Paper: Rugby Football

Categorization of Rugby Union Players by Performance Characteristics Using Principal Component Analysis and Cluster Analysis

Makoto Kiuchi[*], Hirofumi Maehana[**] and Nobuyoshi Hirotsu[*]

[*]Graduate School of Health and Sports Science, Juntendo University
1-1 Hiragagakuendai, Inzai, Chiba 270-1695 Japan
makotos621115@yahoo.co.jp
[**]Department of Psychological Counseling Faculty of Human Sciences, Mejiro University
[Received October 4, 2019 ; Accepted August 4, 2020]

The aim of this study was to categorize rugby players by performance characteristics irrespective of positions using principal component analysis (PCA) and cluster analysis (CA). Data were drawn from the Japan Rugby Top League 2015–2016 season; a sample of 231 players and 16 items were used in the analysis. We used PCA to identify the players' ability and CA to categorize players' performance characteristics, regardless of position. PCA reduced the 16 items to four principal components: penetration, defense and competition, ball handling and kicking, and turnovers. We categorized players into six clusters based on the four component scores: playmaker, ball carrier, tackler and supporter, competitor, passer and kicker, and spoiler. Although player categorization by CA was achieved regardless of position, the players' performance characteristics also influenced categorization, and specific positions have performance characteristics. Unique players were identified for some positions. In addition, we clarified a difference in performance characteristics between Japanese and foreign players in the league. Since the categorization of players by performance characteristics irrespective of position enables teams to find necessary players, this study may be advantageous for team management and may contribute to more useful analysis of players and clarifying performance characteristics.

Keywords: game data, performance characteristics, player categorization, principal component analysis, cluster analysis

[Football Science Vol.17, 86-97, 2020]

1. Introduction

In rugby, players with varied performances is increasing in many domestic competitions due to the rule allowing foreign players to be selected for national teams of other countries, providing certain conditions are met. For example, the number of players from other countries such as New Zealand, Australia, and South Africa has increased in the Japan Rugby Top League (JRTL). The JRTL currently has a its most diverse collection of players ever, with each player also having different performance characteristics.

The performance characteristics of players determine their role in the team. In many cases, individual players may be assigned specific roles on a team depending on their performance characteristics. When a team is constructed with the aim of strengthening it, it is useful to categorize players by performance characteristics, because the players who can perform the role on the team could be identified. This means that a team that performs an open attack requires a player who can have characteristics of playing a pass to succeed. Since the success of a team greatly depends on individual performance (Lim et al., 2009), categorizing players by performance characteristics may lead to greater success. Therefore, categorizing players by their performance characteristics can be helpful in creating a more effective selection process.

In rugby, there are ten positions, divided into forwards (FW) and backs (BK). In the FW pack, there is a prop (PR), a hooker (HO), a lock (LO), a flanker (FL), and number 8 (NO8); and the BK consists of a scrum-half (SH), a fly-half (FH), center (CTB), wing (WTB), and fullback (FB). A player is generally chosen depending on the position and the trend of performance (Duthie et al., 2006; Parsons and Hughes, 2001; Quarrie et al.,

2013; Villarejo et al., 2013; Vivian et al., 2001). The performance characteristics of players might relate to their positions. For example, an FH performs many passes and kicks to move forward to the ball and territory for other members of the team. Additionally, an LO is performing many tackles and providing support during a field play (Parsons and Hughes, 2001).

Although the performance characteristics of players relate to positions, categorizing players firstly based on performance characteristics without considering players' positions, could contribute to finding a more varied team composition. For example, a team that has a weak point in its defense might find players that have characteristics to tackle in the SH. Or, a team that is weak in ball possession might select players, such as LOs, who are good ball carriers. The players who have this kind of characteristics may not be usual players, but they could be needed in a specific team. As the variation of the performance is increasing in modern rugby with the aim of strengthening a team, categorizing players by their performance characteristics regardless of position, i.e., irrespective of position, could be helpful to understand performance characteristics of individual players and contributes to strengthen a team. This is important in modern rugby, especially now that there is increased variation among the performance of the players.

Categorization of players by performance characteristics may be conducted using game data. Game data contain important information related to performance during a game. Therefore, game data were used to evaluate players. James et al. (2005) developed a position-specific evaluation indicator to assess compare the players of each position. Their evaluation method has been used to identify the players' playing patterns through a comparison of such data. However, since rugby is complicated and multifaceted (International Rugby, 2011), a comprehensive analysis of the performance of each player needed to be conducted.

