – top 25%; 2–25–50%; 3–50–75%; and 4 – bottom 25%) and described their performance. U16 S&C coaches displayed fair agreement when assessing players body mass (39.3%, $\kappa = 0.20$). U18 rugby coaches demonstrated fair agreement for strength and size (42.5%, $\kappa = 0.23$) and body mass (48.7%, $\kappa = 0.31$) whilst both U18 rugby and S&C coaches showed fair agreement levels for endurance (39.8%, $\kappa = 0.25$, 44.3%, $\kappa = 0.29$), respectively. Three higher-order themes were identified from coaches' descriptions of players including physical, rugby and attitude characteristics when evaluating performance. Overall, coaches cannot accurately assess players physical performance against fitness testing data. Though, findings suggest coaches adopt a multidimensional approach when evaluating players performance. Practitioners within talent development systems should utilise both objective and subjective assessments when making decisions regarding players performance.

KEYWORDS

Talent identification; fitness testing; performance; coach rating; evaluation

Highlights

- Rugby and S&C coaches cannot accurately assess all aspects of players physical performance.
- The greatest assessment agreement was for body mass, strength and size, and endurance, while the poorest were for strength, acceleration, and maximum speed.
- Rugby and S&C coaches considered rugby, physical and attitude attributes when evaluating players.
- Findings highlight the complex nature of physical profiling. Subjective and objective measures are required to provide an accurate description of players physical performance.

Introduction

Physical performance is commonly measured using fitness tests (Jones, Till, Manley, & Mcguigan, 2017) to monitor and evaluate athletes, inform training programmes (Mccormack, Jones, & Till, 2020), support decision making (Reeves, Mcrobert, Lewis, & Roberts, 2019), and enhance athlete motivation (Mccormack, Jones, Scantlebury, Rotheram, & Till, 2020). Anthropometric (e.g. height, body mass) and physical (e.g. speed, strength, power, endurance) qualities are typically assessed to provide comparisons against reference values (Till, Morris, Emmonds, Jones, & Cobley, 2018a), facilitating the interpretation of scores and supporting players through performance pathways. Moreover, fitness testing data are useful for player development staff as it can be used to provide an indication of future professional career attainment (Till et al., 2016; Till et al., 2017b). However, fitness testing has been termed highly individual and variable in nature, and not fully representative of sporting

CONTACT Sam McCormack sam.mccormack@leedsbeckett.ac.uk Carnegie Applied Rugby Research (CARR) Centre, Carnegie School of Sport, Leeds Beckett University, Cavendish G12, Headingley Campus, Leeds LS6 3QU, UK

performance (Mccormack et al., 2020a). Therefore, evaluating current and future athletic ability is a persistent challenge for researchers and practitioners (Rees et al., 2016).

With sporting performance being underpinned by a plethora of attributes (i.e. physical, technical, tactical, psychological; (Larkin & O'connor, 2017)), singular isolated measures of performance (e.g. fitness testing) may not provide sufficient information on a player's ability (Cripps, Hopper, & Joyce, 2016; Mccormack et al., 2020a). Typically, sport coaches are tasked with selecting or de-selecting players based on their subjective perception of a players long-term potential (Cripps, Hopper, & Joyce, 2019). Moreover, such decisions are informed by objective data whilst using the "coaches' eye" (Höner, Leyhr, & Kelava, 2017). However, to date, limited research has examined such processes. For example, a recent investigation by Dugdale, Sanders, Myers, Williams, and Hunter (2020) examined levels of agreement between objective (i.e. fitness measures) and subjective (i.e. coach ratings of fitness measures) measures within youth soccer (Dugdale et al., 2020). Findings suggested that coaches' subjective assessments met objective criteria with the highest- and lowest-performing players, but results differed when assessing players with similar physical qualities. Moreover, a combination of coach assessment and performance tests were the most accurate predictors of career attainment in youth soccer (Sieghartsleitner, Zuber, Zibung, & Conzelmann, 2019). Additionally, recent findings from rugby league practitioners also suggested they are able to subjectively assess physical performance without the use of objective measures (Mccormack et al., 2020a).

There has been numerous investigations examining the physical qualities of rugby league players (e.g. (Dobbin, Highton, Moss, & Twist, 2019; Dobbin, Hunwicks, Highton, & Twist, 2017a; Ireton, Till, Weaving, & Jones, 2019; Till et al., 2016; Till, Darrall-Jones, Weakley, Roe, & Jones, 2017a)). Typically, objective measures (i.e. fitness test batteries) are used to compare between positions, playing level and age grade to support talent identification processes. Moreover, there has been previous research investigating coaches opinions on different factors influencing rugby league development and performance (Cupples & O'connor, 2011; Mccormack et al., 2020a; Mccormack et al., 2020b). However, it remains to be determined if rugby league and strength and conditioning (S&C) coaches can subjectively assess the physical performance of players, and the importance coaches place on certain attributes when assessing player performance and potential. Typically, subjective coach decisions are classified as the gold standard and have a significant role in talent identification (Roberts, Greenwood, Humberstone, & Raynor, 2020). With both rugby and S&C coaches being routinely involved in recruitment, decision making processes (Till, Muir, Abraham, Piggott, & Tee, 2019), talent identification, programming and long-term athletic development, such information is required to support multidimensional approaches within talent development.

Therefore, by adopting a mixed-method study design, the primary aim of this study was to assess the level of agreement between player rankings based on objective physical performance data and subjective rankings made by S&C and rugby league coaches. A secondary aim was to examine coaches' subjective descriptions of their players physical performance.

Materials and methods

Study design

A fitness testing battery was completed at nine rugby league club academies during a pre-season period (November – December 2019). This fitness testing battery included anthropometric (height, body mass, body fat) and physical performance (isometric mid-thigh pull [IMTP], countermovement jump [CMJ], 10–40 m speed and prone YoYo intermittent recovery test level 1) measures. Following testing, each club received their results alongside an anonymised national dataset for comparison purposes. The rugby and S&C coaches of each club were then subsequently recruited via email to take part in an online survey. Approximately two months after fitness testing was completed (January – February 2020), coaches were asked to rank each of their players physical performance and subjectively explain and describe their players performance and ranking. An encrypted personalised survey was administered online and contained their players names alongside a drop-down menu for rankings, and a section to provide a short rationale for their assessment. Prior to all experimental procedures, ethics approval was granted (application reference 72900).

