International Journal of Sports Science & Coaching

Are you lookin' at me? A mixed-methods case study to investigate the influence of coaches' presence on performance testing outcomes in male academy rugby league players

Journal:	*International Journal of Sports Science & Coaching*
Manuscript ID	SPO-22-0259.R1
Manuscript Type:	Original Research Article
Date Submitted by the Author:	28-Jun-2022
Complete List of Authors:	Richardson, Ben; 1School of Science, Technology and Health, York St John University, York Dobbin, Nick; Manchester Metropolitan University, Health Professions White, Christopher; Glyndwr University Bloyce, Daniel; University of Chester Faculty of Medicine Dentistry and Life Sciences, Department of Sport and Exercise Sciences Twist, Craig; University of Chester Faculty of Medicine Dentistry and Life Sciences, Department of Sport and Exercise Sciences
Keywords:	Fitness, Monitoring, Observer, Power, Habitus
Abstract:	The study used a mixed-methods approach to examine how the presence of coaches influenced male academy rugby league players' performance during physical performance testing. Fifteen male rugby players completed two trials of 20 m sprint, countermovement jump and prone Yo-Yo test; one with only the lead researcher present and a second where the lead researcher conducted the battery with both the club's lead S&C coach, academy manager, and the first team assistant and head coach present. Players and coaches then completed one-to-one semi-structured interviews to explore their beliefs, attitudes and opinions towards physical performance testing. In all tests, the players' performance was better when the coaches were present compared to when this was conducted by the sport scientist alone. Interviews revealed performance testing was used by coaches to exercise their power over players to socialise them into a desired culture. Players' own power was evident through additional effort during testing when coaches were present. Practitioners should ensure consistency in the presence of significant observers during performance testing of male rugby players to minimise their influence on test outcome.

Title

Are you lookin' at me? A mixed-methods case study to investigate the influence of coaches' presence on performance testing outcomes in male academy rugby league players

Submission Type: Original article

Authors: Ben Richardson[1], Nick Dobbin[2], Christopher White[3], Daniel Bloyce[4] Craig Twist[4]✉

Affiliations:

[1]School of Science, Technology and Health, York St John University, York, YO31 7EX, UK
[2]Department of Health Professions, Manchester Metropolitan University, Manchester, UK M15 6GX.
[3]Wrexham Glyndwr University, Mold Road, Wrexham, LL11 2AW.
[4]Department of Sport and Exercise Sciences, University of Chester, Parkgate Road, Chester, UK, CH1 4BJ.

✉**Corresponding Author:** Craig Twist, University of Chester, Parkgate Road, Chester, CH1 4BJ, England.
Work Tel: 01244 513441
Work Email: c.twist@chester.ac.uk

Abstract Word Count: 178
Manuscript Word Count: 5202
Tables: 0
Figures: 1

Title

Are you lookin' at me? A mixed-methods case study to investigate the influence of coaches' presence on performance testing outcomes in male academy rugby league players

Abstract

The study used a mixed-methods approach to examine how the presence of coaches influenced male academy rugby league players' performance during physical performance testing. Fifteen male rugby players completed two trials of 20 m sprint, countermovement jump and prone Yo-Yo test; one with only the lead researcher present and a second where the lead researcher conducted the battery with both the club's lead S&C coach, academy manager, and the first team assistant and head coach present. Players and coaches then completed one-to-one semi-structured interviews to explore their beliefs, attitudes and opinions towards physical performance testing. In all tests, the players' performance was better when the coaches were present compared to when this was conducted by the sport scientist alone. Interviews revealed performance testing was used by coaches to exercise their power over players to socialise them into a desired culture. Players' own power was evident through additional effort during testing when coaches were present. Practitioners should ensure consistency in the presence of significant observers during performance testing of male rugby players to minimise their influence on test outcome.

Key words: Fitness, monitoring, observer, power, habitus

Introduction

Team sport athletes require well-developed physical qualities to tolerate the demands of training and match-play [1–3]. Such qualities include speed, strength, power, agility (or change of direction) and intermittent running capacity, as well as possessing the appropriate technical, tactical, psychological and social attributes for success. The assessment of physical qualities is routine practice within team sports, and enables sport scientists, strength and conditioning coaches and skills coaches to assess the development of players over time [4–6], quantify the effectiveness of training interventions [7,8] and reduce the risk of injury [9].