We believed that principal component analysis (PCA) and cluster analysis (CA) would be effective categorizing the players by performance characteristics from a comprehensive analysis. PCA can reduce data into several factors through a grouping process and emphasize the key factors that should be analyzed, and can create the synthetic variable that could present the ability of performance of the player related to each component. Thus, by using PCA to analyze the game data, the player can be evaluated comprehensively with their synthetic variables. On the other hand, CA can be used for classifications in which similar items are partitioned into groups. Thus, by using CA with the data of their synthetic variables, players who have similar performance characteristics during the match could be categorized. This will help us to identify the performance characteristics of the individual players. Hence, through a comprehensive analysis using PCA and CA, we could categorize players by performance characteristics.

The usefulness of such analysis has been seen in other sports: PCA analysis has been used for soccer (football) (Barros et al., 2006; Fernandez-Navarro et al., 2016), tennis (Kramer et al., 2017), and Olympic records data (Naik and Khattree, 1996). For example, Fernandez-Navarro et al. (2016) analyzed the playing style of a soccer team comprehensively using PCA. In their study, 19 items of game data were reduced to six components that displayed the team performance characteristics from each component score. Likewise, CA has been used to categorize the performance trend of a position in rugby (Quarrie et al., 2013), which displays the differences between the activities and time-motion of the positional groups, both between and within the FW and the BK. Moura et al. (2014) conducted a study that applied both PCA and CA, summarizing the game data of soccer teams using PCA and then classifying them into winning and losing teams based on their component scores using CA. Similarly, through PCA and CA, rugby players can be classified depending on their performance characteristics.

The advantage of categorizing players by performance characteristics, irrespective of position, is the potential to identify suitable players for teams. Clarifying the kind of performance characteristics each player can perform contributes to the acquisition and appointment of players by the team's management. Therefore, the aim of this study was to categorize players by performance characteristics, irrespective of positions using PCA and CA.

In the JRTL, there are players with a various range of performance characteristics from not only Japan but also elsewhere in the world, and they have a variety of skills. In this study, the classification of players by performance characteristics in the JRTL was used to highlight such variation.

2. Methods

2.1. Data

Data from the JRTL 2015–2016 season was provided by Data Stadium Inc., which was certified as the official statistics provider by the JRTL. There were ten games in the JRTL 2015–2016 season. Officially, each game lasts 80 minutes. Since each team played ten games, the total time played per season was 80 × 10 = 800 minutes. For this study, we analyzed the data of 231 players who had played for more than 400 minutes, or about one-half of the total time that season. The players who played different positions for each game were categorized, depending on the position that they played most often.

We selected items that can be played in any position to the comprehensive analysis of all players enabled. Therefore, the field play was analyzed, except for set pieces such as place kicks, lineouts, and scrums. Data analysis was conducted based on the 16 items indicated in **Table 1**.

Ethical approval for this study was granted by the ethics committee of Juntendo University (Faculty of Health and Sports Sciences, Juntendo University, Chiba, Japan).

2.2. Statistical Analysis

We used PCA to extract components from the 16 items (see **Table 1**). The criterion for component extraction was that the eigenvalue exceeded 1.0. Play items with principal component loadings greater than 0.5 displayed a positive or negative correlation and indicated a substantial value for each component. In addition, the component scores were quantified by categorizing them into six clusters using CA by the k-means method. We used the SPSS statistics software (Version 22.0) for this analysis.

Table 1 Variables studied in player assessment

Items	Operational Definition	Mean	SD	Max	Min
Try scored	5 points awarded to the scoring team when the ball is placed down in the try area.	1.8	1.9	10	0
Ball touch	Catching the ball under opposition pressure. Includes picking the ball from base of ruck.	93.4	71.3	482	7
Ball carry	A player touching the ball is deemed to make a carry if they have made an obvious attempt to go forward and attack the opposition with the ball in hand.	45.8	23.8	134	6
Pass	Passing the ball under opposition pressure.	69.6	120.3	616	0
Kick	Team regains possession after a tactical kick.	10.6	20.9	114	0
Contact	Carrying the ball into the opposition defensive line and opposition players to commit to a tackle situation.	52.2	28.7	155	6
Line break	A ball carrier on attack breaks through the defensive line.	2.1	2.4	12	0
Off-load	Passing the ball on to a supporting player when being tackled, thus, maintaining the forward flow of play.	4.3	4.6	27	0
Tackle break	The ball-carrier breaks through an attempted tackle.	8.0	7.5	36	0
Support	Supporting player to arrive in a tackle situation to lend attacking support.	91.1	59.6	296	2
Turnover lost	Results in turnover of possession to the opposition.	2.0	1.7	8	0
Tackle	Tackling of opposition player.	61.0	25.0	146	12
Tackle assist	Assisting in a tackle situation.	17.4	10.4	48	1
Jackal	Attempt to steal the ball from a tackle situation in a turnover attempt.	6.0	6.7	37	0
Turnover won	Results in turnover of possession.	3.1	2.7	13	0
Penalty	Infringement of the laws of the game resulting in penalty.	6.8	3.8	25	1

3. Results

3.1. Clarification of Player Ability

To clarify the performance characteristics of the players, we analyzed 16 data items using PCA and these to four components. The coefficients from the first to the fourth principal components are shown in **Table 2**. The contribution ratio of the four components was 69.93%.