Participants

Coaches

Twenty-two male rugby league practitioners (mean ± SD, age: 32.0 ± 5.7 years; coaching experience: rugby; 12.1 ± 3.6, S&C; 7.0 ± 3.2 years) participated in this study. All participants were classified as either a rugby (U16 $n = 4$, U18 $n = 7$) or S&C coach (U16 $n = 4$, U18 $n = 7$) and had worked with their players during the time of fitness testing and for at least one full

season. Coaches provided informed consent prior to competing the survey.

Players

Five hundred and eight male rugby league players (age 16.7 ± 1.4 years) participated in this study and were categorised as either U16 ($n = 255$) or U18 ($n = 253$). All players were associated with a Rugby Football League Licenced Academy and provided informed consent. Parental consent was obtained when a player was under 18 years at the time of data collection.

Procedures

All fitness testing measures were gathered within a single session at each clubs training ground, and testing procedures were completed by the research team to ensure standardisation. Players provided information on their date of birth and playing position and completed a standardised warm-up.

Fitness testing battery

Anthropometry

Body mass and height were measured to the nearest 0.1 kg and 0.1 cm using a portable stadiometer (Seca 213, Hamburg, Germany) and analogue scales (Seca, Hamburg, Germany). Body fat was measured using a bio-impedance analyser (Tanita BF-350, Tokyo, Japan). Participants were instructed to wear minimal clothing (i.e. light t-shirt, shorts).

Muscular power

The CMJ was performed with hands on hips and each foot placed on an individual force plate (PS-2141, Pasco, Roseville, California, USA) (Lake et al., 2018). Participants completed two trials and were instructed to start in a standing position and drop to a self-selected depth before immediately jumping as high as possible (Weakley et al., 2019). The intraclass correlation coefficient (ICC) and coefficient of variation (CV) for CMJ height were 0.85 and 3.8%, respectively. The highest jump height (cm) was recorded and used as a measure of "power".

Muscular strength

The IMTP was used to provide a measure of full-body strength using a custom-built dynamometer (Takei Scientific Instruments, Niigata, Japan). Participants followed the protocol outlined by Till and colleagues (2018). A strong significant relationship has been identified between the peak force derived from a dynamometer and a force platform ($r = 0.92$, $p < 0.001$) (Dobbin et al., 2017b) within a similar cohort. In addition, the dynamometer has shown acceptable between day reliability (CV = 5.5% [4.5–6.9]) (Sawczuk et al., 2018). The highest absolute score (N) was used as a measure of "strength".

Speed

Speed was measured over 10, 20, 30 and 40 m using photoelectric timing gates (Brower Timing Systems, Draper, UT, USA). Participants stood with their front foot 0.5 m from the first timing gate (Darrall-Jones, Jones, Roe, & Till, 2016) in a two-point start and set off in their own time. All participants completed 2 maximal sprints with 3 min rest between repetitions. Maximum speed was calculated by dividing the fastest 10 m split time by the distance between splits (10 m). All times were measured to the nearest 0.01 sec. Previous research has reported Brower timing systems to be reliable when measuring 10, 20, 30 and 40 m sprints with mean typical errors expressed as CV of 2.5%, 2.2%, 2.2% and 1.8%, respectively (Sawczuk et al., 2018). The fastest 10 m time and maximum speed achieved were used as measures of "acceleration" and "maximum speed", respectively.

Endurance

The prone Yo-Yo IR1 was completed according to protocols outlined previously (Dobbin, Highton, Moss, Hunwicks, & Twist, 2018a). The final level and distance achieved (m) were recorded following the second failed attempt to complete the shuttle in the allocated time, or volitional exhaustion. The reliability (CV = 9.9%) (Dobbin, Hunwicks, Highton, & Twist, 2018b) and concurrent validity have previously been reported (Dobbin et al., 2018a). The final distance achieved was used as a measure of "endurance".

Coach rankings

Each coach was provided a personalised spreadsheet (Google sheets) containing the players' names from their club who had previously completed all tests. The coaches were asked to rank their physical performance based on their retrospective fitness test results. Coaches used a 4-point Likert scale (1 – top 25%; 2–25–50%; 3–50–75%; and 4 – bottom 25%) to rank eight physical qualities; "strength and size"; "power and speed"; "body mass"; "strength"; "power"; "acceleration"; "maximum speed"; and "endurance". Similar studies using coach rankings have previously been implemented (Dugdale et al., 2020; Hill, Scott, Mcgee, & Cumming, 2020b). In addition, the survey was pilot tested with four rugby league and S&C coaches prior

to the commencement of the study to identify any issues with reliability.

U16 coaches were required to complete two questions: (1) rank each of your player's current physical performance within the U16 playing group and (2) rank each of your player's physical performance versus an U18 player. U18 coaches were required to complete one question: (1) rank each of your player's current physical performance within the U18 playing group. U16 and U18 players were compared against the national dataset at their respective age group. In addition, the U16 players were compared against the U18 dataset to provide an indication of future performance. Participants were asked for a subjective qualitative description for each player to justify their rankings (e.g. why did you give this player these ratings?). This open-ended question invited coaches to further explain their assessment.

Coaches' subjective rankings were subsequently compared with the players objective quantile data. "Strength" and "power" were classified as peak force (N) from the IMTP and maximum jump height attained during the CMJ, respectively. "Acceleration" and "maximum speed" were gathered from 10 m time and maximum velocity attained during sprint tests, respectively. "Endurance" was classified as final level achieved during the prone YoYo IR1. "Body mass" – derived from body mass results. Lastly, "strength and size" and "power and speed" were derived from principal component analysis. All testing variables were subjected to principal component analysis as part of the fitness testing project. Two principal components (PC) were extracted and explained the most variance within the dataset. PC1 explained strength and size characteristics (i.e. "strength and size"), while power and speed qualities contributed to PC2 (i.e. "power and speed"). These results were presented to practitioners alongside descriptive fitness testing data to provide an overall fitness test result (Mccormack et al., 2021). Each club received their results alongside an anonymised national dataset allowing for comparison to be made against respective positions and age groups. In addition, coaches understood the classification of PCs which allowed for the concise visualisation of physical qualities for practitioners to enhance longitudinal evaluation and development. As such, "strength and size" and "power and speed" were ranked separately to the individual qualities of "strength", "body mass", "power" and "speed".

