Examining the evolution and classification of player position using performance indicators in the National Rugby League during the 2015–2019 seasons

Examining the evolution and classification of player position using performance indicators in the National Rugby League during the 2015-2019 seasons

C. Wedding[a]*, C. T. Woods[b], W. H. Sinclair[a], M. A. Gomez[c] and A. S. Leicht[a]

[a]*Sport and Exercise Science, James Cook University, Douglas, Queensland, Australia.*

[b]*Institute for Health and Sport, Victoria University, Melbourne, Australia.*

[c] *Faculty of Physical Activity and Sport Sciences –INEF, Polytechnic University of Madrid, Madrid, Spain*

*Corresponding Author

Corey Wedding, Sport and Exercise Science, James Cook University, Townsville, Queensland, Australia

Mob: +61 402 446 190 Email: corey.wedding@my.jcu.edu.au

Abstract

Objectives: This study aimed to: 1) examine recent seasonal changes in performance indicators for different National Rugby League (NRL) playing positions; and 2) determine the accuracy of performance indicators to classify and discriminate positional groups in the NRL.

Methods: 48 performance indicators (e.g. passes, tackles) from all NRL games during the 2015-2019 seasons were collated for each player´s match-related performance. The following analyses were conducted with all data: (i) one-way ANOVA to identify seasonal changes in performance indicators; (ii) principal component analysis (PCA) to group performance indicators into factors; (iii) two-step cluster analysis to classify playing positions using the identified factors; and (iv) discriminant analysis to discriminate the identified playing positions.

Results: ANOVA showed significant differences in performance indicators across seasons ($F = 2.3–687.7$; $p = 0-0.05$; partial $\eta2 = 0.00–0.075$). PCA pooled all performance indicators and identified 14 factors that were included in the two-step cluster analysis (average silhouette = 0.5) that identified six positional groups: forwards, 26.7%, adjustables, 17.2%, interchange, 23.2%, backs, 20.9%, interchange forwards, 5.5% and utility backs, 6.5%. Lastly, discriminant analysis revealed five discriminant functions that differentiated playing positions.

Conclusions: Results indicated that player's performance demands across different playing positions did significantly change over recent seasons (2015-2019). Cluster analysis yielded a high-level of accuracy relative to playing position, identifying six clusters that best discriminated positional groups. Unsupervised analytical approaches may provide sports scientists and coaches with meaningful tools to evaluate player performance and future positional suitability in RL.

Keywords: Team sports; sport analytics; classification; data visualisation; performance analysis

Practical implications

- PCA is a useful model to associate and group performance indicators into factors that may explain RL player´s performances.
- Clustering techniques (e.g. two-step cluster) using unsupervised approaches allow analysts to classify player´s performance into different profiles that account for related performance indicators and roles during competition.
- The identification of specific playing positions and the discrimination among them via performance factors may enable establishment of player's performance profiles, critiquing of player´s performances over seasons and identify player´s recruitment potential and suitability.
- Further application of the results of this study could assist sports practitioners in providing greater decisional support with the design and implementation of various training and game-play strategies

Introduction

Rugby league (RL) is a demanding team invasion sport, requiring players to possess a range of physical[1,2] and technical[3-5] qualities. Specifically, National Rugby League (NRL) teams have to perform at the highest level during a very competitive tournament that requires the integration of performance analysis with the intention of describing and identifying teams and player´s performances.[5-7] The integration of these processes could continue to yield a variety of benefits for high performance staff within the NRL, such as understanding the current team performance trends among the league.[6] This may assist with coaching strategies specifically related to game planning and subsequent player selection. Similarly, the ability to understand current positional performance trends could provide a team with advantages during their player recruitment process, such that they can identify and appropriately assess the value of potential player acquisitions – an avenue that is yet to be explored within the NRL.