We interpreted each component based on the group of associated items. The scores of the first component were associated with seven items (i.e., contact, ball carry, tackle break, off-load pass, line break, try scored, and turnover lost) and represented the performance associated with "penetration." The second component was associated with five items (i.e., support, tackle assist,

Table 2 Coefficients from the first to fourth principal components for each item

Component Items	First (Penetration)	Second (Defense and Competition)	Third (Ball Handling and Kicking)	Fourth (Turnover)
Contact	**.917**	.123	-.147	-.181
Ball carry	**.891**	.179	-.130	-.206
Tackle break	**.843**	-.235	-.193	-.017
Off-load	**.737**	-.213	.072	.153
Line break	**.638**	-.424	-.216	.154
Try scored	**.605**	-.180	-.250	.275
Turnover lost	**.530**	-.249	-.233	-.242
Support	-.007	**.844**	-.083	-.219
Tackle assist	.206	**.784**	.255	-.170
Tackle	.264	**.646**	.446	-.183
Jackal	.425	**.569**	.151	.208
Kick	.102	**-.627**	**.609**	-.271
Pass	-.005	-.466	**.581**	.363
Ball touch	.435	-.406	**.559**	-.457
Turnover won	.399	.284	.354	**.523**
Penalty	.390	.388	.237	.325

Principal component loadings in bold indicated a strong positive or negative correlation

tackle, jackal, and kick) and related to "defense and competition." The third component was associated with three items (i.e., kick, pass, and ball touch) and related to "ball handling and kicking." The fourth component was associated with one item (i.e., turnover won) and related to "turnover." Each factor related to the performance ability of the player.

Calculating the data using PCA could facilitate the identification of the comprehensive ability of individual players related to each factor, by calculating synthetic variables. **Table 3** represents the raw data and each component score of PR players. For example, reviewing at the raw data, we can infer that, since Toyoda does not have the highest values for all items related to the first component, he may not be a player whose specialization is penetration in PR. However, Toyoda is a player who is identified by the greatest number of penetrations because he has the highest first component score in PR. Although it may be difficult to grasp comprehensively the ability of players from the individual raw data, by calculating raw data using PCA, the ability of the players could be understood easily. Players who received a high score for each component performed in a way related to each factor. A player with a high score in the first component performed "penetration" more than a low-score player. A player with a high score in the second component also performed "defense and competition" more than a low-score player.

Thus, it was possible to calculate the principal component scores for each player to more easily explain the ability of the player compared to using comparisons of single items as in previous studies (Enrique et al., 2009; James et al., 2005; Michele et al., 2006).

3.2. Categorization of Players by Performance characteristics

We analyzed player data using CA to categorize them by performance characteristics. We used their principal component scores to categorize the players into six groups. The average component scores of each cluster are shown in **Table 4**. Cluster 1 (CL 1) has a first component score of 0.36, a second component score of 1.33, a third component score of 1.00, and a fourth component score of 1.05. Note that component scores were normalized as an average 0 and variance 1 in the SPSS output.

Players categorized in CL 1 had positive first and third component scores and had negative second and fourth component scores (**Table 4**). That means that although those players sometimes penetrated, mainly performed ball handling and kicking. In other words, they approached the opposing defensive line and distributed the ball during the game, and they usually assisted in the

Table 3 Raw data and each component score in individual player of PR

Player Name	The raw data of the individual player																Component Score			
	Try scored	Ball carry	Contact	Line break	Off load	Tackle break	Turnover lost	Support	Tackle	Tackle assist	Jackal	Ball touch	Pass	Kick	Penalty	Turnover won	First	Second	Third	Fourth
Toyoda	2	70	78	1	1	9	0	142	59	16	6	83	24	0	20	4	0.38	0.98	0.12	0.81
Hirahara	0	74	80	2	1	5	2	202	90	20	6	95	32	0	10	1	0.17	1.1	-0.08	-1.3
Inagai	0	59	60	0	0	1	4	182	97	38	9	80	15	0	11	2	-0.03	1.64	0.4	-1.33
Morikawa	1	52	62	2	2	10	4	195	58	10	2	60	10	0	7	0	-0.08	0.25	-1.17	-1.14
・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・
Ueda	0	9	9	0	0	0	1	118	52	12	1	14	12	0	6	1	-1.52	0.2	-0.48	0.1
Kitagawa	0	11	10	0	0	0	1	85	33	14	0	10	3	0	3	3	-1.55	-0.06	-0.65	0.35
Asahara	0	10	10	0	0	0	0	191	61	25	2	12	3	0	3	0	-1.58	0.81	-0.4	-0.61
Yamamura	0	6	6	0	0	0	0	106	19	8	0	8	7	0	1	0	-1.88	-0.33	-1.08	0.02
Mean	0.43	31.50	33.50	0.50	0.97	2.13	0.93	139.90	57.23	20.83	4.30	36.90	9.50	0.03	7.53	1.87	-0.86	0.64	-0.28	-0.10
SD	0.80	19.22	21.34	0.81	1.25	2.73	1.18	39.80	20.28	10.47	5.44	22.99	6.81	0.18	4.32	1.09	0.58	0.53	0.43	0.68
Max	3.00	74.00	80.00	3.00	6.00	10.00	4.00	202.00	104.00	48.00	30.00	95.00	32.00	1.00	20.00	4.00	0.38	1.95	0.74	1.00
Min	0.00	6.00	6.00	0.00	0.00	0.00	0.00	63.00	19.00	6.00	0.00	8.00	0.00	0.00	1.00	0.00	-1.88	-0.33	-1.17	-1.33

