A comparison of game-play characteristics between elite youth and senior Australian National Rugby League competitions

Woods T. Carl¹*, Robertson Sam², Sinclair H. Wade¹, Till Kevin³, Pearce Leesa¹, Leicht S. Anthony¹

¹Discipline of Sport and Exercise Science, James Cook University, Townsville, Queensland, Australia
²Institute of Sport, Exercise & Active Living (ISEAL), Victoria University, Melbourne, Australia
³Institute for Sport, Physical Activity and Leisure, Leeds Beckett University, Leeds, West Yorkshire, United Kingdom

*Corresponding Author

Carl Woods, Discipline of Sport and Exercise Science, James Cook University, Townsville, Queensland, Australia.

Ph: +61 08 4781 6550 Mob: +61 421254329 Email: carl.woods@jcu.edu.au
Abstract

Objectives: To compare game-play characteristics between elite youth and senior Australian National Rugby League (NRL) competitions.

Design: Longitudinal observational.

Methods: The dataset consisted of 12 team performance indicators (e.g., ‘all runs’, ‘offloads’ and ‘tackles’) extracted from all 2016 national under 20 (U20) competition (elite youth; n = 372 observations) and National Rugby League (NRL) (elite senior; n = 378 observations) matches. Data was classified according to competition (Two levels: U20 and NRL) and modelled using two techniques. Firstly, non-metric multidimensional scaling resolved multivariate competition (dis)similarity, visualised using a two-dimensional ordination. Secondly, a conditional interference (CI) classification tree was grown to reveal the performance indicators most capable of explaining competition level.

Results: Non-metric multidimensional scaling revealed high competition dissimilarity, with U20 and NRL teams orienting distinctive positions on the first dimension of the ordination surface. Five team performance indicators were retained within the CI tree (‘all runs’, ‘tackle breaks’, ‘tackles’, ‘missed tackles’, and ‘kicks’), which correctly classified 79% of the U20 observations and 93% of the NRL observations.

Conclusions: Multivariate differences between elite youth and senior rugby league competitions were identified. Specifically, NRL game-play was classified by a greater number of ‘all runs’, and ‘tackles’ and a lower number of ‘missed tackles’ relative to the U20 competition. Given the national U20 competition is purported to assist with the development of prospective NRL players, junior coaches may consider training interventions that primarily aid the tackling capacities of players. This may subsequently assist with talent development and player progression in Australian rugby league.

Key words: Performance analysis; talent development; classification tree; multidimensional scaling
1. Introduction

Developing talent is a complex and typically non-linear process,\(^1\) influenced by a range of intrapersonal, environmental and situational catalysts.\(^2\) In an attempt to positively augment this process, it is common for national sporting organisations to establish talent development academies or ‘pathways’, intended to offer longitudinal player development opportunities for talent identified juniors.\(^3,4\) The unifying goal of these development pathways is often to bridge performance discrepancies between junior and senior competitions, creating an expedited developmental transition for participants towards the elite senior level.\(^5\)

Talent development in team sports, such as rugby league, is often further complicated given their innate multidimensionality (i.e., physical, technical and perceptual requisites).\(^6-9\) Accordingly, it is unsurprising to note the quantity of work that has investigated a range of performance qualities discriminant of developmental level specific to rugby league.\(^8,10-12\) Till et al.\(^8\) identified the anthropometric and physical fitness characteristics of English academy rugby league players in relation to their career attainment (professional or academy). Their results demonstrated differences (lower body strength) and changes (10 m momentum) in the physical characteristics of players as they progressed through the academy according to career attainment level, with these observations being of use for the establishment of training interventions assistive with talent development.\(^8\)

In Australia, the premier youth rugby league competition is the national under 20 (U20) competition; currently referred to as the Holden Cup. Fundamentally, the premise of this competition is to provide talent identified youths with a pathway into the elite senior competition, the National Rugby League (NRL). Accordingly, each of the 16 NRL teams has a representative U20 team who competes within a 26-week competition. Although statistics regarding player progression from the U20 to the NRL competition are not available, it is widely known within the NRL community that this national youth competition offers a critical developmental environment for prospective NRL players.
However, despite this, the game-play characteristics (e.g., ball carries and tackles) between competitions has yet to be compared. Resolving this competition difference would likely provide coaches at the U20 level with an objective basis to minimise performance gaps between the U20 and NRL competitions and aid player progression and development within Australian rugby league.