Previous work in RL has identified performance indicators capable of differentiating playing position (backs, forwards, fullback, hooker, and service players).[4] It was observed that forwards, hookers and service players (halfbacks and five-eight players) completed more tackles per minute than both backs and fullbacks.[4] When each of the groups was compared for offensive involvements, hookers had the highest count of ball touches, whereas both backs and fullbacks completed more runs with the ball than all other positional groups.[4] A similar study also compared the total number of offensive and defensive actions performed by three different positional groups (forwards, backs, and adjustables) amongst junior RL athletes.[3] In this study, forwards (props, lock, and back rowers) completed the greatest number of both offensive and defensive actions compared to both adjustables and backs.[3] Further, adjustables (halves, hooker, fullback) completed a significantly greater number of defensive and total technical skills compared to the backs.[3] Collectively, these studies demonstrated that player's game involvements were likely to vary according to playing position. The implications of this are likely to extend towards practice design, enabling a level of positional representativeness. However, despite these initial findings, it remains unknown whether positional specific attributes in the NRL have evolved over time.

Several studies have identified player's performance from a medium-term perspective in soccer[8] and Australian football (AF).[9,10] In soccer for example, compared to attackers and wide players, central players increased their involvement in play through a greater increase in the number of passes made and pass success rate.[8] More specifically, centre midfielders and fullbacks increased the number of short and medium distance passes from the 2006-07 to 2012-13 season.[8] Furthermore, despite large player homogeneity across various positional demands in junior AF,[9] when combined with physical performance measures, clearer associations between higher and lower ranked draftees were identified.[10] Understanding that the demands of sport may change over time,[6, 11] and having systems in place to monitor and adapt to these changes, is crucial to ensure that contemporary training and game strategies are implemented to enhance a team's chances of success.[12]

Due to the large number of performance indicators available to NRL teams, it is important to understand which of these are explanatory of a successful performance. Performance modelling involving analytical approaches such as factor reduction, clustering and discriminant analysis have previously been used to differentiate playing positions and the importance of various performance indicators in multiple sports.[13-15] These approaches enable the closer inspection of the relationships that exist between both performance variables and positional groups.[13,14] Pertinently, such analytical approaches are capable of resolving clusters of attributes that explain specific aspects of performance, as well as identifying different positional types that may not be typically understood by coaching staff.[13,14] For example, three positional groups (guards, forwards, and centres) have been historically identified within basketball. However, using clustering techniques, six different positional groups were identified via technical basketball performance data[13] and five different groups using only anthropomorphic data.[14] As such, it may be important to consider novel performance modelling techniques when exploring the various demands of RL performance in order to better understand the relationships between different positional groups and their performance indicators. Previous work in RL has observed differences in positional technical performance demands using a select number of technical variables.[3,4] Additionally, changes in collective team performance indicators have been identified between the 2005-2011 and 2012-2016 NRL seasons.[6] However it is unclear whether the positional specific demands of RL athletes differed across seasons, or whether there was relative positional stability over time. Overall technical performance demands of teams in the NRL were reported to have evolved,[6] which may subsequently have led to a change in the positional demands of NRL athletes, however, this is yet to be identified.

The aim of this study was to investigate whether technical performance demands of different positional groups in the NRL had changed over recent years (2015-2019), and whether playing positions could be accurately classified and discriminated using performance indicators from the NRL. Findings could assist coaches in understanding the current trends of positional technical performance demands, and subsequently improve decision making with regards to game strategy, training planning and personnel selection.

Methods

Forty-eight performance indicators were collected from a licensed central database (Analyzer; The League Analyst, Version V4.14.318) containing indicators from all NRL games during the 2015-2019 seasons (34, 047observations)(Supplementary Table 1). The performance indicators were chosen based on consultation with current NRL coaching staff and were similar to those previously examined and normalised against playing time.[5,6] Players were *a priori* classified based on their coach-selected starting line-up and playing number, and then further classified per game into four playing groups.[4] These positional groups have previously been reported to exhibit different physical[16-18] and technical skill demands[3,4,19] in RL athletes (Supplementary Table 2). Data was collated and then analysed in accordance with approval from the local institutional Human Research Ethics Committee.

Statistical Analysis

All statistical analyses were carried out using the statistical software IBM SPSS for Windows version 25 (Armonk, NY, USA, IBM Corp.).[20] One-way analysis of variance (ANOVA) was performed to examine changes in the selected technical performance indicators between 2015 to 2019, for each positional group to identify consistency over time and enable subsequent cluster analysis with differences identified via Bonferroni post-hoc analysis.