Table 4 The average principal component scores of each cluster

Cluster / Component	(CL 1)	(CL 2)	(CL 3)	(CL 4)	(CL 5)	(CL 6)
First component (Penetration)	0.36	0.45	-0.62	0.79	-0.03	-0.44
Second component (Defense and Competition)	-1.33	-0.73	0.49	1.06	-1.38	-1.14
Third component (Ball Handling and Kicking)	1.00	-1.06	-0.21	0.40	3.10	1.35
Fourth component (Turnover)	-1.05	0.04	-0.30	0.58	-1.64	1.90

creation of breaks for their teammates and in creating opportunities for scoring a try. On the other hand, CL1 players performed little defense, competition, and turnover. Players categorized in cluster 2 (CL 2) had a positive first component score. Those players aimed to break through the opposing defensive line. Players in cluster 3 (CL 3) had a positive second component score and had other negative components. Since those players supported the ball carrier, defending against opposing attacks. However, those players could not play holding the ball in the game. Players in cluster 4 (CL 4) had all positive component scores, and the first and second component scores, in particular, were related. The players of CL 4 performed ball carrying, competing, and tackling opposite players. Players in cluster 5 (CL 5) had a positive third component score and had other negative components. Those players selected specialized in the option of kicking or passing in attack. Players in cluster 6 (CL 6) had positive third and fourth component scores, and had negative first and second component scores. Since they distributed their own team's ball and stole the opponents' ball. On the other hand, they could not conduct penetration, tackle, and competition. In this way, based on the principal component scores, the players were categorized into six performance characteristics depending on their ability.

3.3. Relationship Between Performance Characteristics and Position

The performance characteristics of players seemed to roughly depend on positions; however, specific positions were also associated with certain performance characteristics. The number and percentage of players in each cluster are indicated in **Table 5** according to their positions. For example, of the 30 players whose PR were analyzed, 28 (93.3%) were classified as CL3, and 2 (6.7%) were classified as CL4. First, the relationship between the performance characteristics and position for the FW will be explained.

The majority of the FW players were roughly categorized as CL3 and CL4. Of the PR players, 28 out

Table 5 Number of players in each position in each cluster

Position \ Cluster		(CL 1)	(CL 2)	(CL 3)	(CL 4)	(CL 5)	(CL 6)	Total
FW	PR	0	0	28(93.3%)	2(6.7%)	0	0	30(100%)
	HO	0	0	10(83.3%)	2(16.7%)	0	0	12(100%)
	LO	0	2(6.7%)	22(73.3%)	6(20.0%)	0	0	30(100%)
	FL	0	1(2.7%)	12(32.4%)	24(64.9%)	0	0	37(100%)
	NO8	0	3(30.0%)	2(20.0%)	5(50.0%)	0	0	10(100%)
BK	SH	0	0	1(6.2%) *	0	0	15(93.7%)	16(100%)
	FH	12(63.1%)	0	1(5.3%)	0	6(31.6%)	0	19(100%)
	CTB	1(3.4%)	10(34.5%)	14(48.3%)	4(13.8%)	0	0	29(100%)
	WTB	0	32(100%)	0	0	0	0	32(100%)
	FB	6(37.5%)	6(37.5%)	3(18.8%)	0	0	1(6.2%)	16(100%)
Total		19	54	93	43	6	16	231

* Takashi Suzuki

of 30 were classified as CL3, as were 83% of the HO and 73% of the LO. The FL and the NO8 were mainly classified as CL4.

In contrast, in the BK, the positions of SH and WTB were clearly classified. For instance, 94% of the SH were categorized as CL6 and all the WTB players as CL2. However, the FH, CTB, and FB positions might have certain performance characteristics. Most of the FH players were categorized as CL1, but 32% were also classified as CL5. Additionally, the CTB and FB positions had four performance characteristics. Thus, the performance characteristics of players seemed to roughly depend on positions.