Given the multidimensional dynamicity of rugby league match play,\textsuperscript{13} singular linear statistical approaches may not adequately reveal multivariate patterns within a dataset,\textsuperscript{14} constraining the practical utility of the observations.\textsuperscript{15} Given this, recent research in rugby league has adopted machine learning approaches to assist with the explanation of match outcomes using a diverse range of performance indicators.\textsuperscript{16} As such, it would seem appropriate to consider machine learning approaches when examining differences between elite junior and senior rugby league competitions to assist with the resolution of non-linear interactions between diverse datasets.

The aim of this study was to compare game-play characteristics between elite youth (national U20) and senior (NRL) Australian rugby league competitions. The subsequent results of this work will offer coaches at the U20 level with objective guidance with respect to the establishment of targeted training interventions intended to expedite talent development and player progression in rugby league.

2. Methods

Team performance indicators from the 2016 Holden Cup and NRL seasons were extracted from a publicly available source (http://www.nrl.com/stats) (Supplementary Table 1). These 12 indicators were chosen owing to their relevance in the explanation of rugby league match outcomes,\textsuperscript{17} as well as their availability at the time of analysis. Both competitions consisted of 26-rounds, equating to 372 observations in the U20 competition (n = 186 games) and 378 observations in the NRL (n = 189 games). The observational differences were due to the number of match-free rounds ('byes') integrated within both competitions. The absolute game-times were the same across both
competitions. The relevant Human Ethics Committee provided ethical approval prior to data acquisition.

All of the following analyses were performed using the computing environment R, version 3.2.5.\(^{19}\) Prior to modelling, data was classified according to competition (Two levels: U20 and NRL). A correlation matrix was built to assess the level of collinearity between each team performance indicator. Descriptive statistics (mean ± standard deviation) were calculated for each team performance indicator relative to developmental level. A multivariate analysis of variance (MANOVA) was then used to test the main effect of competition (Two levels: U20 and NRL) on each team performance indicator, with the Type-I error being set at \(P < 0.05\). Additionally, the effect size and subsequent 90% confidence interval of competition was calculated using Cohen’s \(d\) statistic\(^{20}\) in the ‘MBESS’ package,\(^{21}\) with interpretations being in accordance with established recommendations.\(^{22}\)

To reveal the level of competition dissimilarity, non-metric multidimensional scaling (nMDS) was used. This is an indirect gradient analysis that produces an ordination based on a dissimilarity matrix.\(^{23}\) This matrix is resolved via isotopic regression, which is a non-parametric form of regression that iteratively searches for a least squares fit based on the ranked dissimilarities.\(^{23}\) It is the preferred ordination technique when no assumptions are made about the underlying distribution of the data.\(^{23}\) Aggregates of each team performance indicator were used across the season for this analysis. Further, teams across both competitions were sorted according to their ladder position at the conclusion of the season. This enabled insights into the dissimilarity of dominant and less dominant teams across both competition levels. The team performance indicators were used to build a dissimilarity matrix using the Bray-Curtis measure in the ‘vegan’ package via the \textit{metaMDS} function.\(^{24}\) This dissimilarity matrix was then plotted in two-dimensions using generalised additive models employing an isotopic smoother via thin-plate regression splines.\(^{23}\) On the ordination surface, competition was colour coded and teams were labelled, while their subsequent ladder
position was denoted via the size of their ‘point’ using the `geom_label` and `geom_point` functions in the ‘ggplot2’ package.25