Data analysis

Descriptive frequencies and Cohens Kappa (κ) with 95% confidence intervals were used to determine the level of agreement between subjective (i.e. coach rankings) and objective data (i.e. fitness testing data). The level of agreement for κ was determined as; <0 less than chance agreement, 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and 0.81–0.99 almost perfect agreement (Landis & Koch, 1977). All data analyses were completed in R Studio (version 4.0.2) using the CohenKappa function from the irr package (Matthias Gamer, Fellows, & Singh, 2019) with statistical significance set at p <0.05. Objective fitness testing data were classified as quantiles; top 25%, 25–50%, 50–75% and bottom 25% (see Table 1).

The open-ended qualitative responses were analysed using hierarchical content analysis (Braun, Clarke, & Weate, 2016). Raw data responses were read on numerous occasions to ensure familiarity and permit data immersion. Next, responses were labelled as codes to represent meaningful information. Codes were connected and ordered inductively from participants responses with the absence of a pre-determined framework. These codes were classified into sub-themes and then into larger, more inclusively meaningful themes. Each theme was given a title that represented the codes and categories it contained. Following this, raw data responses, sub-themes and themes were examined thoroughly again by the primary investigators to ensure all information were represented. The research team engaged in constant discussion to cross check and confirm the distribution of data and ensure correct definition and naming (Braun et al., 2016). Finally, the analysis was reviewed, and results ordered in a figure to display themes generated.

Results

Coach rankings

Table 2 displays the agreement levels between subjective and corresponding objective data. U16 S&C coaches exhibited a fair level of agreement for ranking body mass (39.3%, κ = 0.20, 95% CI = 0.13–0.27, p < 0.001). In addition, U16 S&C coaches demonstrated fair levels of agreement when assessing their players power (43.4%, κ = 0.22, 95% CI = 0.15–0.29, p < 0.001) and body mass (43.4%, κ = 0.23, 95% CI = 0.15–0.30, p < 0.001) versus an U18 player. U18 rugby coaches displayed fair agreement for strength and size (42.5%, κ = 0.23, 95% CI = 0.16–0.31, p < 0.001), and body mass (48.7%, κ = 0.31, 95% CI = 0.24–0.38, p < 0.001). U18 rugby and S&C coaches demonstrated 39.8% and 44.3% agreement for endurance (κ = 0.25, 95% CI = 0.17–0.32, p < 0.001; κ = 0.29, 95% CI = 0.22–0.37, p <

Table 1. Proportion (%) of agreement between practitioners' (rugby coach, S&C coach) subjective assessment and objective fitness testing data.

	U16s				U16s & U18s				U18s			
	Coach %	Cohen's Kappa (k)	S&C %	Cohen's Kappa (k)	Coach %	Cohen's Kappa (k)	S&C %	Cohen's Kappa (k)	Coach %	Cohen's Kappa (k)	S&C %	Cohen's Kappa (k)
Strength and size (PC1)	25.5	0.04 (−0.03 – 0.11) $p = 0.379$	31.0	0.10 (0.03–0.17) $p = 0.01$	21.7	−0.003 (−0.07–0.06) $p = 0.948$	37.9	0.14 (0.07–0.21) $p = 0.002$	42.5	0.23 (fair; 0.16–0.31) $p = <0.001$	36.4	0.15 (0.08–0.22) $p = <0.001$
Power and speed (PC2)	41.5	0.19 (0.11–0.26) $p = <0.001$	35.9	0.15 (0.08–0.22) $p = 0.0004$	29.3	0.07 (0.0–0.14) $p = 0.181$	37.9	0.16 (0.09–0.23) $p = 0.0003$	28.3	0.045 (−0.02–0.11) $p = 0.377$	39.3	0.19 (0.12–0.26) $p = <0.0001$
Body mass	27.4	0.08 (0.0–0.15) $p = 0.08$	39.3	0.20 (fair; 0.13–0.27) $p = <0.001$	28.3	0.08 (0.00–0.15) $p = 0.114$	43.4	0.23 (fair; 0.15–0.30) $p = <0.001$	48.7	0.31 (fair; 0.24–0.38) $p = <0.001$	27.9	0.04 (−0.02–0.11) $p = 0.327$
Strength	25.5	−0.04 (−0.12–0.02) $p = 0.39$	31.7	0.10 (0.03–0.18) $p = 0.014$	23.9	−0.004 (−0.07–0.06) $p = 0.942$	33.8	0.10 (0.02–0.17) $p = 0.027$	35.4	0.17 (0.09–0.24) $p = <0.0001$	28.6	0.04 (−0.02–0.11) $p = 0.344$
Power	32.1	0.11 (0.03–0.18) $p = 0.0227$	35.2	0.13 (0.06–0.21) $p = 0.002$	35.9	0.14 (0.07–0.22) $p = 0.009$	43.4	0.22 (fair; 0.15–0.29) $p = <0.001$	23.0	−0.01 (−0.08–0.05) $p = 0.751$	32.9	0.11 (0.04–0.18) $p = 0.015$
Acceleration	17.9	−0.02 (−0.09–0.04) $p = 0.562$	19.3	−0.08 (−0.16–0.01) $p = 0.059$	16.3	−0.10 (−0.17–0.03) $p = 0.055$	22.1	−0.003 (−0.07–0.06) $p = 0.935$	23.9	−0.03 (−0.10–0.03) $p = 0.538$	17.1	−0.08 (−0.15—0.01) $p = 0.064$
Max speed	18.9	0.01 (−0.05–0.08) $p = 0.729$	18.6	−0.06 (−0.13–0.00) $p = 0.134$	21.7	−0.003 (−0.07–0.06) $p = 0.95$	33.1	0.12 (0.04–0.19) $p = 0.006$	33.6	0.12 (0.05–0.19) $p = 0.012$	38.6	0.19 (0.12–0.26) $p = <0.0001$
Endurance	37.7	0.16 (0.08–0.23) $p = 0.002$	36.6	0.16 (0.09–0.23) $p = 0.0002$	34.8	0.14 (0.07–0.21) $p = 0.008$	39.3	0.19 (0.12–0.27) $p = <0.0001$	39.8	0.25 (fair; 0.17–0.32) $p = <0.001$	44.3	0.29 (fair; 0.22–0.37) $p = <0.001$

Note: Data are presented as % agreement, Cohen's Kappa (κ) (95% CI) and p value. Fair inferences (0.21–0.40) highlighted in bold.