Since this study categorizes players by performance characteristics regardless of position, all players are analyzed under similar conditions. Therefore, in the results of this study, the influence of "position" that could not be considered in the analysis may be reflected. However, among all players, one with the same position and different performance characteristics could be indicated as a particularly unique player in that position. For example, the PR and HO players categorized in CL4 were different players in the same position. Similarly, the LO and FL players in CL2 and the SH players in CL3 were also different players. Categorizing players holistically could highlight that players with different performance characteristics exist even in the same position.

4. Discussion

In this section, the results of the categorization of players by performance characteristics are discussed, and certain applications for the selection of suitable players for a team are suggested.

4.1. Categorization of Players by Performance Characteristics

The players in each cluster were classified depending

on their characteristics. The component scores of the players used are shown in **Table 6**. Some players were used to explain each cluster. Tusi Pisi, a Samoan national team player in CL 1, has a score of 1.83 for the first component, 2.04 for the second, 0.57 for the third, and 1.33 for the fourth. Based on the results of PCA, he has the performance characteristics "penetration" and "ball handling and kicking" (**Table 6**). A Japan national team player, Hendrik Tui, is categorized as a CL2. Based on the results of PCA, he has the performance characteristic "penetration." Keita Inagaki, a Japan national team player, is categorized as a CL3. He can be regarded as specialized in "defense and competition." In this way, based on the principal component score, the players in each cluster were classified depending on their characteristics.

Each performance characteristic seems to display the player's main role, meaning the functional role. Strictly speaking, the functional roles defined by IRB (2011) are not the same as the performance characteristics, but each cluster could be named by reference to the functional role of the IRB (2011) to help our understandings: CL1 is "play maker," CL2 is "ball carrier," CL3 is "tackler and supporter." CL4 is "competitor," CL5 is "kicker and passer," and CL6 is "spoiler." Additionally, the names of these clusters were used by references as MacDonald and Rees (1938), Kraak and Welman (2014), and Olds (2001), under a similar context.

Representative players of each performance characteristics are indicated in **Table 7**. Players in CL1

Table 6 Component scores of players described in each cluster

Cluster	Player	Component score			
		First	Second	Third	Fourth
1	T. Pisi	1.83	-2.04	0.57	-1.33
2	H. Tui	3.90	-0.44	-0.97	0.48
3	K. Inagaki	-0.09	1.64	0.40	-1.33
4	M. Leitch	2.28	0.69	0.55	1.26
5	T. Monji	-0.10	-1.08	3.07	-1.73
6	Du Preez	-0.01	-2.19	1.12	2.57

Table 7 Representative players of each cluster

Cluster	Player				
1	T. Pisi	B. Barnes	F. Steyn	B. Foley	E. Jantjies
2	H. Tui	E. Etzebeth	J. P. Pietersen	M. Sau	F. Anderson
3	K. Inagaki	K. Inagaki	S. Makabe	Y. Ohdo	S. Ito
4	M. Leitch	A. Thomson	A. Bekker	R. Kahui	A. Mafi
5	T. Monji	H. Moriwaki	D. Konishi	T. Ootao	Y. Shigemitsu
6	Du Preez	A. Ellis	Y. Yatomi	C. Kim	T. Ogawa

are famous "playmakers" players. For example, Berrick Barnes, a former member of the Australian national team, was categorized as a "playmaker" (Bret, 2012). In addition, Eben Etzebeth of the South African national team was categorized as a "ball carrier" (Hefin, 2018). The players classified in each cluster had the same performance characteristics as the players described (**Table 6**). For that reason, it was suggested that players could be categorized depending on their performance characteristics.

The relationship between performance characteristics and position is explained as follows. The performance characteristics of a player may be roughly categorized depending on the position (**Table 5**). The roles of the FW and BK are different (Duthie et al., 2006; Villarejo et al., 2013). Many of the FW players are classified as "tacklers and supporters" or "competitors," and since the PR, HO, and LO players have few opportunities related to ball possession (Parsons and Hughes, 2001), they are considered to be "tacklers and supporters." The PR, HO, and LO players also take on the main role of set pieces such as scrums and line-outs (Bompa and Claro, 2009; Quarrie and Wilson, 2000). Most of the FL and NO8 players may be classified as "competitors" because they tackle and compete at the breakdown more than at other positions (Villarejo et al., 2013). In contrast, the BK players are mainly playmakers, ball carriers, passers and kickers, and spoilers.