To determine the combination of performance indicators that provided the greatest competition classification, a conditional interference (CI) classification tree was grown in the ‘party’ package.26 This type of classification tree was chosen as it estimates a regressive relationship through binary partitioning by testing the null hypothesis between a set of explanatory variables (team performance indicators) and a binary response variable (competition).26 Partitioning ceases when the null hypothesis cannot be rejected (P ≥0.05). A benefit of this analysis is that its fitting algorithm corrects for multiple testing, avoiding overfitting, resulting in the growth of an unbiased classification tree that does not require pruning.26

3. Results
Collinearity was noted between ‘all runs’ and ‘all run metres’ (r >0.5), with the latter being removed from further analysis given its dependence on the ‘all run’ indicator. The descriptive statistics and between group effects are presented in Table 1. A significant effect of competition was noted (V = 0.936, F = 23.194, P <0.05), with ten of the 11 modelled indicators yielding significance (Table 1). Further, ‘all runs’ expressed the largest effect relative to competition (d = 1.22, Table 1).

The dissimilarity matrix solution was reached after 20 iterations (stress = 0.09, rmse = 3.1 x 10^{-4}, maximum residual = 1.4 x 10^{-3}). The subsequent two-dimensional ordination showed distinct competition dissimilarity along dimension one (MDS1, Figure 1). Notably, the profiles of the NRL teams oriented on the right panelling of the first dimension, while the profiles of the U20 teams oriented on the left panelling (Figure 1). Despite this multivariate competition dissimilarity, it was of interest to note the similar positioning of dominant and less dominant teams across both competitions along the second dimension (MDS2). Notably, the higher ranked teams (i.e., closer to

****INSERT TABLE ONE ABOUT HERE****
‘1’ on the ladder, denoted by a smaller ‘point’) oriented the top half of MDS2 on the ordination surface, while the lower ranked teams (i.e., closer to ‘16’ on the ladder, denoted by a larger ‘point’) generally oriented the lower half of MDS2 on the ordination surface (Figure 1).

***INSERT FIGURE ONE ABOUT HERE***

The CI classification tree successfully classified 293 of the 372 U20 observations (79%) and 350 of the 378 NRL observations (93%). Of the 11 team performance indicators modelled, five were retained within the full tree (Figure 2); these being ‘all runs’ (root node 1), ‘tackle breaks’, ‘tackles’, ‘missed tackles’, and ‘kicks’. Accordingly, 14 terminal nodes were grown from these performance indicator combinations (Figure 2).

***INSERT FIGURE TWO ABOUT HERE***

Progressing to the right of the root node (>167 ‘all runs’), leaf node 19 partitions the data based on ‘tackle breaks’ at a count of 36 (Figure 2). Progressing to leaf node 20 (≤36 ‘tackle breaks’), the data was partitioned based on ‘tackles’ at a count of 288. Of the 144 observations within terminal node 22 (>167 ‘all runs’, ≤36 ‘tackle breaks’ and >288 ‘tackles’), 98.6% were classified as NRL observations, while only 1.4% were classified as U20 observations. This combination provided the most accurate classification of NRL observations and lowest classification of U20 observations within each terminal node on the right of the tree.

Progressing to the left of the root node (≤167 ‘all runs’), leaf node 2 partitioned the data based on ‘tackles’ at a count of 314. Progressing to leaf node 3 (≤314 ‘tackles’), the data was partitioned based on ‘tackle breaks’ at a count of 30. Of the 177 observations in terminal node 7 (≤167 ‘all runs’, ≤314 ‘tackles’ and >30 ‘tackle breaks’), 95.5% were classified as U20 observations, while only 4.5% were NRL games. Although terminal nodes 15 and 18 resolved 100% classification accuracy for U20 observations, it is important to note that these nodes consisted of a small number of observations (12 and 11, respectively).
4. Discussion