Classification of positional groups was achieved via a three-step process: (1) principal component analysis (PCA); (2) two-step cluster analysis; and (3) discriminant analysis.[13] PCA is commonly used as a dimension reduction technique that involves reducing the total number of observed variables into 'n' number of factors.[21] This is achieved by transforming a set of possibly linear variables into a separate set of linearly uncorrelated variables (principal components; Table 1). These factors were determined using eigenvalues above 1 (Table 2) and further extracted from the rotated component matrix for values above 0.60 (see supplementary Table 3).[21,22]

The factors obtained from the PCA were then incorporated into a two-step cluster analysis to model natural positional groups within the dataset. Two-step cluster analysis automatically determines the "optimal" number of clusters (positional groups) by using the Schwartz's Bayesian Information criterion.[23] In order to determine the "goodness" of the determined solution, the silhouette coefficient was used as a measure to cluster cohesion and separation.[23,24] Additionally, the log-likelihood distance measure was used to calculate the similarity between clusters.[23] Finally, discriminant analysis was used to better differentiate the positional groups determined by the two-step cluster. This approach provides classification functions that best discriminate among clusters (i.e., check which cluster each player best fits).[21] Structure coefficient (SC) values greater than |0.30| were considered significant for identifying the variance of positional technical performance.[14]

Results

The results of one-way ANOVA revealed significant changes in 36 of 48 (73%) technical performance characteristics ($F = 2.3 – 687.7$; $p = 0 - 0.05$; partial $\eta^2 = 0.00 – 0.075$) across the chosen time-period (2015-2019 NRL seasons). The performance characteristics which differed across seasons were: runs, run metres, line break, line break assist, hit ups, kick breaks, try assist, tackle break, play the ball win, play the ball loss, botch try, handling errors, penalty conceded, penalty won, decoy, support, metres after contact, tackle made, tackle miss, tackle forced turnover, scarps, kick pressure, intercepts, try saves, penalty conceded (def), conceded line break, try cause failed kick defusal, kick metres, field goal made, field goal miss, penalty made, kick errors, kick dead, kick caught in goal. PCA revealed fourteen factors (principal components, Table 1) that explained the variance of different performance outcomes based on the performance indicators (Supplementary Table 1). Factor 1 (forward attacking play) explained 13.7% of the total variance, while factor 2 (general play kicking) accounted for 6.4% and factor 3 (kick pressure) explained 5.4%. The cumulative loading for all fourteen factors accounted for 58.5% of the variance of positional technical performance across the competition.

[Insert Table 1 approximately here]

The two-step cluster analysis (Figure 1 and Supplementary Table 3) achieved a good silhouette measure of cohesion and separation (average silhouette = 0.5) revealing six clusters as opposed to the four *a priori* positional classifications. The clusters were: cluster 1 'backs' (20.9% of all players; 100% accuracy); cluster 4 'adjustables' (17.2% of all players; 100% accuracy); cluster 3 'interchange' (23.2% of all players; 99.9% accuracy); and cluster 6 'forwards' (26.7% of all players; 100% accuracy). The two additional clusters which were identified were cluster 2 labelled as 'utility back' (6.5% of all players) which consisted of a combination of two *a priori* classified groups, 'adjustables' (74.7%) and 'backs' (17.3%) players; and cluster 3 labelled as 'interchange forwards' (5.5% of all players) consisting of a combination of 'interchange' (50.8%) and 'forwards' (30%).

[Insert Table 2 approximately here]

The discriminant analysis revealed that 62.9% of the originally grouped clusters (i.e. two-step clustering) were correctly classified using the 14 factors obtained via PCA. The greatest level of classification accuracy occurred in cluster 1 (backs; 93.2%), followed by cluster 2 (utility back; 84.2%), cluster 3 (interchange forwards; 65.5%), cluster 4 (adjustables; 64.1%), cluster 6 (forwards; 50.4%, interchange; 37.7%) and cluster 5 (interchange; 42.3%, forwards; 41.6%). The discriminant analysis identified five significant discriminant functions (accounting for variance of kick conversions, general attacking play, penalties, general play kicking and scoring attacking play, respectively). The significant factors were forward attacking play (functions 2 and 5: SC=0.39 and SC=-0.34, respectively), general play kicking (function 4: SC=-0.63), kick pressure (function 4: SC=-0.51), conversions (function 1: SC=0.44), penalties (function 3: SC=-0.51), try causes (function 5: SC=-0.33), try assists (function 5: SC=0.38) and supports (function 5: SC=0.63).