0.001), respectively. All other agreements between objective and subjective data were slight.

Subjective evaluations

Open-ended responses to the participants assessment of performance, identified three distinct higher-order themes: (1) physical, (2) rugby and (3) attitude (Figure 1). The higher-order theme of physical related to players physical performance and was composed of sub-themes including "two sides of a coin", "good or bad" and "the all-rounder". Similarly, the theme of rugby encompassed factors related to players sport-specific performance and consisted of sub-themes of "transfer to performance", "good or bad", and "potential". Lastly, the theme of attitude was inclusive of aspects associated with players mindset and included sub-themes of "positive" and "negative".

Physical

When assessing players' performance, coaches typically acknowledged their physical qualities. The most

Table 2. Quantile classification of fitness testing data for each age group.

	Strength and size (PC1)	Power and speed (PC2)	Body mass (kg)	Strength (N)	Power (CMJ)	Acceleration (10 m; sec)	Max speed (m/s)	Endurance (Prone IR1 Level)
U16 and U18								
Quantile 1	−0.75	−0.78	74.1	1540	28.4	1.76	7.75	14.4
Quantile 2	−0.03	−0.07	81.9	1754	32.6	1.82	8.20	14.8
Quantile 3	0.74	0.68	90.2	2005	37.4	1.88	8.62	15.5
U16								
Quantile 1	−1.20	−1.05	69.9	1400	26.7	1.78	7.63	14.2
Quantile 2	−0.46	−0.36	77.7	1574	30.3	1.85	8.06	14.6
Quantile 3	0.15	0.34	86.1	1763	34.8	1.91	8.47	15.3
U18								
Quantile 1	−0.15	−0.38	79.3	1728	31.9	1.74	8.20	14.5
Quantile 2	0.39	0.26	85.1	1949	36.3	1.79	8.55	15.0
Quantile 3	0.90	0.84	92.3	2178	40.1	1.85	8.85	15.6

Note: Data are presented as mean for each quantile. E.g. U16 body mass – bottom 25%; 0–69.9 kg, 50–75%; 70–77.7 kg, 25–50%; 77.8–86.1 kg, top 25%; 86.2 kg and above.

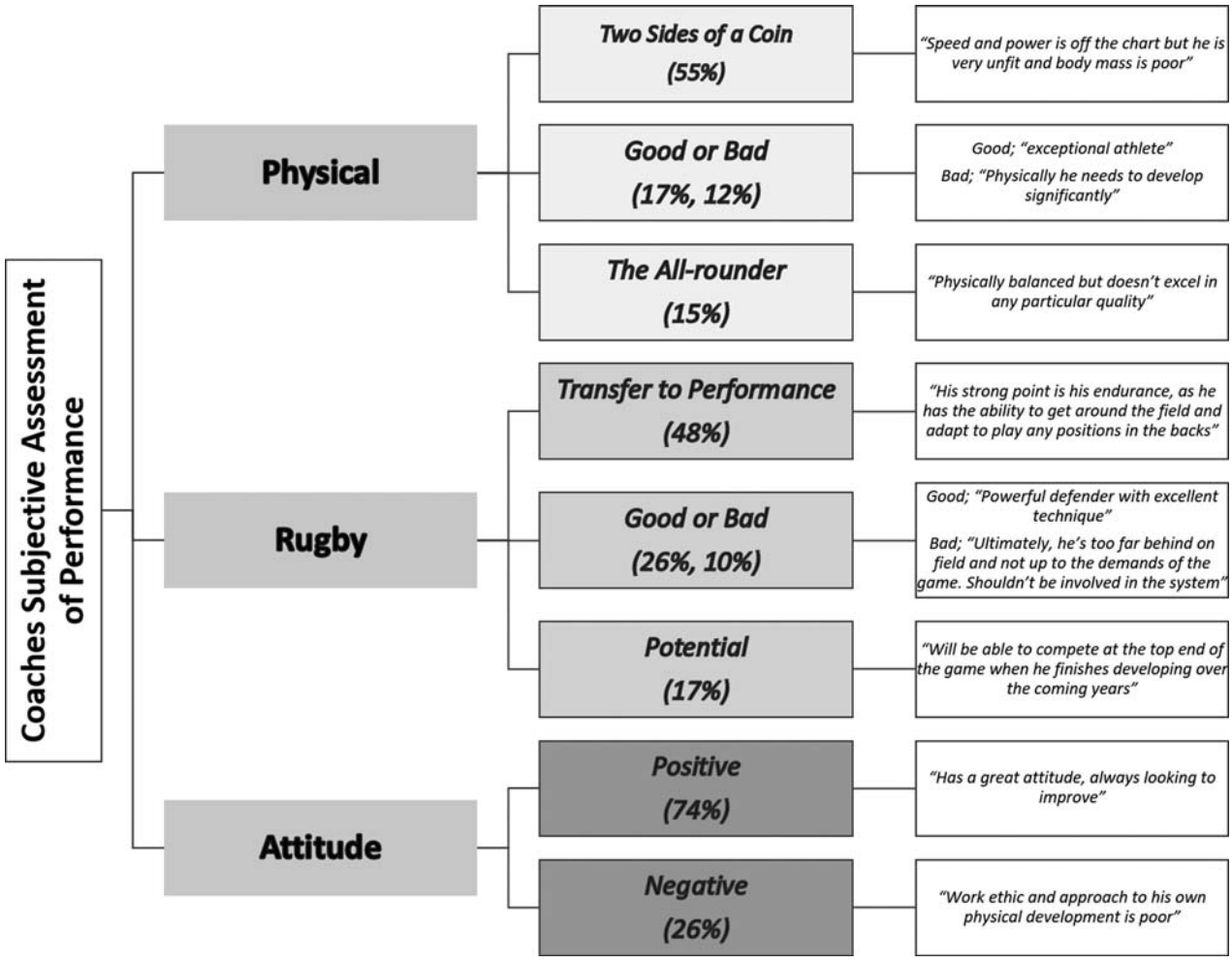

Figure 1. Schematic of themes and sub-themes (percentage of responses) associated with practitioner's subjective assessment of performance.