The aim of this study was to compare game-play characteristics between the U20 and NRL competitions within Australia. As shown in Figure 1, there was a clear dissimilarity in the multivariate (performance indicator) profiles of U20 and NRL teams. Given this competition dissimilarity, it was expected to note the high classification accuracy shown by the CI classification tree. A unique combination of ‘all runs’, ‘tackle breaks’ and ‘tackles’ provided the greatest classification of both the NRL and U20 observations (Figure 2). These results demonstrated that there were clear differences in game-play behaviours between competitions, primarily manifested via ‘all runs’, ‘tackle breaks’ and ‘tackles’. Coaches at the U20 level could utilise these results to base training interventions to address the apparent performance gaps in talent development for the NRL.

The nMDS analysis revealed apparent dissimilarity between the two competitions, indicating a difference in multivariate team profiles. These multivariate differences may have arisen from a range of factors, such as coaching strategies and team tactics implemented across competitions, or the functional capacities of players across both competitions. Given the U20 competition offers a feeder of talent for the NRL, it would seemingly be beneficial for coaches at this level to implement tactics that attempt to enable game-play characteristics that express similarity to the NRL (e.g., minimising missed tackles and generating greater ball carries), assisting with player development. However, the dissimilar positioning of U20 and NRL teams from the same club suggests distinctive profiles generated by teams from the same club. Nonetheless, it is possible that these multivariate differences stemmed from functional differences between the players in both competitions. Specifically, given their current stage of development, U20 players may be unable to perform the same types of actions as frequently as their NRL counterparts.

The apparent similarity between higher and lower ranked U20 and NRL teams along MDS2 on Figure 1 was of note. Higher ranked teams (closer to ‘1’ on the ladder) in both the U20 and NRL competitions generally clustered along the top half of MDS2, while lower ranked teams (closer to
'16’ on the ladder) generally clustered along the lower half of MDS2. Although speculative, this may indicate that superior teams across both competitions produce unique multivariate profiles relative to their respective competitions. This is supported by the recent findings of Woods et al.\textsuperscript{13} who noted that superior NRL teams (defined via grand final representation) generated highly dissimilar profiles relative to their less superior counterparts. It is, however, important to note that this data signature was more prominent within the NRL as opposed to the U20 competition (noted in the positioning of the West Tigers and the Roosters on Figure 1). Subsequently, it is possible that the focus of the U20 competition is not only to win games, but also to promote player development, which may not always be synonymous with match success.

The results of the CI classification tree are likely to yield practical utility in the development of prospective NRL players. This analysis indicated that NRL games were classified by a greater number of ‘all runs’, ‘tackles’ and a lower number of ‘tackle breaks’ relative to the U20 level. Accordingly, U20 players entering the NRL may not have been exposed to the type of physicality (e.g., tackling capacity) needed to compete within this elite senior competition.\textsuperscript{27} This suggestion is somewhat supported by the work of Ireton et al.\textsuperscript{12} who examined the differences in physical fitness and athletic movement between elite senior and youth rugby league players. Their results demonstrated differences between elite senior and youth rugby league players for measures of body mass, lower body power and athletic movement.\textsuperscript{12} Although a multitude of factors could have implicated the performance indicator characteristics resolved here, when the work of Ireton et al.\textsuperscript{12} is coupled with the current study, the competition differences could be, in part, due to lower body power and movement quality discrepancies. However, in addition to physical precocity being an important element in determining tackling efficiency,\textsuperscript{28} the differences noted in our study may also be underpinned by differences in tackling technique.\textsuperscript{29} Accordingly, it is recommended that coaches at the U20 level integrate a holistic approach to the development of tackling derivatives that consists of both physical and technical elements.\textsuperscript{28,29}
Additionally, NRL coaches and talent recruitment managers could consider ‘bridging’ an U20 player’s transition into the NRL by encouraging their participation in the State League, which is a state-based senior competition. Given that the State League is classified as a senior competition, it is possible that it offers an environment slightly more conducive to the physicality needed within the NRL. However, given the paucity of data supporting this recommendation, continued work is required to compare game-play characteristics between the State League competition and the U20 and NRL competitions.