The assessment of such qualities are often practised by the strength and conditioning (S&C) coaches, who have a vested interest in the results, with a positive change in performance indicative of a successful training programme. However, it is often the case that all those involved in the player's development as they progress from youth to professional status (i.e. coach, peers, owner, manager), also have an interest. Furthermore, the use of sport scientists from outside of the professional club, who might collect these data as part of research or wider talent development profiling on behalf of the sport's governing body, increases the surveillance network beyond the central locality of that squad (i.e. Academy) [10]. Nonetheless, the collection of such data tends to be perceived as positive by the players and staff [11], allowing the club to collect constant and normalised data creating a personalised profile that is then used to inform training, recovery or medical treatments [12]. For the players, partaking in performance testing is believed necessary to attain success, and is potentially a mechanism through which they can distinguish individual excellence and demonstrate adherence to 'professional' ideals [10,11,13]. The ability to manipulate the environment in which players complete performance testing is likely to alter their behaviour [14]. For example, if players feel that performance testing

is a suitable opportunity to display superiority over their peers, the presence of a coach, who has the ability to select or deselect players into the senior squad, might, to some degree, alter the players' approach and effort. Several researchers have also highlighted the influence of others' (e.g. coaches) presence during training [15–19], suggesting this can reduce effort perception, increase exercise adherence, intensity, motivation, and improve the adaptive response. While the anthropometric and physical qualities of rugby league players has received considerable interest, few, if any, studies report who was present during these assessments or the attitude of the players and coaches to these practices [10,12]. Such insight might well be important when interpreting any change in performance [15].

Whilst performance testing is common practice, little is currently known about the players' and coaches' views towards these practices as well as their purpose and how this might vary in accordance to the coaches' role. Furthermore, how the data collected is used throughout the club by players, coaches and other members of staff is of particular interest given they are likely to have different uses for this data [20]. McCormack and colleagues' observations revealed the multi-demensional use of objective physical testing beyond its intended use to complement other subjective assessments (e.g. tactical, technical) that inform athlete preparation, selection, standardisation and player motivation [20,21]. Jones et al.[12] also noted the head coach might use particular variables to ensure players were meeting her/his expectation and as a form of disciplinary power. In contrast, they found that the performance analysts and S&C coaches used the data to monitor training load in an attempt to minimise injury risk. A thorough understanding of the views and purpose of performance testing will help practitioners to understand the various challenges of appropriately implementing such tests. Ultimately, this could improve the application of performance testing and promote effective athlete development.

The aim of this study was to use a mixed-methods approach to investigate if, and to what extent, the coaches' presence influenced performance during a standarised testing battery for the assessment of physical qualities. A secondary aim was to explore the opinions of players and coaches towards performance testing within a professional academy rugby league environment, drawing upon the figurational sociological concepts of habitus, power balances and unintended consequences. It was hypothesised that the coach being present during testing would positively impact upon on players' performance through the interaction of power between coach and player. It is also possible that the opinions players held towards testing would influence their perfromance.

Methods

Participants and design

All participants (players n = 15, coaches n = 5) were, at the time of the study, registered to a professional rugby league club with data collected during the preseason training period for the upcoming competitive season. All participants provided written consent to participate and the study was approved by the authors' relevant research ethics committee (1372/17/ND/SES).

In the first stage of this study, 15 male players (stature: 178.9 ± 6.1 cm; body mass: 85.9 ± 10.2 kg) participated in the physical performance testing and deemed themselves to be free from injury, which was confirmed by the club's medical team. The required sample to detect a positive impact of the coaches' presence was derived based on the data presented by Dobbin et al. [22]. The standardised mean difference (SMD) that ranged from 0.45 to 0.75 were inserted into G*Power [23] with α at 0.05 and β at 0.80. The required sample for 20 m sprint time, countermovement jump (CMJ) and prone Yo-Yo Intermittent Recovery Test level 1 (prone

Yo-Yo IR1) was 13, 14 and 32, respectively. However, as we only had access to 15 players, we acknowledge a lower than desired statistical power for the Yo-Yo IR1 test (β = ~50) and the probability of any effect for this outcome should be interpreted with caution. To support interpretation, we have also reported the SMD and 95%CI to indicate the magnitude of the difference and precision of the estimate. Participants were required to abstain from performance-enhancing supplements (e.g., caffeine) and not to have completed any club-based or leisure-time activity of high intensity in the 2 and 24 hours before testing, respectively. The players completed two trials of tests (20 m sprint, CMJ and prone YoYo IR1) selected from the Rugby Football League profiling testing battery [22]; one with only the lead researcher present and a second where the lead researcher conducted the battery with both the club's lead S&C coach, academy manager, and the first team assistant and head coach present. Coaches did not offer any verbal encouragement or motivation to partcipants, simply standing proximal to the testing and within view of all partcipants for the duration of all tests. Trial order was not randomised because of the inablity to blind particpants to the trial conditions. However, all participants were habituated to the physical performance testing procedures having completed the tests several times before as part of routine monitoring at the club. Both trials were conducted 14 days apart in an indoor training facility (Trial 1 = temperature: 6°C, humidity: 89%, pbar: 1001 mbar; Trial 2 = tempearture: 4°C, humidity: 82%, pbar: 1012 mbar) on an artificial grass surface.