[Insert Figure 1 approximately here]

Discussion

This study investigated whether there have been any changes over seasons in the technical performance indicators of different positional groups in the NRL, and whether these positional groups could be classified and discriminated based on performance indicators. The results identified significant changes in the performance indicators over the selected time-period with 27% of indicators (e.g. tries, kick defused, conversion made/miss, kick 40/20) stable across the 2015-2019 NRL seasons. Further, a model was created, which accurately classified playing position based upon a series of factors derived from commonly used performance indicators.[13,14] Collectively, these findings identified a newly developed model confirming the efficacy of unsupervised classification analysis for positional technical performance in RL. As such, with the large amount of data available to sports teams, the use of an unsupervised classification approach such as PCA, sports practitioners will be able to refine the vast amount of data available to them, into information that they may find more useful. Subsequently, the positional classification characteristics identified in this study may also allow sports practitioners to better prepare current players for their specified role, manage recruitment, and potentially identify new positions better suited for current players.

A major finding of this study was the observed variation in technical performance characteristics over the chosen time-period (2015-2019). This finding is supported by previous research which had observed changes in league-wide technical performance over 11 seasons (2005-2016).[6] The authors suggested that the introduction of a series of new rules by the NRL prior to the commencement of the 2016 NRL season, namely the reduction in interchanges (from 10 to 8) and the introduction of a 'shot clock' (35 seconds for scrums and 30 seconds for dropouts) may have augmented the subsequent outputs of players.[6] Potentially, the individual playing style of teams and how playing positions were utilised within that style, rather than the specific role of each playing position, may have contributed to the contrasting different result.[4,6,11] Regardless, it is evident that the technical performance demands in the NRL is constantly evolving, which has been further supported by the results of this study. As such, it is important that teams are constantly monitoring these changes, such that coaching staff can make informed decisions regarding training and strategizing game tactics.

The model produced in this study was successful in identifying six positional clusters, with a good level of accuracy (i.e. successfully assigning 89.4% of the players to their *a priori* cluster). This result highlights the suitability of clustering analysis to assist performance staff with accurate classification of RL playing positions using competition performance. As such, this approach may be further applied to talent identification or recruitment strategies, as it may identify players in other competitions (e.g. Super League, Reserve Grade, U20s) through comparisons of their performance against other players in the NRL (and possibly their most suited position). Combining match technical performance characteristics with other important physical measures could form part of a robust talent identification tool.[25]

Another intriguing result from the cluster analysis was the identification of two additional clusters. The first additional cluster (cluster 2) consisted of a combination of adjustables (74.7%) and backs (17.3%), who exhibited a unique set of technical performance characteristics which have been labelled as a 'utility back' group. The main features of this group were kicking (including goal kicking and kick breaks), try assists, intercepts, try causes and botched tries. The other additional positional cluster (cluster 3) consisted primarily of a combination of interchange (50.8%) and forward players (30%), which have subsequently been labelled as 'interchange forwards'. The main features of this 'interchange forwards' group were forward attacking play, defensive decisions, penalties, kick pressure, try assists, try saves and handling errors. Discriminant analysis further revealed that 84.2% of players classified as a 'utility back' would have been reclassified in the same cluster and 65.5% 'interchange forwards' reclassified in the same cluster with the remainder primarily reclassified amongst adjustables (9.5%) and forwards (13.5%). One of the most representative examples of the 'utility back' playing group was Player X who would traditionally be considered a 'fullback' (adjustable) but was re-classified as 'utility back' for 36% of matches and an 'adjustable' for 64% of 112 matches. Whereas one of the more representative examples of the 'Interchange Forward' group was Player Y (47 % of 98 matches as 'Interchange Forward', 22% as 'Interchange' and 31% as 'Forward'). It is however unclear whether one or both of these additional positional groups were commonly featured amongst all teams, or whether successful (or unsuccessful) teams consisted of these types of players. As such, further investigation into the influence of this positional group on match outcome may be of value to coaching and performance staff regarding tactical game planning and player development and recruitment strategies.