These performance characteristics, excluding "ball carrier," have high third-component scores (ball handling and kicking), as indicated in **Table 4.** According to previous studies (Parsons and Hughes, 2001; Quarrie et al., 2013), the BK perform more kicks and passes than the FW in games. For this reason, most BK players are categorized as "playmakers", "passers and kickers," and "spoilers.". For example, the FH players were predominantly categorized as "playmakers" and "passers and kickers." In a study by James et al., (2005), two FH players were compared to identify their play patterns and then divided into players who were good at ball carrying or passing and kicking. Therefore, the categorization of the FH in this case was considered a reasonable result. It is possible to conclude that the role of the WTB may be to penetrate the opposition because they perform line breaks and off-loads more than other positions (Quarrie et al., 2013). Therefore, the WTB players may be categorized as "ball carriers." Based on the above information, the performance characteristics of players may be roughly categorized depending on the positions they play.

4.2 Application to Suitable Team Player Selection

Categorization of players by performance characteristics can be helpful in discovering the most suitable players for a team. First, the positions that have certain performance characteristics will be explained, followed by the identification of unique players associated with these positions. Finally, the difference between the performance characteristics of Japanese and foreign players will be discussed.

Although players in many positions were categorized depending on their performance characteristics, specific positions do have certain performance characteristics: the NO8, FH, CTB, and FB positions. The team tactics and relations to other players may influence players in these positions. In the case of CTB players, they may be "tacklers and supporters" when the team uses a kick to gain territory or when the team uses many ball carriers. In contrast, other CTB players may be needed to break through as "ball carriers" by the team. This study did not consider the "team" to which the subject player belongs, as it instead analyzed the entire player with raw data. As a result of this study, the variability in the classification of specific positions such as NO8, FH, CTB, and FB may indicate that those positions have a unique role for the team. Therefore, this method may be most useful in the analysis of these positions. Since the player requirements differ depending on the team, the categorization of players by their performance characteristics might be helpful in finding suitable players for a team.

Moreover, this method may assist in the discovery of interesting players with different performance characteristics in the same position such as the SH player, Takashi Suzuki, in "tackler and supporter" (**Table 5**). An SH generally has the role of passing between the FW and BK (International Rugby Board, 2011; Parsons and Hughes, 2001; Quarrie et al., 2013) and is not a "tackler and supporter." However, Suzuki performed a tackle and support role rather than a passing role; that is, his performance characteristics as an SH seems to be different from other SH players. This makes it possible to suggest that Takashi Suzuki would be a useful choice for the Japanese national team, which is weak at tackling compared to international top-level leagues (Kiuchi and Hirotsu, 2016). The Japanese national team needs to strengthen its tackling, and good tacklers are required in all positions. Although Takashi Suzuki is an SH and a tackler and supporter, our analysis showed that he has the potential to improve this weak point for Japan. Other

players of interest in this study included players such as the PR and HO players categorized as "competitors" and the LO and FL players categorized as "ball carriers." The benefit of this method is that players may be distinguished by different performance characteristics in the same position, which may contribute in finding players needed for a team.

However, take precaution when interpreting this analysis. A player such as Suzuki—who has different performance characteristics than a player in the same position—may be performing a role given by the team. Possibly, the LO and FL players categorized as "ball carriers," may be given special roles to help break through for the team. Although the player performance characteristics may depend on team tactics, the performance characteristics they were employing during the game could be somehow identified; this could also help to find the difference between the roles of Japanese and foreign players.

Since this method could be used to classify the differences between the performance characteristics of Japanese and foreign players, it may assist in the selection of foreign players. In particular, the results of the FH are notable. FH players play a decision-making role in the game (Greenwood, 1997; Parsons & Hughes, 2001). We categorized FH players as "playmakers" (CL 1) or "passers and kickers" (CL 5) (see **Table 8**). There were 12 FH players classified as "playmakers," and half of them were foreign players. The players categorized as "playmakers" perform more "penetration" and "ball handling and kicking" (**Table 4**), and the foreign FH players seem to have more of these characteristics. On the other hand, all the FH players classified as "passers and kickers" were Japanese players. The players classified as "passers and kickers" mostly perform "ball handling and kicking." For that reason, half of the Japanese FHs were assumed to set the team's playing style, especially with the use of passes and kicks.

The performance characteristics of "playmakers" and "passers and kickers" are different. Furthermore, most of the FHs who were foreign players categorized as "playmakers" also had experience playing on a national team. The results suggest an international trend in the performance characteristic of FH players to include not only ball handling and kicking but also penetration. If Japanese FH players improve their penetration skills and play more like playmakers, this might strengthen the Japanese national team. This might also contribute to nurturing better FH players. The analyses can aid in the selection of foreign players who can fill the performance characteristics of playmaker.