In the second part of the study, approximately one week after the testing, 10 of the players who had taken part in the trials and 5 coaches completed one-to-one semi-structured interviews with a single researcher experienced in qualitative research methods. Within qualitative research, sample sizes are determined by the richness of data which is collected, as opposed to a specific number of participants [24,25]. Therefore, alongside the limitations imposed by having access to

only those players and coaches within the one club, a pragmatic approach was taken to determining the number of interview participants required, with the researcher halting data collection once sufficient data had been collected to provide a rich exploration of the research topic [26]. This allowed for the detailed exploration of the players' and coaches' beliefs, attitudes and opinions towards physical performance testing for the assessment of anthropometric and physical qualities. For the players, a second part of the interview also focused on their perceptions towards their performance during the physical performance testing in trial 1 and 2.

The use of quantitative and qualitative methods in conjunction can be problematic due to the differing philosophical assumptions of each approach. Quantitative methods are underpinned by positivism, the belief that reality is singular and that its 'true' nature can be known by researchers maintaining objectivity [27]. Meanwhile, qualitative methods are often underpinned by interpretivism, the belief that realities are multiple and created through the individual's subjective interpretation, therefore a 'truth' of reality cannot be known [27]. To bridge this philosophical divide, a realist approach was taken, whereby reality is viewed as singular and external but is influenced by an individual's subjective interpretation [28,29]. Therefore, this paper sought to establish whether, and why, observer presence impacted performance during the standardised testing battery whilst still assuming that this knowledge may be fallible and is not a definitive 'truth' [28].

Procedures

Stretch stature was measured using a portable stadiometer (Seca, Leicester Height Measure, Hamburg, Germany) to the nearest 0.1 cm, and body mass (Seca, 813, Hamburg, Germany) to the nearest 0.1 kg.

Participants completed a thorough warm-up based on the RAMP principles. Initially activities that increased heart rate, muscle, and blood temperature (e.g., jogging) were performed before activation (e.g., overhead lunge) and mobilisation (e.g., spinal extension). Finally, participants completed a series of activities (e.g., squat jumps) to potentiate the muscle, and concluded with several accelerations (50-100% effort), sprints (75 to 100% effort) and decelerations (100% effort). A period of 5 minutes passive recovery was provided before the initial assessment of sprint performance.

Sprint performance was measured using electronic timing gates (Brower, Speedtrap 2, Brower, Utah, USA) positioned at 0 and 20 m, 150 cm apart and at the height of 90 cm. Participants began each sprint from a two-point athletic stance 30 cm behind the start line. Two maximal 20 m sprints were recorded to the nearest 0.01 s with two minutes between each attempt and the best 20 m sprint time was used for analysis possessing a CV of 3.6% [22].

Participants completed two countermovement jumps with 2-minutes passive recovery between each attempt. Participants placed their hands on their hips and started upright before flexing at the knee to a self-selected depth and extending up for maximal height, keeping their legs straight throughout. Jumps that did not meet the criteria were not recorded, and participants were asked to complete an additional jump. Jump height was recorded using a jump mat (Just Jump System, Probotics, Huntsville, Alabama, USA) and corrected for accuracy[30] before peak height was used for analysis, with a CV of 5.9% [22].

The prone Yo-Yo IR1 required participants to start each 40 m shuttle in a prone position with their head behind the start line, legs straight and chest in contact with the ground. Shuttle speed was dictated by an audio signal commencing at 10 $km \cdot h^{-1}$ and increasing 0.5 $km \cdot h^{-1}$

approximately every 60 s to the point at which the participants could no longer maintain the required running speed. The final distance achieved was recorded after the second failed attempt to meet the start/finish line in the allocated time. The reliability (CV% = 9.9%)[22] and concurrent validity of this test have been reported [3].