Discriminant analysis revealed the difficulty of reclassifying ‘interchange’ players into the same cluster, with 42.3% of interchange players successfully reclassified in cluster 5, and 41.6% assigned to cluster 6 (forwards). Given it is common practice for NRL teams to assign multiple (often three out of four) spots on their interchange towards forward-positional players, it is unsurprising that there was a level of misclassification that occurred during this analysis process. Given this, it could be assumed that ‘interchange’ players were expected to be able to make similar performance contributions to the team as ‘forwards’. An example of this would be Player Z (97.1% of 110 matches as ‘Forward’; 2.9% of matches as ‘Interchange’), who was traditionally considered a ‘Interchange Forward’ compared to Player Y (47 % of 98 matches as ‘Forward’, 22% as ‘Interchange’ and 31% as ‘Forward’) who would also be considered a ‘Forward’. Both of these players would be considered to be within the ‘Forward’ group, as classified *a priori* however, the individual match performance of Player Z was variable compared to that of a ‘Forward’ and fluctuated between a starting and reserve role. As such, coaches should ensure any positional specific training that is planned, gives similar opportunity to players that undertake similar roles irrespective of start position (field or bench).

Limitations

The current study highlighted the efficacy of unsupervised classification for positional technical performance in RL over recent seasons through the use of PCA, two-step clustering and discriminant analysis. However, in contrast to previous research, this study only sampled five seasons worth of data compared to previous research which observed changes over 11 seasons.[6] In saying this, changes noted in this study are similar to prior research[6], confirming that the NRL is evolving and that a larger observational periods may be required to gain a deeper insight into the evolution of playing position in the NRL. Additionally, it is important to note that the *a priori* classification of NRL playing positions was determined by how players were initially listed when their teams were announced prior to the game. As such, players named outside of the 17 initially intended to be playing, were assigned numbers beyond 17 (e.g. 18, 19, 20, etc.). For example, a player who was replaced from outside the original 17 at late notice due to injury (e.g. back) was unable to be differentiated from the interchange group, and as such may have resulted in some initial *a priori* misclassification. However, the unsupervised approaches used in this study overcome this issue, as the analysis determines which positional group each player falls into, rather than coaches.

Conclusion

This study identified changes in the technical performance demands of NRL players across the sampled seasons in the NRL (2015-2019). The current study also demonstrated the usefulness of both clustering (two-step) and classification (discriminant analysis) approaches to understanding the positional technical performance characteristics of NRL players. The high level of classification accuracy achieved from these approaches indicated that the chosen analytical techniques could be used to support sports practitioners in their evaluation of player performance and future positional suitability (e.g. talent identification, personnel recruitment). More importantly, this study highlighted the utility of unsupervised analytical approaches for sports practitioners, as they can offer insights into queries that they may not be able to resolve using traditional analytical approaches.