In this study, players were categorized depending

Table 8 Foreign and Japanese FH players

Cluster	(CL 1)		(CL 5)
	Foreign	Japanese	Japanese
Player	B. Barnes*	T. Fukuda	K. Hashino
	B. Foley*	K. Morita	D. Konishi
	M. Gerrard	J. Ogura*	T. Monji
	E. Jantjies*	K. Ono*	H. Moriwaki
	L. Timothy*	Y. Sato	T. Otao*
	T. Pisi*	R. Yamanaka*	Y. Shigemitsu

* National Team Player

on their performance characteristics. Players could be selected according to team tactics as an advantage of the method of this study. In terms of player acquisition, in the case of the team that does not have many players that can penetrate, they may acquire "ball carriers" according to in any position. In terms of assigning players, the team may find it easier to select a lot of passers and kickers in any position in the case that the team conducts more open play. Thus, the method applied in this study could assist rugby teams' management in acquiring and assigning players. However, the limitations of this study should be noted. Since this study analyzed to focused on ball-in-play, all players were analyzed with the same item. Therefore, set pieces such as scrums and line-outs have been excluded. Because some players may play a significant role in the performance of those set pieces, future research may also consider the performance of those set pieces. Additionally, since the subject players of this study were limited to players who had played for over 400 minutes, players under this time were not analyzed. In modern rugby, substitute players also play an important role as impact players. Therefore, in a future study, the range of subjects should be expanded by increasing the play time.

5. Conclusion

In rugby, the role of a player on a team is determined by their performance characteristics. When constructing a rugby team, categorization of players by their performance characteristics regardless of position may benefit the team, as variation among players is increasing. Additionally, since rugby is complicated and multifaceted (International Rugby, 2011), a comprehensive analysis of the performance characteristics of players should be conducted. Therefore, the aim of this study was to categorize players by their performance characteristics, irrespective of positions, using PCA and CA. Data from the JRTL 2015–2016 season, focusing on sample of 231 players and 16 items, were analyzed. Using PCA, the 16 items were reduced to four principal components (i.e., penetration, defense and competition, ball handling and kicking, and turnover) to produce a principal component score that enabled to an easier analysis of the players and clarified performance characteristics. The players' component scores were categorized into six clusters using CA according to their performance characteristics: playmaker, ball carrier, tackler and supporter, competitor, passer and kicker, and spoiler. Although categorization of the players by CA was achieved regardless of position, the performance characteristics of the players suggested that these were roughly categorized by the position. However, specific positions also had certain performance characteristics, and unique players were identified for certain positions.

We also identified a difference in roles between Japanese and foreign players in the JRTL. Since this method may enable team management to identify its own weak spots and highlight the strengths of players it analyzes, this study may contribute to the improving the success rate of scouting, assessment, acquisition, and the appointment of more suitable players by teams' management.

Acknowledgment

This study was supported by the joint research program of the Faculty of Health and Sports Science, Juntendo University, and Grants-in-Aid for Scientific Research (C) of Japan (No. 26350434). Data on the Japan Rugby Top League used in this study were provided by Data Stadium Inc.

References

Barros, R.M.L., Cunha, S.A., Magalhaes, W.J., and Guimarães, M.F. (2006). Representation and analysis of soccer players' actions using principal components. J. Hum. Mov. Stud., 51: 103–116.

Bompa, T., and Claro, F. (2009). Periodization in rugby. Aachen, Germany: Meyer & Meyer Sport Publishing.

Bret, H. (2012). Wallabies five-eighth Berrick Barnes is becoming perfect playmaker's foil in Test rugby. Fox Sports. https://www.foxsports.com.au/rugby /wallabies/wallabies-five-eighth-berrick-barnes-is-becoming-perfect-playmakers-foil-in-test-rugby/news-story/d7346f79f7197d3e3350082eb09121cb?sv=3d73d210868e40bd4b04fd3207c5325c. (accessed 2019-04-03).

Duthie, G.M., Pyne, D.B., Marsh, D.J., and Hooper, S.L. (2006). Sprint patterns in rugby union players during competition. J. Strength Cond. Res., 20: 208–214. doi:10.1519/R-16784.1.

Enrique, O., Diego, V., and Jose, M.P. (2009). Differences in game statistics between winning and losing rugby teams in the Six Nations Tournament. J. Sport Sci. Med., 8: 523–527.

Fernandez-Navarro, J., Fradua, L., Zubillaga, A., Ford, P.R., and McRobert, A.P. (2016). Attacking and defensive styles of play in soccer: Analysis of Spanish and English elite teams. J. Sports Sci., 34: 2195–2204. doi:10.1080/02640414.2016.1169309.

Greenwood, J. (1997). Total rugby. London: A & C Black.

Hefin, J. (2018, January 18). South Africa star Eben Etzebeth faces lengthy spell on sidelines after injuring shoulder. Pundit Arena. https://punditarena.com/rugby/hjo-nes/south-africa-star-eben-etzebeth-faces-lengthy-spell-on-sidelines-after-injuring-shoulder/. (accessed 2019-03-27).