common form of assessment (55%; i.e. percentage of responses that referred to sub-theme) was classed as "two sides of a coin" and refers to coaches offering a positive comment as well as highlighting an apparent weakness. For example, coaches descriptions included "Struggles with strength and power, but performs well in sprinting", "some great physical traits, but too slow", "very strong and powerful, struggles with fitness side of the game" and "excellent athlete who is very fast and powerful with good relative strength and condition. However, he lacks in size and robustness".

The second sub-theme identified within coaches' assessments of physical performance was "good or bad" and highlights the inverse nature of coaches' opinions. Coaches classified players as either good (17%); "Very strong for his size", "exceptional athlete", "top end player physically" and "very good physically". In contrast, players physical performance was identified as being at the opposite end of the spectrum (i.e. bad [12%]); "not really strong in any physical or rugby areas", "very small and slight", "another pivot who is too slow", "needs to develop a lot still, physically", "his physical fitness is his biggest developmental need", and "too high body mass, struggles with all physical qualities".

Lastly, the coaches referred to "the all-rounder" when assessing players performance (15%), which made up the final sub-theme of physical performance. For example, "XXX is a great all-round athlete", "physically balanced, but doesn't excel in any particular quality" was the way in which two S&C coaches described player's physical performance. This theme was reinforced with two coaches stating, "not bad at anything but doesn't excel in any particular quality either" and "good physical qualities across the board".

Rugby

This second higher-order theme encompassed practitioners referring to players' "rugby" ability within their assessment. Both rugby and S&C coaches

attempted to link both physical and rugby performance. This "transfer to performance" (48% of responses) was typically outlined as being positive "is powerful and strong. Players at this age struggle with him", "extremely strong and powerful. Very hard to put down", "powerful and deceptively strong on the field in games" and "strengths are generally around speed, especially in specific situations as a hooker". The coaches identified how there was a "positive transfer" from physical to field performance. However, in contrast, coaches recognised that physical performance does not necessarily transfer to rugby performance with one S&C coach citing; "is a strong and big player but lacks the acceleration and power on the field". Moreover, one rugby coach stated how a player has; "good size and very strong but needs to transfer into speed and acceleration as a winger". In addition, another rugby coach highlighted how their player's "(rugby) performance doesn't align to (fitness) data" indicating a mismatch between fitness testing and rugby performance. This was further justified by one rugby coach as he described a player as

> quick and agile – very elusive. Also, strong compared to bodyweight. His bodyweight is nowhere near where it should be so struggles with demands of the game. Will be his physical side and approach that lets him down – not footballing ability.

Similar to the first high-order theme, rugby was usually characterised as "good or bad", with coaches recognising the paradoxical nature of players performance. Good rugby performance (26%) was defined by rugby and S&C coaches as; "extremely strong in close contact", "deceptively strong on the field and in games", "plays above his weight and is very quick around the ruck" "very big and powerful, but also has great agility and skill" "just a very, very good athlete on the field" and "very consistent in games and training". On the contrary, coaches identified the poorer performing players (10%) as "ultimately, he's too far behind on field and not up to the demands of the game. Shouldn't be involved in the system". One S&C coach cited how a player was "quite often dominated and put on his back" and "lacks in his endurance in terms of playing long spells in at 9". Lastly, a rugby coach expressed how a players' "developmental needs are rugby IQ related".

The final sub-theme associated with rugby was players perceived "potential" (17%). "Excellent prospect, developing nicely" and how a player was "still developing but will be very good" were ways in which rugby coaches classified players. In addition, one S&C coach articulated that a player was a "late developer and late to the game but on a very hopeful pathway". Moreover, S&C coaches stated, "top athlete with high potential" and another; "will be able to compete at top end of game when finished developing over the coming years".

Attitude

Both rugby and S&C coaches routinely referred to the final higher-order theme of "attitude" throughout their responses. The most common responses were categorised as either "positive" or "negative". The subtheme of "positive" (74%) was associated with players' "motivation" for training. One S&C coach highlighted how a player completed "extra training outside scholarship sessions" and how a player "works hard but not athletically gifted". Moreover, rugby coaches indicated how some players are "competitive" and "work hard" with mindsets of "always looking to improve". One S&C coach stated, "very consistent and hard working on all physical areas". Comparatively, "negative" (26%) attitude was also represented within coaches' responses with some players "lacking motivation", needing "external motivation to achieve and develop" and not having "much drive in him to succeed". Lastly, one rugby coach stated how a player's negative attitude may have a negative influence in the future: "a great talent physically and a great player. If anything, it's his attitude off-field that may hinder him long-term".

Discussion

Using a mixed method approach, 22 coaches evaluated 508 players to examine the accuracy of subjective assessments of players physical performance. Additionally, we aimed to explore the factors coaches consider when making subjective assessments of player performance. Overall, results showed that coaches cannot accurately rate all aspects of players physical performance. Though, there was a fair level of agreement between rugby and S&C coaches' assessment of players body mass, power, strength and size, and endurance qualities. Furthermore, novel insights were identified from coaches' subjective evaluations, which included physical and rugby performance, and attitude evaluations. To our knowledge, this is the first study to examine rugby league coach's assessment of performance and suggests they typically subjectively evaluate players holistically. It is proposed that rugby league practitioners refer to objective data and utilise multidisciplinary assessments when making decisions regarding players physical performance.