Interviews

One-to-one semi-structured interviews were employed using an interview guide but also allowing a considerable degree of flexibility during the interview process to explore new areas that emerged throughout the process. Each interview was recorded using a dictaphone and transcribed *verbatim*. Interviews were used to attempt to "generate data which gives authentic insights into peoples' experiences".[26] Interviews explored the experiences and views of the players and coaches with regards to the physical performance tests. [28]

Statistical analysis

Descriptive statistics for the physical performance tests were presented as the mean ± standard deviation (SD). The Shapiro-Wilk test was used to assess assumptions of normality, with all data meeting this assumption (p = 0.309 to 0.887). Separate paired sample t-tests were used to determine differences ($p < 0.05$) in performance between trials with and without coach presence. Standardised mean differences (SMD) with 95% confidence intervals were also calculated using the difference in the means over the pooled standard deviation. SMD thresholds were: 0.0 – 0.2, trivial; 0.2 - 0.6, small; 0.6 – 1.2, moderate; 1.2 – 2.0, large; >2.0, very large. All statistical analysis was performed using SPSS (IBM SPSS Statistics for Windows, Version 25.0, Armonk, NY, USA).

Thematic Analysis

Consistent with the philosophical assumptions of the study, a realist thematic analysis approach [28] was used to analyse interview data. From this analysis, four themes were identified: 1) The perceived value of physical performance testing; 2) Coaches' use of power to promote hard-work and determination; 3) Players' use of power to achieve career progression; and 4) Players respond differently to observer presence.

Results

Analyses revealed players were -1.2 ± 1.1% (mean ± 95% CI) faster during the 20 m sprint with (3.03 ± 0.08 s) compared to without (3.07 ± 0.10 s) the coaches present (p = 0.035, SMD ± 95% CI: -0.31 ± 0.28). Players' CMJ jump height was 4.7 ± 3.4% higher with (42.0 ± 5.3 cm) compared to without (40.2 ± 6.0 cm) the coaches (p = 0.006, SMD ± 95% CI: 0.27 ± 0.19). Finally, the players' prone Yo-Yo distance increased by 7.2 ± 5.2% (979 ± 223 cf. 1040 ± 186 m; p = 0.015, SMD ± 95% CI: 0.28 ± 0.19) when coaches were present. Mean and individual values are shown in Figure 1.

***** Insert Figure 1 about here *****

Discussion

For the first time, we report the influence of coaches' presence on the outcomes of physical performance testing conducted by a sport scientist on professional team sport athletes. In all tests the athletes' performance was improved when the coaches were present compared to when this was conducted by the sport scientist alone. Further data gathered from interviews build upon these quantitative findings by providing insight into the perceptions of players and coaching staff towards performance testing, alongside highlighting potential explanations for

changes in player performance in the presence of observers. Within this discussion, the four themes that were developed from these data are explored and used to suggest explanations for these changes in performance.

The perceived value of physical performance testing

Primarily, head coaches placed considerable emphasis on these tests and viewed them as an important barometer of a player's ability and felt that they helped to inform "a good recognition of what the strengths and weaknesses are of each individual" (Head Academy Coach). In fact, performance testing was held in such high regard by head coaches that it impacted on team selection. Jones et al.[12] similarly found that coaches used Global Positioning System (GPS) data on high-speed running to ascertain if players were training at an intensity that reflected match play and would not select those who did not. In the present study, the Head of Performance was questioned about whether testing impacted the head coach's team selection and stated:

> I have seen it impact on selection before and heard a lot about it impacting on selection before. Yeah, lads not meeting their markers and it impacts on selection, and at some clubs, it impacts when they have to return from an off season. If they come back and meet certain markers then they will get some time off, if they meet them again, they might get some more time off. So yeah, it impacts on selection but also when they come back.

Ultimately, the use of testing by head coaches to shape team selection might be explained by the belief that this would contribute to a potential competitive advantage. For example, both head coaches suggested that the nature of rugby league required players who were resilient and strong-willed and testing gave them "an indication of where players are mentally" (Academy

Head Coach). Additionally, testing informed training to improve players' physical performance. For example, the Head Coach stated, "we can get better… by implementing certain strength techniques in the gym and certain aspects of our training are designed around testing, you know, to get stronger, to get faster" and that it was important to "implement [the results] and get better and better and better each week, each month, each year continuously". Also, testing informed the tactics employed by head coaches, with the Academy Head Coach stating "it enables me to identify some positional specific stuff. So, for instance if I have got halves and they are coming up on the sprint tests constantly slow, I might alter the way I play the game".