International Rugby Board. (2011). International rugby board level 2 coaching developing rugby skills. https://webpoint.usarugby.org/files/Event_PDFs/1794_Event_OtherDoc3.pdf. (accessed 2017-04-10).

James, N., Mellalieu, S., and Jones, N. (2005). The development of position-specific performance indicators in professional rugby

union. J. Sports Sci., 23: 63–72. doi:10.1080/02640410410001730106.
Kiuchi, M., and Hirotsu, N. (2016). Analysis of game statistics on teams and players in Japan rugby top league. Proceedings of the 13th Australasian Conference on Mathematics and Computers in Sport. Melbourne, Australia: Melbourne University Australasian, 145–149.
Kraak, W.J., and Welman, K.E. (2014). Ruck-play as performance indicator during the 2010 Six Nations Championship. Int. J. Sports Sci. Coach, 9: 525–537. doi:10.1260/1747-9541.9.3.525.
Kramer, T.B., Huijgen, C.H., Elferink-Gemser, M.T., and Visscher, C. (2017). Prediction of tennis performance in junior elite tennis players. J. Sport Sci. Med., 16: 14–21.
Lim, E., Lay, B., Dawson, B., Wallman, K., and Anderson, S. (2009). Development of a player impact ranking matrix in super 14 rugby union. Int. J. Perf. Anal. Spor., 9: 354–367. doi:10.1080/24748668.2009.11868492.
MacDonald, H.F., and Rees, J.I. (1938). Rugger practice and tactics. London: Edward Arnold and Co.
Michele, K., Van, R., and Timothy, D.N. (2006). An analysis of the movements, both duration and field location, of 4 teams in the 2003 Rugby World Cup. Int. J. Perf. Anal. Spor., 6: 40-56.
Moura, F.A., Martins, L.E.B., and Cunha, S.A. (2014). Analysis of football game-related statistics using multivariate techniques. J. Sports Sci., 32: 1881–1887. doi:10.1080/02640414.2013.853130.
Naik, D.N., and Khattree, R. (1996). Revisiting Olympic track records: Some practical considerations in the principal component analysis. Am. Stat., 50: 140–144. doi:10.1080/00031305.1996.10474361.
Olds, T. (2001). The evaluation of physique in male rugby union players in the twentieth century. J. Sports Sci., 19: 253–262. doi:10.1080/026404101750158312.
Parsons, A., and Hughes, M.D. (2001). Performance profiles of male rugby union players. In M. D. Hughes & I. M. Franks (eds.), Pass. com. Fifth World Congress of Performance Analysis of Sport (pp. 129–136). Cardiff, Wales: Centre for Performance Analysis, University of Wales Institute.
Quarrie, K.L., and Wilson, B.D. (2000). Force production in the rugby union scrum. J. Sports Sci., 18: 237–246. doi:10.1080/026404100364974.
Quarrie, K.L., Hopkins, W.G., Anthony, M.J., and Gill, N.D. (2013). Positional demands of international rugby union: Evaluation of player actions and movements. J. Sci. Med. Sport, 16: 353–359. doi:10.1016/j.jsams.2012.08.005.
Villarejo, D., Palao, J.M., and Toro, E.O. (2013). Match profiles for establishing position specific rehabilitation for rugby union players. Int. J. Perf. Anal. Spor., 13: 567–571. doi:10.1080/24748668.2013.11868670.
Vivian, R., Mullen, R., and Hughes, M. (2001). Performance profiles at League, European Cup and international levels of male rugby union players, with specific reference to flankers, no. 8 and no. 9. In M. D. Hughes and I. M. Franks (eds.) Pass.com (pp. 137–143). Cardiff, Wales: UWIC, Computer Science and Sport III and Performance Analysis of Sport V.

Name:
Makoto Kiuchi

Affiliation:
Graduate School of Health and Sports Science, Juntendo University

Address:
1-1 Hiragagakuendai, Inzai, Chiba 270-1695 Japan

Brief Biographical History:
Juntendo University (BS, MS, PhD)

Main Works:
- Kiuchi, M., Washiya, K and Hayasaka, K. (2015).Comparison of the tackles in sevens rugby and 15-a-side rugby. Research Journal of Sports Performance, 7: 334-345.
- Kiuchi, M., Shimozono, H., Murakami, J., Hayasaka, K and Hirotsu, N. (2019).Identification of team characteristics in rugby union by using principal component analysis. Japanese Journal of Rugby Science 31(2): 40-45.
- Kiuchi, M. (2020).Characteristics of Oceania Regional Players' Performance in Rugby Union -Focusing on Fiji, Tonga and Samoa Players-. Global communication studies 9: 243-260.

Membership in Learned Societies:
- Japanese Society of Science and Football
- Japan Society of Physical Education, Health and Sport Sciences
- The Operations Research Society of Japan
- Japan Society of Rugby