Despite the practical utility of this work, it is not without limitations. Firstly, it did not investigate the physical activity profiles of competition levels. Future work may integrate global positioning system (GPS) analysis with these results to gain a deeper insight into the differences between the U20 and NRL levels. Despite being the premier youth rugby league competition in Australia, U20 coaches are likely to have limited access to their players, as they are not full time athletes like their NRL counterparts. Thus, it is likely that NRL players have a greater potential to develop their performance qualities relative to their U20 counterparts, potentially exacerbating the observed performance differences. Notably, a greater number of technical errors in the U20 competition (potentially underpinned by limited training time in contrast to the NRL) may result in less runs, which slows the speed of play. Nonetheless, these preliminary results may offer coaches with the capability to develop targeted training interventions in time-constrained environments.

5. Conclusion

This work demonstrates differences in game-play characteristics between the U20 and NRL competitions, manifested via certain team performance indicators. The multivariate profiles of teams from both competitions was distinctly dissimilar. Additionally, NRL game-play was characterised by a greater number of ‘all runs’, ‘tackles’ and a lower count of ‘tackle breaks’ in comparison to game-play at the U20 level.

6. Practical Applications
Coaches at the U20 level should look to integrate training interventions that enable players to engage in a high number of ‘all runs’ and ‘tackles’, while equipping them with the defensive capabilities to minimise the number of ‘missed tackles’.

Alternatively, NRL coaches and talent recruitment managers may consider encouraging U20 players to ‘bridge’ the gap between the U20 and NRL levels by participating in State League games. Given the State League is a senior competition, it is possible that it will afford an environment more conducive to that seen within the NRL.

Performance analysts working with rugby league teams may consider the analytical techniques used here to examine performance differences across developmental levels or competitions given their visualisation capacities and consideration of non-linear phenomena within datasets.

References


Figure 1. The ordination plot using nMDS of a dissimilarity matrix calculated from team performance indicators of each U20 and NRL team in the 2016 season.

Figure 2. The CI classification tree showing the classification of U20 and NRL observations.

Note: ‘n’ represents the number of observations in each node (minimum of five observations was set). The first ‘y value’ denotes the classification of U20 observations and the second denotes that classification of NRL observations (e.g., 0.8 = 80%). ‘AR’ all runs; ‘TB’ tackle breaks; ‘TK’ tackles; ‘MT’ missed tackles; ‘K’ kicks. Certain performance indicators appear twice in this tree as the recursive partitioning performed by the fitting algorithm searches for classification rules between each variable using a non-linear approach.
**Table 1.** Descriptive and between group effects relative to developmental level