When compared to national fitness testing data, rugby league and S&C coaches displayed the most agreement for body mass, endurance, size and strength, and power, however, levels of agreement were relatively

low (i.e. <50% correct assessment). Nonetheless, it may be posited that these qualities are more easily recognised by coaches during rugby league training and match play (Dugdale et al., 2020). However, as coaches demonstrated fair levels of agreement for four of the qualities, the remaining four were not accurately assessed. For example, strength, acceleration and maximum speed qualities typically demonstrated the lowest agreement levels (e.g. ~17–35%), and therefore suggest coaches should refer to respective objective data to either support or defend their opinion, and make evidence-based decisions where possible (Dugdale et al., 2020). Though, it must be acknowledged that such qualities (e.g. acceleration, maximum speed) display marginal differences which may have influenced subjective evaluations.

A lack of agreement between coach's subjective assessment and objective data may be due to several reasons. For instance, evaluating players ability during adolescence is associated with numerous challenges (Cripps et al., 2019 Hill, Scott, Mcgee, & Cumming, 2020a;) due to the individual nature of testing, and the myriad of factors associated with sporting performance (Mccormack et al., 2020a; Ryan et al., 2018). Secondly, the disparity in results may be due to coaches identifying and selecting players for a timepoint in the future versus current performance, and therefore, focussing on potential. Coaches are required to make judgements on players at a young age and select potential talent that may emerge in the future (Roberts et al., 2020). Indeed, coaches often referred to player's "potential" within their assessment, highlighting a focus toward the future. Such findings are important as consistent over or under estimation of a player's physical performance may result in inadequate training prescription by coaches (e.g. prescription of excess training load) (Scantlebury, Till, Sawczuk, Weakley, & Jones, 2018). Therefore, it is important that coaches utilise objective data as evidence to inform training and long-term athletic development plans.

Though our results indicate that coaches display somewhat low agreement accuracy when assessing players' physical performance, unique insights are provided into the multiple variables included within their assessment of performance. Both rugby and S&C coaches considered physical and rugby characteristics when evaluating player performance, which suggests holistic evaluations. Furthermore, talent identification models inclusive of physical, technical and tactical abilities have greater prediction accuracy of talent than single measures alone (Cripps et al., 2019). As such, coaches attempted to establish a link between physical and rugby performance by alluding to the indicators they associate with good or bad performance. These findings are similar to research in Australian rules football (Mcintosh, Kovalchik, & Robertson, 2018) where subjective evaluations typically depend on distinct characteristics. Aligned with previous research (Larkin & O'connor, 2017; Reeves et al., 2019), coaches also considered player's "attitude" within their assessments. The positive influence of attitude was associated with a player's willingness to learn and develop, and such positive attributes were "hard working", "looking to improve" and "motivated". In contrast, a negative attitude was classified as "lacking motivation", "approach to training" and "personality off the pitch". And previous research indicates these psychological traits to be negatively associated with team performance and dynamics (Larkin & O'connor, 2017), and selection in Australian rules football (Larkin, Marchant, Syder, & Farrow, 2020). Therefore, it may be posited that coaches' perceptions of player's (positive or negative) attitude may have influenced their rankings (i.e. low agreement). As such, these findings further support the use of multidimensional methods to assess physical performance.

Both rugby and S&C coaches referred to their players' physical qualities throughout, and often attempted to categorise them (e.g. "one of the strongest [in their team]"). However, according to the quantitative data, these assumptions are often incorrect, which questions the validity of the "coaches' eye" (Schorer, Rienhoff, Fischer, & Baker, 2017). As such, the results highlight the importance of using objective data over subjective opinion when making inferences regarding players physical performance. Moreover, as objective data are routinely gathered and stored to support practice (e.g. longitudinal evaluation, training prescription), we propose the same for subjective data. This process may encourage practitioners to be more aware of their assessments,hile also quantifying their opinions. Anecdotally, subjective coach opinions may be discussed and shared within a multidisciplinary team, however, are not recorded in a similar fashion as objective data. Therefore, the collation of subjective reports may further support talent identification and development systems by providing a multidimensional, rigorous evidence base. In addition, recording subjective coach assessments may reduce the number of non-measured attributes (e.g. technical ability), while providing further detail on such.

Interestingly, there were limited differences in the agreement levels between rugby and S&C coaches' assessments of players performance. Though, U18 rugby coaches typically displayed greater agreement than S&C coaches. Such findings are noteworthy, especially given the S&C coaches' role is primarily

focussed on the physical development of players. Additionally, rugby coaches displayed greater coaching experience which may have increased the accuracy of their subjective assessment. The results potentially highlight poor accuracy of an S&C "coaches' eye". A possible explanation of this is an overreliance of objective data, and a reduction in general observation by S&C coaches. On the other hand, rugby coaches are likely more discerned with players' technical/tactical ability, anticipatory skills, and decision-making, and therefore, consider more in-depth, contextual information from other dimensions, whilst not focussing solely on physical attributes. As a result, may influence their coaches' eye, and subsequently subjective evaluations. Therefore, the rugby coach's primary role offers an explanation for the lack of synchronicity between subjective and objective data. However, both rugby and S&C coaches require a multidimensional perspective, covering all aspects of player development, and adopting an individual focus is likely to miss pieces of the puzzle.

The current results highlight the complexity of assessing players physical performance. With an absence of unequivocal agreement between coach's perceptions and actual performance, both subjective and objective measures to evaluate performance is suggested. Owing to the fact that selection and decision making largely hinge on coaches subjective assessment of their players, objective data are needed to provide support for or against such decisions (Jokuschies, Gut, & Conzelmann, 2017) and should be prioritised over subjective opinion where possible. Furthermore, subjective coach evaluation should be used where objective data are not available Additionally, coaches ought to be cognisant of the apparent mismatch between their opinion and actual performance, alongside their subjective tendencies, which may lead to regular over or under rating of players (Mcintosh et al., 2018). Coaches should firstly consider their own subjective assessments of physical performance and then discuss their beliefs and philosophies with colleagues to better understand such complex processes. As a result, a common physical performance assessment or performance profile (Jones et al., 2017) which has been subjectively informed by coaches would contribute to the systematic assessment of players and improve decision-making processes.