However, for other members of coaching staff, testing was considered less useful, and they placed less value on it. For example, Assistant Coach A felt that performance testing was "probably not really too relevant to my role…I am probably looking at the smaller picture stuff". These views were shared by Assistant Coach B: "I have got more things to worry about, I let the head coach worry about that". These observations corroborate those reported previously that the support around, and selection of, players adopts a multidimensional approach of which physical performance testing is only part of the process [20,21]. Meanwhile, players often stated that they simply did not like testing and viewed it as "one of those things you got to do. It's part and parcel of it [being an academy rugby player]" (Player C). Some players argued that performance testing was not important to them as it did not truly reflect their rugby league ability and, therefore, they placed little value on it. However, some players did feel that the testing informed whether their physical performance had improved. Also, testing was viewed as valuable for players when coaching staff were present. Often, this value came from the opportunity to impress the head coaches and "to stand out from the rest of the team… [to] have the best opportunity to get into the first team" (Player D). Consequently,

players seem to value testing more in the presence of coaching staff primarily due to its perceived use in team selection. According to figurational sociologists, individuals are interdependent and form networks of social relations called 'figurations' [31]. Within these figurations, relations between interdependent individuals shape social culture, meanwhile this social culture simultaneously shapes individuals' socially constructed "personality make-up" (p.59) known as habitus [31]. Therefore, by players valuing testing more in the coaching staff's presence it seems that players' habituses were built around a desire to play the game itself and progress into the first team.

Coaches' use of power to promote hard-work and determination

The role of physical performance tests in rugby league is multifaceted [20]. Whilst testing did seem to inform training, our interview data suggest that these tests were frequently used by coaches to exercise their power over players. This was primarily done by coaches surveilling players during testing to assess their character. For coaches, testing was viewed as an opportunity to identify athletes who were mentally resilient and "who can hang on the longest" (Head Coach). This is summarised by the Head Academy Coach who stated:

> it gives me an indication of where players are mentally… when we do two bouts of testing close together like we did, I can see, I can recognise social loafing and stuff like that and recognise people who want to improve and want to put it in and realise what the testing is about, and other people who may have that attitude that it's not as important as the rugby side.

To this end, coaches appeared to make judgements on whether a player was of sufficient standard to compete for the club based upon this testing, supported by the previous point

highlighting that team selection would be influenced by testing performance. Players who were seen as showing 'poor' character during these tests were subsequently challenged by coaches and viewed as not being at the required playing standard. For example, the Head Academy Coach recalled saying to one of his players:

> Do you think you are ever going to become a professional athlete if you don't do what's required of you from all aspects of training? And that's about all aspects of training we require from you.

Fundamentally, the coaches placed such emphasis on testing as a tool to measure the players' character with the aim of creating a culture within the club that promoted hard-work and determination. This is encapsulated by Assistant Coach A who felt that testing was "for building some mental toughness resilience for the journey". Similarly, Assistant Coach B stated:

> You would want the players to hit as hard as they can with the coach there or not there. I live in the real world; I know that some won't, so what you are trying to do is put the right people there in the club… to create the right culture

Power is constantly in flux and is balanced between interdependent humans within figurations [31]. In this instance, power appeared to be balanced towards coaches due to their ability to shape team selection based upon testing performance. As coaches used testing to assess a player's character, it seems that they tried to use this power to encourage a social culture that promotes hard-work and determination, which they hoped players would be socialised into and internalise. This aligns with previous research, which suggests that surveillance techniques are used by coaches to normalise athlete behaviour and shape athletes into the desired 'type' of person [10–12].

Players' use of power to achieve career progression

Power balances between coaches and players provided an extra incentive for players to work hard and show character, however, such additional effort was performative. Indeed, players' performance improved, on average, for the sprint, jump and intermittent running tests by 1.2, 4.7 and 7.2%, respectively with the coaches present, compared to when tests were conducted by the sport scientist alone. Furthermore, our interview data show that players would put in additional effort in the presence of coaches to display a 'strong' character. For example, Player C stated:

> They [the coaching staff] want to see how mentally tough you are and how much you dig in. Like, when you're bolloxed, they don't want you to just give up, you know, like they want you to keep going and going and going until you can't go anymore.

Meanwhile, Player G felt that "you have to push yourself to try and impress him [the head coach]" and make sure "it *looks* like you are putting the effort in to do it". However, these data, considered alongside the point that many players only saw value in testing when coaches were present, suggest that this additional effort was performative. This highlights how power was also balanced towards the players who seemingly chose whether they gave maximum effort in testing. Additionally, this suggests that players did not internalise a 'hard-working' habitus in respect to testing that the coaching staff desired. Rather, this reinforces the point that players' habituses were built around progressing into the professional game and testing was used as an opportunity to achieve this as they saw fit. Therefore, coaches attempting to exercise their power and use testing to assess the character of players did not promote the hard-working

culture in the manner they had hoped. Instead, this caused players to exercise their own power and merely perform differently in the presence of coaches during such tests, which further highlights how power is in flux within the figuration. Taken together we suggest that when using performance tests to appraise players' physical qualities as part of longitudinal monitoring, practitioners should ensure consistency in the personnel observing the procedures.