Disclosure of interest: The authors report no conflict of interest

References

1. Twist C, Highton J, Waldron M, Edwards E, Austin D, Gabbett TJ. Movement demands of elite rugby league players during Australian National Rugby League and European Super League matches. *Int. J. Sports Physiol. Perform.* 2014; 9(6):925-930.
2. Hulin BT, Gabbett TJ, Kearney S, Corvo A. Physical Demands of Match Play in Successful and Less-Successful Elite Rugby League Teams. *Int. J. Sports Physiol. Perform.* 2015; 10(6):703-710.
3. Bennett KJ, Fransen J, Scott BR, Sanctuary CE, Gabbett TJ, Dascombe BJ. Positional group significantly influences the offensive and defensive skill involvements of junior representative rugby league players during match play. *J. Sports Sci.* 2016; 34(16):1542-1546.
4. Sirotic AC, Knowles H, Catterick C, Coutts AJ. Positional match demands of professional rugby league competition. *J. Strength Cond. Res.* 2011; 25(11):3076-3087.
5. Woods C, Sinclair W, Robertson S. Explaining match outcome and ladder position in the National Rugby League using team performance indicators. *J. Sci. Med. Sport.* 2017.
6. Woods C, Robertson S, Sinclair WH, Collier NF. Non-metric multidimensional performance indicator scaling reveals seasonal and team dissimilarity within the National Rugby League. *J. Sci. Med. Sport.* 2018.
7. Kempton T, Kennedy N, Coutts AJ. The expected value of possession in professional rugby league match-play. *J. Sports Sci.* 2016; 34(7):645-650.
8. Bush M, Barnes C, Archer DT, Hogg B, Bradley PS. Evolution of match performance parameters for various playing positions in the English Premier League. *Hum, Movement Sci.* 2015; 39:1-11.
9. Woods C, Veale J, Fransen J, Robertson S, Collier NF. Classification of playing position in elite junior Australian football using technical skill indicators. *J. Sports Sci.* 2018; 36(1):97-103.
10. Woods C, Veale JP, Collier N, Robertson S. The use of player physical and technical skill match activity profiles to predict position in the Australian Football League draft. *J. Sports Sci.* 2017; 35(4):325-330.
11. Woods C, Leicht A, Jones B, Till K. Game-play characteristics differ between the European Super League and the National Rugby League: Implications for coaching and talent recruitment. *Int. J. Sports Sci. & Coach.* 2018:174795411878844.
12. Robertson S, Bartlett JD, Gastin PB. Red, amber or green? Athlete monioring in team sport: the need for decision support systems. *Int. J. Sports Physiol. Perform.* 2017; 12(2).
13. Sampaio J, Goncalves B, Mateus N, Shaoliang Z, Leite N. Basketball, Chapter 6, in *Modelling and simulation in sport and exercise scienc*, Routledge, 2018.
14. Zhang S, Lorenzo A, Gomez MA, Mateus N, Goncalves B, Sampaio J. Clustering performances in the NBA according to players' anthropometric attributes and playing experience. *J. Sports Sci.* 2018:1-10.
15. Lago-Peñas C, Gómez-Ruano M, Yang G. Styles of play in professional soccer: an approach of the Chinese Soccer Super League. *Int. J. Perform. Anal. Sport.* 2018; 17(6):1073-1084.
16. Austin D, Gabbett T, Jenkins D. Repeated High-Intensity Exercise in Professional Rugby League. *J. Strength Cond. Res.* 2011; 25(7):1898-1904.

17. Gabbett T, Jenkins D, Abernethy B. Physical collisions and injury during professional rugby league skills training. *J. Sci. Med. Sport.* 2010; 13(6):578-583.
18. Gabbett T, Jenkins D, Abernethy B. Physical demands of professional rugby league training and competition using microtechnology. *J. Sci. Med. Sport.* 2012; 15(1):80-86.
19. Gabbett T, Kelly J, Pezet T. A comparison of fitness and skill among playing positions in sub-elite rugby league players. *J. Sci. Med. Sport.* 2008; 11(6):585-592.
20. Corp. I. IBM SPSS Statistics for Windows. In: Corp. I, ed. 25 ed. Armonk, NY, USA: IBM Corp.; 2017.
21. Jolliffe I. Principal Component Analysis, in *International Encyclopedia of Statistical Science*. Lovric M, Berlin, Heidelberg, Springer Berlin Heidelberg, 2011.
22. Weaving D, Beggs C, Dalton-Barron N, Jones B, Abt G. Visualising the complexity of the athlete monitoring cycle through principal component analysis. *Int. J. Sports Physiol. Perform.* 2019:1-31.
23. Wendler T, Gröttrup S. Cluster Analysis, Chapter 7, in *Data Mining with SPSS Modeler: Theory, Exercises and Solutions*, Springer, 2016.
24. Norusis M. Cluster Analysis, Chapter 17, in *IBM SPSS Statistics 19 Statistical Procedures Companion*, Prentice Hall, 2011.
25. Pion J, Lenoir M, Vandorpe B, Segers V. Talent in Female Gymnastics: a Survival Analysis Based upon Performance Characteristics. *Int. J. Sports Med.* 2015; 36:935-940.

Figures and table captions

Figure 1. Two-Step Cluster analysis results identifying six distinct playing positions clusters.

Table 1. Principal components and the associated technical performance characteristics.

Table 2. Eigenvalues for principal components and total variance explained.

Supplementary Table 1. Description of assessed technical skill performance metrics.

Supplementary Table 2. Description of *a priori* playing positions.

Supplementary Table 3. Rotated component matrix of technical performance indicators.

Supplementary Table 4. Descriptive statistics (mean ± SD) for the technical performance characteristics across seasons normalized to playing time.