Although this is the first study to use a mixed method approach to player evaluations in rugby league, the study has several limitations. Firstly, as the study was carried out approximately two months after fitness testing was conducted, the retrospective nature may have reduced coaches' prediction accuracy of their players physical qualities. Moreover, fitness testing was conducted in pre-season, therefore results may likely change throughout the season. As such, subjective evaluations should be implemented in conjunction with fitness testing to promote synchronicity. Lastly, the fitness tests used in this study are independent to the sport-specific nuances associated with rugby league and should not be considered as isolated measures of performance.

Conclusion

Rugby league and S&C coaches cannot accurately assess all aspects of players physical performance. Our findings suggest coaches considered physical and rugby performance, and attitude when making subjective evaluations. The assessment of rugby league players' physical performance is a complex process, and a mismatch between coaches' assessment and objective fitness testing data may lead to several issues. Coaches displayed the highest agreement when assessing players body mass, strength and size, power, and endurance, and assessments were most erroneous for strength, acceleration, and maximum speed qualities. Due to the somewhat lack of agreement between coaches' subjective assessment and objective fitness testing data, our results demonstrate that practitioners should use objective data over subjective opinion to provide a more accurate description and support decision making when evaluating players performance. Additionally, implementing multidimensional assessments are suggested to optimise talent development. Therefore, the combination of subjective coach assessment and a battery of objective tests would contribute to a dynamic context and provide a more holistic evaluation of player performance.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Sam McCormack http://orcid.org/0000-0001-6994-178X
Ben Jones http://orcid.org/0000-0002-4274-6236
Kevin Till http://orcid.org/0000-0002-9686-0536

References

Braun, V., Clarke, V., & Weate, P. (2016). *Using thematic analysis in sport and exercise research. Routledge handbook of qualitative research in sport and exercise*. London: Routledge.

Cripps, A. J., Hopper, L., & Joyce, C. (2016). Maturity, physical ability, technical skill and coaches' perception of semi-elite adolescent Australian footballers. *Pediatric Exercise Science*, *28*, 535–541. doi:10.1123/pes.2015-0238

Cripps, A. J., Hopper, L. S., & Joyce, C. (2019). Can coaches predict long-term career attainment outcomes in adolescent athletes? *International Journal of Sports Science & Coaching*, *14*, 324–328.

Cupples, B., & O'connor, D. (2011). The development of position-specific performance indicators in elite youth rugby league: A coach's perspective. *International Journal of Sports Science & Coaching*, *6*, 125–141. doi:10.1260/1747-9541.6.1.125

Darrall-Jones, J. D., Jones, B., Roe, G., & Till, K. (2016). Reliability and usefulness of linear sprint testing in adolescent rugby union and league players. *The Journal of Strength & Conditioning Research*, *30*, 1359–1364.

Dobbin, N., Highton, J., Moss, S. L., Hunwicks, R., & Twist, C. (2018a). Concurrent validity of a rugby-specific Yo-Yo intermittent recovery test (Level 1) for assessing match-related running performance. *Journal of Strength and Conditioning Research*, *35*(1), 176–182.

Dobbin, N., Highton, J., Moss, S. L., & Twist, C. (2019). Factors affecting the anthropometric and physical characteristics of Elite Academy Rugby League Players: A multiclub study. *International Journal of Sports Physiology and Performance*, *14*, 958–965. doi:10.1123/ijspp.2018-0631

Dobbin, N., Hunwicks, R., Highton, J., & Twist, C. (2017a). Reliable testing battery for assessing physical qualities of Elite Academy Rugby League Players. *Journal of Strength and Conditioning Research*, doi:10.1519/JSC.0000000000002280

Dobbin, N., Hunwicks, R., Highton, J., & Twist, C. (2018b). A reliable testing battery for assessing physical qualities of elite academy rugby league players. *The Journal of Strength & Conditioning Research*, *32*, 3232–3238.

Dobbin, N., Hunwicks, R., Jones, B., Till, K., Highton, J., & Twist, C. (2017b). Criterion and construct validity of an isometric mid-thigh-pull dynamometer for assessing whole-body strength in professional rugby league players. *International Journal of Sports Physiology and Performance*, *13*, 235–239.

Dugdale, J. H., Sanders, D., Myers, T., Williams, A. M., & Hunter, A. M. (2020). A case study comparison of objective and subjective evaluation methods of physical qualities in youth soccer players. *Journal of Sports Sciences*, *38*, 1304–1312.

Hill, M., Scott, S., Mcgee, D., & Cumming, S. (2020a). Coaches' evaluations of match performance in academy soccer players in relation to the adolescent growth spurt. *Journal of Science in Sport and Exercise*, *2*(4), 359–366.

Hill, M., Scott, S., Mcgee, D., & Cumming, S. P. (2020b). Are relative age and biological ages associated with coaches' evaluations of match performance in male academy soccer players? *International Journal of Sports Science & Coaching*, 1747954120966886.

Höner, O., Leyhr, D., & Kelava, A. (2017). The influence of speed abilities and technical skills in early adolescence on adult success in soccer: A long-term prospective analysis using ANOVA and SEM approaches. *PloS one*, *12*, e0182211.

Ireton, M. R. E., Till, K., Weaving, D., & Jones, B. (2019). Differences in the movement skills and physical qualities of elite senior and Academy Rugby League Players. *Journal of Strength and Conditioning Research*, *33*, 1328–1338. doi:10.1519/JSC.0000000000002016

Jokuschies, N., Gut, V., & Conzelmann, A. (2017). Systematizing coaches' 'eye for talent': Player assessments based on expert coaches' subjective talent criteria in top-level youth soccer. *International Journal of Sports Science & Coaching*, *12*, 565–576.

Jones, B., Till, K., Manley, A., & Mcguigan, M. R. (2017). A multidisciplinary approach to the profiling and interpretation of fitness testing data: A case study example. *Journal Of Australian Strength And Conditioning*, *25*, 31–36.