Players respond differently to observer presence

Collective test performance with the coaches in attendance were statistically better $p = 0.006$ to 0.035), albeit changes were small (SMD = 0.27 to 0.31) and within the reported trial to trial error for these tests [22]. For example, the improvement in sprint performance was better than the ~0.3% change reported in academy rugby league players after a pre-season training period [32], but smaller than the 5.9% faster times observed in senior players after an 8-week resistance training intervention [33]. Likewise, the observed increases in CMJ and prone Yo-Yo between trials were lower than changes reported in a similar group of academy rugby league players after pre-season training (~10% and ~20%, respectively; [32]). However, when the data are taken on an individual basis, six of the fifteen players reported improvements in prone Yo-Yo performance between 120 and 200 m when the coaches attended. Therefore, 40% of the players made improvements in intermittent running performance that would be deemed beneficial based on required change in running performance reported for this test [22]. Given the strong associations between prone Yo-Yo distance and rugby league match running performance [3], and that a better prone Yo-Yo performance differentiates between playing standards [34], players not providing a true maximal effort in the test means coaches would be misinformed when using this data to inform training or team selection. The fewer number of players that showed beneficial improvements in CMJ (n = 2) and sprint performance (n = 1) when the coach was in attendance perhaps reflects the nature of the tests compared to the prone Yo-Yo test, suggesting

that players are more capable of replicating short duration, all-out tests and that these are less susceptible to the influence of observers. Alternatively, these larger fluctuations in Yo-Yo test performance under the presence of observers might be due to the emphasis participants placed upon this test. Players reported that the Yo-Yo test was the hardest of the battery and that it was viewed as "the big one" (Player A). As such, the players focussed upon this test and directed their energy towards it. For example, Player I stated that "the Yo-Yo test is the hardest because that's like conditioning but the rest of them aren't too bad". Furthermore, players implied that the Yo-Yo test specifically provided an opportunity for coaches to see which players had good character, supporting the idea that such testing is viewed as a character test. For instance, Player A was asked what they felt the coaches gained from being present at testing and responded by saying: "Well, [to see] how people deal with it, right. When, say when the Yo-Yo starts getting hard, yeah, how people react to it". Indeed, coaches seemed to purposefully attend some testing sessions to see how players would react, for example Assistant Coach A stated:

> If the head coach is present, it's kind of like "oh shit" you know what I mean? To give them a bit of .. "oh, the head coach is there we will work a bit harder". So that would be the purpose of why he would be there.

Consequently, this provided an opportunity for players to demonstrate their character and be seen positively by coaches, which in turn appeared to significantly influence the effort players gave during this specific test. This is highlighted by statements made by players such as "I tried to do better because someone like him [the head coach] was there" (Player F) and "I think, as a whole, all of us, like, try harder if you know what I mean? Because they were there watching, so I think that had an impact on us" (Player D). However, for some participants this observer influence had an adverse effect. For example, Player I displayed a reduction in Yo-Yo

performance with observers present and felt that there was an "added pressure". Also, the Head Academy Coach felt that they "noticed more tension really and that could have had a negative effect on some of the testing results". Therefore, it seems that the coaching staff sometimes manipulate their power by observing testing with the intention of encouraging players to try harder and improve their performance. However, through the interdependence of the players and coaches, and the power balances between them, an unintended outcome occurred [31], whereby performance improved for some players but worsened for others. Consequently, the variability of individual performance in the presence of observers during these tests, especially the Yo-Yo, severely undermines their validity. Our findings are the first to report such impact of observer presence on performance testing.

Our study is not without limitations. Using a single club meant the player and coach recruitment was restricted to those employed by the organisation. Thus, the club's and individual coaching philosophies will have influenced perceptions and attitudes of both players and coaches towards performance testing. Restriction to a finite number of players within the club means we acknowledge the potential overestimation of the population effect for the prone YoYo IR1. Albeit players and coaches were well familiarised with the testing, we were also unable to conduct the study using a randomised crossover design to limit any order effects. Future studies using larger sample sizes from a range of clubs, with a more truly experimental design are warranted.

Conclusions

Our findings show that athlete performance during a standardised testing battery for the assessment of physical qualities generally improved when coaches were present. Performance testing was mainly used by coaches to exercise their power-chances over players in the hope

of generating a social culture that promoted hard-work and dedication, which players would be socialised into and adopt. Performance testing was used to assess the 'character' of the players and decide if they were of sufficient standard to be at the club, which consequently informed team selection. However, athletes exercised their own power by providing additional effort in the presence of coaches exclusively. This reinforces the argument that their habitus was centred around being selected for the team and progressing within the sport rather than around persistent hard-work that the coaches tried to promote. The interdependence of coaches and players, and the power balances between them, led to the unintended consequence of some athletes' performance worsening during this test. Ultimately, our findings challenge the validity of such tests under the presence of coaches and encourage practitioners to strive for consistency in observing staff members when conducting these procedures as part of longitudinal monitoring.