<table>
<thead>
<tr>
<th>Performance indicator</th>
<th>NRL</th>
<th>U20</th>
<th>$d$ (90% CI)</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>All runs*</td>
<td>170.2 ± 19.8</td>
<td>147.2 ± 17.4</td>
<td>1.22 (1.09 – 1.35)</td>
<td>‘large’</td>
</tr>
<tr>
<td>Line breaks*</td>
<td>4.0 ± 2.3</td>
<td>5.6 ± 2.5</td>
<td>0.64 (0.51 – 0.76)</td>
<td>‘moderate’</td>
</tr>
<tr>
<td>Try assists*</td>
<td>2.8 ± 1.8</td>
<td>3.4 ± 1.9</td>
<td>0.34 (0.22 – 0.46)</td>
<td>‘small’</td>
</tr>
<tr>
<td>Offloads*</td>
<td>10.3 ± 4.4</td>
<td>7.0 ± 3.3</td>
<td>0.82 (0.69 – 0.95)</td>
<td>‘moderate’</td>
</tr>
<tr>
<td>Tackles*</td>
<td>325.0 ± 39.7</td>
<td>283.4 ± 35.6</td>
<td>1.10 (0.97 – 1.22)</td>
<td>‘moderate’</td>
</tr>
<tr>
<td>Missed tackles*</td>
<td>27.6 ± 8.4</td>
<td>35.7 ± 10.9</td>
<td>0.82 (0.70 – 0.95)</td>
<td>‘moderate’</td>
</tr>
<tr>
<td>Errors*</td>
<td>9.2 ± 2.7</td>
<td>10.4 ± 3.2</td>
<td>0.39 (0.27 – 0.51)</td>
<td>‘small’</td>
</tr>
<tr>
<td>Total kicks*</td>
<td>19.0 ± 3.8</td>
<td>14.7 ± 3.6</td>
<td>1.13 (1.00 – 1.26)</td>
<td>‘moderate’</td>
</tr>
<tr>
<td>Line break assists*</td>
<td>3.0 ± 2.0</td>
<td>3.6 ± 2.0</td>
<td>0.30 (0.18 – 0.42)</td>
<td>‘small’</td>
</tr>
<tr>
<td>Dummy half runs</td>
<td>11.1 ± 4.5</td>
<td>10.3 ± 4.6</td>
<td>0.15 (0.03 – 0.27)</td>
<td>‘trivial’</td>
</tr>
<tr>
<td>Tackle breaks*</td>
<td>27.6 ± 8.4</td>
<td>35.7 ± 10.9</td>
<td>0.83 (0.70 – 0.95)</td>
<td>‘moderate’</td>
</tr>
</tbody>
</table>

* denotes $P<0.05$. 
### Supplementary Table 1. Team performance indicator descriptions

<table>
<thead>
<tr>
<th>Team performance indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>All runs</td>
<td>The total number of times a player runs while in possession of the ball. A higher number of ‘runs’ is likely to indicate an efficient offence.</td>
</tr>
<tr>
<td>All run metres</td>
<td>The total number of metres accumulated following each run a player attempts while in possession of the ball. A greater accumulation of ‘run metres’ is likely to indicate an efficient offence.</td>
</tr>
<tr>
<td>Line breaks</td>
<td>The number of times a player was able to run through (or ‘break’) an opposition’s defensive line while in possession of the ball. A greater number is likely to indicate an efficient offence.</td>
</tr>
<tr>
<td>Try assists</td>
<td>A pass (via hand or foot) to a teammate who scores a ‘try’</td>
</tr>
<tr>
<td>Offloads</td>
<td>Passing or ‘offloading’ the ball to a teammate whilst being tackled</td>
</tr>
<tr>
<td>Tackles</td>
<td>The use of physical contact to prevent or hold up an opposition player in possession of the ball</td>
</tr>
<tr>
<td>Missed tackles</td>
<td>The number of times an attacking player penetrates a team’s defensive line without being tackled or incurring physical contact</td>
</tr>
<tr>
<td>Errors</td>
<td>Instances during game-play in which a team loses possession through a dropped ball, high hit, or knock-on</td>
</tr>
<tr>
<td>Total kicks</td>
<td>Disposing of the ball with any part of the leg below the knee including kicks off the ground</td>
</tr>
<tr>
<td>Line break assists</td>
<td>The number of times an attacking player passed the ball to a teammate who then was able to run through (or ‘break’) an opposition’s defensive line either with or without opposition contact</td>
</tr>
<tr>
<td>Dummy half runs</td>
<td>The number of times the acting half collects the ball and runs before passing</td>
</tr>
<tr>
<td>Tackle breaks</td>
<td>The action of breaking a team’s defence despite incurring physical contact from the opposition defensive</td>
</tr>
</tbody>
</table>
Acknowledgements

The authors would like to acknowledge the many analysts who recorded and coded the publicly available data over the course of the 2016 Holden Cup and NRL seasons.