Lake, J., Mundy, P., Comfort, P., Mcmahon, J. J., Suchomel, T. J., & Carden, P. (2018). Concurrent validity of a portable force plate using vertical jump force–time characteristics. *Journal of Applied Biomechanics*, *34*, 410–413.

Landis, J. R., & Koch, G. G. (1977).. The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Larkin, P., Marchant, D., Syder, A., & Farrow, D. (2020). An eye for talent: The recruiters' role in the Australian Football talent pathway. *PLOS ONE*, *15*, 1–17.

Larkin, P., & O'connor, D. (2017). Talent identification and recruitment in youth soccer: Recruiter's perceptions of the key attributes for player recruitment. *PLOS one*, *12*(4), e0175716. https://doi.org/10.1371/journal.pone.0175716

Matthias Gamer, J. L., Fellows, I., & Singh, P. (2019). irr: Various Coefficients of Interrater Reliability and Agreement.

Mccormack, S., Jones, B., Scantlebury, S., Collins, N., Owen, C., & Till, K. (2021). Using principal component analysis to compare the physical qualities between Academy and International Youth Rugby League Players. *International Journal of Sports Physiology and Performance*, *1*, 1–8.

Mccormack, S., Jones, B., Scantlebury, S., Rotheram, D., & Till, K. (2020a). "It's important, but it's not everything": Practitioners' use, analysis and perceptions of fitness testing in Academy Rugby League. *Sports*, *8*, 130. doi:10.3390/sports8090130

Mccormack, S., Jones, B., & Till, K. (2020b). Training practices of Academy Rugby League and their alignment to physical qualities deemed important for current and future performance. *International Journal of Sports Science & Coaching*, doi:10.1177/1747954120924905

Mcintosh, S., Kovalchik, S., & Robertson, S. (2018). Examination of player role in the Australian Football League using match performance data. *International Journal of Performance Analysis in Sport*, *18*, 451–462.

Rees, T., Hardy, L., Güllich, A., Abernethy, B., Côté, J., Woodman, T., … Warr, C. (2016). The great British medalists project: A review of current knowledge on the development of the world's best sporting talent. *Sports Medicine*, *46*, 1041–1058.

Reeves, M. J., Mcrobert, A. P., Lewis, C. J., & Roberts, S. J. (2019). A case study of the use of verbal reports for talent identification purposes in soccer: A Messi affair!. *PloS one*, *14*, e0225033.

Roberts, A. H., Greenwood, D., Humberstone, C., & Raynor, A. J. (2020). Pilot study on the reliability of the coach's eye: Identifying talent throughout a 4-day Cadet Judo Camp. *Frontiers in Sports and Active Living*, *2*, 196.

Ryan, D., Mccall, A., Fitzpatrick, G., Hennessy, L., Meyer, T., & Mccunn, R. (2018). The influence of maturity status on movement quality among English Premier League academy soccer players. *Sport Perf Sci Rep*, *32*, 1–3.

Sawczuk, T., Jones, B., Scantlebury, S., Weakley, J., Read, D., Costello, N., … Till, K. (2018). Between-day reliability and usefulness of a fitness testing battery in youth sport athletes: Reference data for practitioners. *Measurement in Physical Education and Exercise Science*, *22*, 11–18.

Scantlebury, S., Till, K., Sawczuk, T., Weakley, J., & Jones, B. (2018). Understanding the relationship between coach and

athlete perceptions of training intensity in youth sport. *Journal of Strength and Conditioning Research*, *32*, 3239–3245.

Schorer, J., Rienhoff, R., Fischer, L., & Baker, J. (2017). Long-term prognostic validity of talent selections: Comparing national and regional coaches, laypersons and novices. *Frontiers in Psychology*, *8*, 1146.

Sieghartsleitner, R., Zuber, C., Zibung, M., & Conzelmann, A. (2019). Science or coaches' eye? – both! Beneficial collaboration of multidimensional measurements and coach assessments for efficient talent selection in elite youth football. *Journal of Sports Science & Medicine*, *18*, 32.

Till, K., Cobley, S., Morley, D., O'hara, J., Chapman, C., & Cooke, C. (2016). The influence of age, playing position, anthropometry and fitness on career attainment outcomes in rugby league. *Journal of Sports Sciences*, *34*, 1240–1245. doi:10.1080/02640414.2015.1105380

Till, K., Darrall-Jones, J., Weakley, J. J., Roe, G. A., & Jones, B. L. (2017a). The influence of training age on the annual development of physical qualities within academy rugby league players. *The Journal of Strength & Conditioning Research*, *31*, 2110–2118. doi:10.1519/JSC.0000000000001546

Till, K., Morley, D., O'hara, J., Jones, B. L., Chapman, C., Beggs, C. B., … Cobley, S. (2017b). A retrospective longitudinal analysis of anthropometric and physical qualities that associate with adult career attainment in junior rugby league players. *Journal of Science and Medicine in Sport*, *20*, 1029–1033.

Till, K., Morris, R., Emmonds, S., Jones, B., & Cobley, S. (2018a). Enhancing the evaluation and interpretation of fitness testing data within youth athletes. *Strength & Conditioning Journal*, *40*, 24–33.

Till, K., Morris, R., Stokes, K., Trewartha, G., Twist, C., Dobbin, N., … Jones, B. (2018b). Validity of an isometric midthigh pull dynamometer in male youth athletes. *The Journal of Strength & Conditioning Research*, *32*, 490–493.

Till, K., Muir, B., Abraham, A., Piggott, D., & Tee, J. (2019). A framework for decision-making within strength and conditioning coaching. *Strength & Conditioning Journal*, *41*, 14–26. doi:10.1519/SSC.0000000000000408

Weakley, J. J., Till, K., Darrall-Jones, J., Roe, G. A., Phibbs, P. J., Read, D. B., & Jones, B. L. (2019). Strength and conditioning practices in adolescent rugby players: Relationship with changes in physical qualities. *The Journal of Strength & Conditioning Research*, *33*, 2361–2369.