Practical implications

Primarily, the findings of our study can help to inform coaches and practitioners of the challenges of implementing performance testing appropriately. Whilst our findings suggest that the presence of coach observers undermines the validity of various performance tests, this is not to say that performance tests have no value in applied contexts. Instead, practitioners should consider the presence of coaches during performance testing and be wary of the impact this might have on the results of such tests. By doing so, the application of performance testing might be improved and better promote effective athlete development. Ideally, there should be consistency with the presence of observers during testing procedures to minimise this influence, however we are aware that this is often beyond the control of the practitioner and

might not be feasible. Accordingly, further research that investigates the power relations between practitioners and coaching staff, specifically related to the context of physical performance testing, might provide insight into the complexities of ensuring such consistency of observers. Finally, future research may benefit from investigating whether these findings are applicable to senior athletes within the professional game, where the pressures and incentives for players might differ to those of younger players.

Acknowledgments: The authors would like to thank the players and staff who willingly gave up their time to participate in this study.

References

1. Krustrup P, Mohr M, Amstrup T, et al. The Yo-Yo intermittent recovery test: physiological response, reliability, and validity. *Med Sci Sports Exerc* 2003; 35: 697–705.
2. Hogarth LW, Burkett BJ, McKean MR. Influence of Yo-Yo IR2 scores on internal and external workloads and fatigue responses of tag football players during tournament competition. *PLoS One*; 2015; 10: DOI: 10.1371/journal.pone.0140547.
3. Dobbin N, Highton J, Moss SL, et al. Concurrent validity of a rugby-specific yo-yo intermittent recovery test (level 1) for assessing match-related running performance. *J strength Cond Res* 2021; 35: 176–182.
4. Mohr M, Krustrup P. Yo-Yo intermittent recovery test performances within an entire football league during a full season. *J Sports Sci* 2014; 32: 315–327.
5. Waldron M, Worsfold PR, Twist C, et al. A three-season comparison of match performances among selected and unselected elite youth rugby league players. *J Sports Sci* 2014; 32: 1110–1119.

6. Till K, Jones B, Darrall-Jones J, et al. Longitudinal development of anthropometric and physical characteristics within academy rugby league players. *J Strength Cond Res* 2015; 29: 1713–1722.

7. Dobbin N, Highton J, Moss SL, et al. The effects of in-season, low-volume sprint interval training with and without sport-specific actions on the physical characteristics of elite academy rugby league players. *Int J Sports Physiol Perform* 2020; 15: 705–713.

8. Morgan PJ, Callister R. Effects of a preseason intervention on anthropometric characteristics of semiprofessional rugby league players. *J Strength Cond Res* 2011; 25: 432–440.

9. Gabbett TJ, Ullah S, Finch CF. Identifying risk factors for contact injury in professional rugby league players - Application of a frailty model for recurrent injury. *J Sci Med Sport* 2012; 15: 496–504.

10. Manley A, Palmer C, Roderick M. Disciplinary power, the oligopticon and rhizomatic surveillance in elite sports academies. *Surveill Soc* 2012; 10: 303–319.

11. Manley A, Roderick M, Parker A. Disciplinary mechanisms and the discourse of identity: The creation of 'silence' in an elite sports academy. *Cult Organ* 2016; 22: 221–244.

12. Jones L, Marshall P, Denison J. Health and well-being implications surrounding the use of wearable GPS devices in professional rugby league: A Foucauldian disciplinary analysis of the normalised use of a common surveillance aid. *Perform Enhanc Heal* 2016; 5: 38–46.

13. Roderick M. A very precarious profession:: Uncertainty in the working lives of professional footballers. *Work Employ Soc* 2006; 20: 245–265.

14. Taylor WG, Potrac P, Nelson LJ, et al. An elite hockey player's experiences of video-

based coaching: A poststructuralist reading. *Int Rev Sociol Sport* 2017; 52: 112–125.

15. Coutts AJ, Murphy AJ, Dascombe BJ. Effect of direct supervision of a strength coach on measures of muscular strength and power in young rugby league players. *J Strength Cond Res* 2004; 18: 316–323.

16. Mazzetti SA, Kraemer WJ, Volek JS, et al. The influence of direct supervision of resistance training on strength performance. *Med Sci Sports Exerc* 2000; 32: 1175–1184.

17. Minett GM, Fels-Camilleri V, Bon JJ, et al. Peer Presence increases session ratings of perceived exertion. *Int J Sports Physiol Perform*; 17. Epub ahead of print 2022. DOI: 10.1123/ijspp.2021-0080.

18. Kempton T, Sirotic AC, Rampinini E, et al. Metabolic power demands of rugby league match play. *Int J Sports Physiol Perform* 2015; 10: 23–28.

19. Winchester R, Turner LA, Thomas K, et al. Observer effects on the rating of perceived exertion and affect during exercise in recreationally active males. *Percept Mot Skills* 2012; 115: 213–227.

20. McCormack S, Jones B, Scantlebury S, et al. "It's important, but it's not everything": practitioners' use, analysis and perceptions of fitness testing in academy rugby league. *Sports* 2020; 8: 1–18.

21. McCormack S, Jones B, Elliott D, et al. Coaches' assessment of players physical performance: subjective and objective measures are needed when profiling players. *Eur J Sport Sci* 2021; 1–11.

22. Dobbin N, Hunwicks R, Highton J, et al. A reliable testing battery for assessing physical qualities of elite academy rugby league players. *J Strength Cond Res* 2018; 32: 3232–3238.

23. Faul F, Erdfelder E, Lang A-G, et al. G*Power 3. *Heinrich-Heine University - Institute*

for Experimental Psychology.

24. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. *Qual Rep* 2015; 20: 1408–1416.
25. Mason M. Sample size and saturation in PhD studies using qualitative interviews. *Forum Qual Sozialforsch*; 11.
26. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health* 2021; 13: 201–216.
27. Rehman AA, Alharthi K. An introduction to research paradigms. International. *J Educ Investig* 2016; 3: 51–59.
28. Wiltshire G, Ronkainen N. A realist approach to thematic analysis: making sense of qualitative data through experiential, inferential and dispositional themes. *J Crit Realis* 2021; 20: 159–180.
29. Ronkainen NJ, Wiltshire G. Rethinking validity in qualitative sport and exercise psychology research: a realist perspective. *Int J Sport Exerc Psychol* 2021; 19: 13–28.
30. Dobbin N, Hunwicks R, Highton J, et al. Validity of a jump mat for assessing countermovement jump performance in elite rugby players. *Int J Sports Med* 2017; 38: 99–104.
31. van Krieken R. *Norbert Elias*. 2005; Routledge.
32. Dobbin N, Gardner A, Daniels M, et al. The influence of` preseason training phase and training load on body composition and its relationship with physical qualities in professional junior rugby league players. *J Sports Sci* 2018; 36: 2778–2786.
33. Comfort P, Haigh A, Matthews MJ. Are changes in maximal squat strength during preseason training reflected in changes in sprint performance in rugby league players? *J Strength Cond Res* 2012; 26: 772–776.

34. Dobbin N, Highton, Moss L, et al. The discriminant validity of a standardized testing battery and its ability to differentiate anthropometric and physical characteristics between youth, academy, and senior professional rugby league players. *Int J Sports Physiol Perform* 2019; 14: 1110–1116.

Figure legends

Figure 1. Individual changes in 20 m sprint, CMJ and prone Yo-Yo performance trials without and with the coaches present. * denotes difference between trials ($p < 0.05$).

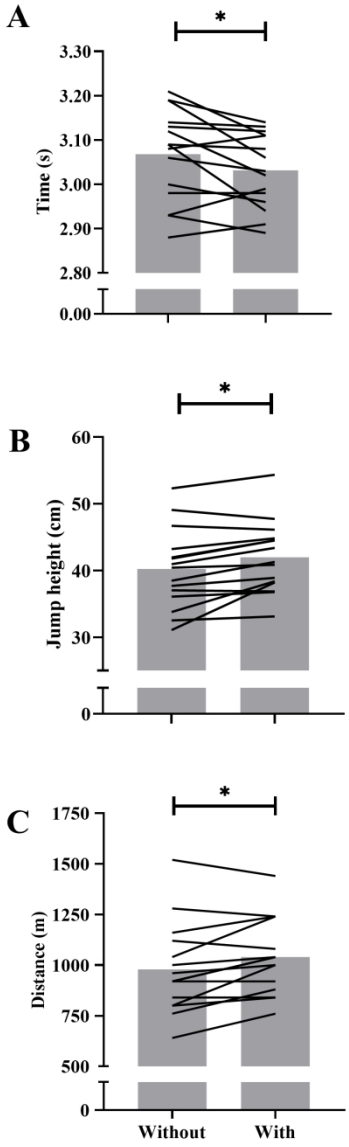

Figure 1. Individual changes in 20 m sprint, CMJ and prone Yo-Yo performance trials without and with the coaches present. * denotes difference between trials ($p < 0.05$).

90x252mm (300 x 300 DPI)