Paper : Game Analysis

Identification of Soccer Defensive Transition Play Using Latent Class Analysis

Hiroki Matsuoka*, Kozue Ando* and Takahiko Nishijima*

*Faculty of Health and Sport Sciences, University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8574 Japan
matsuoka.hiroki.ga@u.tsukuba.ac.jp

[Received February 2, 2021; Accepted July 14, 2021]

The purpose of this study was to identify latent subgroups with respect to defensive transition plays from an achievement pattern data of the defensive transition in the professional soccer game. The research procedure comprised two steps, 1) an analysis of achievement criteria of the defensive transition items by a decision tree analysis, 2) an analysis of latent classes of defensive transitions by latent class analysis (LCA). The 158 defensive transitions in two games of the J-league in 2016 were used. An achievement pattern data of nine defensive transition items consisted of approach, press, and DFL control were used for LCA. Three latent classes, class 1: High achieved transition play (HA), class 3: Moderate achieved transition play (MA), and class 2: Low achieved transition play (LA) were identified. HA (54.25±5.35) showed significantly higher defensive transition skill (DTS) score compared with MA (42.61±5.10) and LA (27.85±3.24), and MA showed significantly higher defensive transition score than LA ($p < 0.05$). In addition, the relationship between LA and the opponent PA entry was found. These results indicated that LA allowed the opponent counter attack leading to failure of defensive transition play. In conclusion, this study successfully identifies latent classes of defensive transition plays by LCA.

Keywords: tracking data, classification of play, criterion-referenced evaluation

[Football Science Vol.18, 60-71, 2021]

1. Introduction

Moments of play in soccer are classified into four phases: offence, transition from offence to defence, defence, and transition from defence to offence. Recently, interest in the study of transition moments has been increasing, and the relationship between transitions to offence after ball gain and offensive outcomes has been analysed (Barreira et al., 2014; Hughes and Lovell, 2019). Barreira et al. (2014) reported that the transition to the offence with ball gain lead to offence success, and Hughes and Lovell (2019) reported that the number of goal scoring opportunities increased with the forward movement in the initial two actions after ball gain.

At the same time, the importance of defensive transitions has increased, and defensive tactics, gegenpressing, and counter-pressing have been developed due to the evolution of offensive transition tactics. Vogelbein et al. (2014) reported that they measured the defensive reaction time in soccer teams in the German Bundesliga and found that the successful teams in the final league table showed a shorter defensive reaction time compared with the unsuccessful teams. Similarly, Casal et al. (2016) analysed defensive transition play in the 2010 FIFA World Cup and investigated the defensive reaction time, an indicator related to the outcomes of the defensive transitions. Hobbs et al. (2018) reported that they successfully identified counter-pressing by classification of the feature of positional information of the soccer players derived from tracking data. However, it was hard to apply this knowledge to improving team performance, because the previous studies only analysed the time to ball recovery or the classification of play, and they failed to measure the behaviours of non-ball-holding players at the defensive transitions.

Methods for measuring soccer skill from soccer game data have been proposed (Suzuki and Nishijima, 2004; Tenga et al., 2009; Lago-Ballesteros et al., 2012; Matsuoka et al., 2020a; 2020b), and criterion-referenced measurement items of soccer defensive skill or soccer defensive transition skill utilizing tracking data and ball-touch data have been developed (Matsuoka et al., 2019; 2020a; 2020b). Soccer skill

is measured using the items of a binary scale with achievement criteria (achieved = 1, not achieved = 0), and an achievement score (ability=θ) of soccer skills is estimated from achievement pattern data using item response theory (IRT). Coaches can evaluate team performance in soccer games by using these scores.

It has been reported that analysis of the characteristics of the play in soccer games has been performed for use in the coaching process. Tenga et al. (2010a; 2010b) defined elaborative attack and counter attack as offence play and balanced defence and imbalanced defence as defence play, and they showed that using a counter attack against an imbalanced defence is an effective play. They suggested that the results shown in their study could be used for planning training or for team selection, taking into account the characteristics of play in the opposing team. Fernandez Navarro et al. (2016) analysed exploratory factor analysis with soccer game statistical data and found 6 factors and 12 styles of play. They reported direct play and possession play as offensive styles of play, and high pressing defence and low pressing defence as defensive styles of play. They suggested that identification of styles of play can be used for training planning for developing a team's playing style for competition or league games. Thus, it is generally considered that analysis and classification of the characteristics of play in soccer games based on achievement pattern data can be used for the coaching process, for example for training planning, and for the development of the team's style of play.

However, the defensive plays that were reported in previous studies are plays such as high or low pressing defence or balanced or imbalanced defence, and are intermingled with styles of play and quality of play. It is hard to determine exactly how many characteristics of play (as measured by soccer game data) exist, and hard to characterise those kinds of plays.

In addition, only soccer experts would recognise the characteristics of the play being latent in the samples of play as tacit knowledge, and this has not been analysed in terms of an achievement pattern of the measurement items. It is thus necessary to identify the latent class, with the characteristics of the play being latent in the samples, from achievement pattern data of play in soccer games by applying a statistical model.

Latent class analysis (LCA) is a statistical method for identifying latent sub-groups from a sample, including heterogeneous groups, and this method classifies individuals who have similar response patterns to the measurement items into the latent sub-group. In latent class analysis, it is assumed that the population includes heterogeneous groups, and hidden groups in a population are identified. It then estimates the probability of latent class membership and the probabilities of responses for each item.

Latent class analysis is often used in research in the social sciences. Lee et al. (2018) identified latent classes of physical activity patterns in elderly people, and analysed the relationship between latent classes and type 2 diabetes. Hue et al. (2011) identified latent classes, which were the possible risk of obesity from eating and the physical activity pattern of children, and they suggested that identification of latent classes made it possible to provide a suitable obesity prevention program for individual classes that had different eating and physical activity habits. Thus, it is possible to identify hidden groups in soccer play data by applying latent class analysis to achievement pattern data of plays in soccer games.

The purpose of this study was to identify latent subgroups with respect to defensive transition plays from achievement pattern data of the defensive transition in professional soccer games. The research procedure comprised two steps, 1) an analysis of the achievement criteria of the defensive transition items using decision tree analysis, and 2) an analysis of latent classes of defensive transitions using latent class analysis (LCA).

2. Methods

2.1. Data and Samples

Tracking data and ball-touch data for two games in the 2016 J-League Championship Finals (Rounds 1 and 2) were used. The scores (home team – away team) of the games were 0 – 1 in the 1st leg and 1 – 2 in the 2nd leg. Ball-touch data included match ID, half ID, time, possession team name, player name, ball action name (e.g., pass, shot, tackle), and ball XY positions. The data were collected by trained staff of a sports analytics company (Kato, 2016). Tracking data included player XY position recorded in every video frame (1/25 fps) using TRACAB (ChyronHego, NY, USA).

The 158 defensive transition plays at the offensive-midfield were collected. The start of defensive transitions was defined as being when the team lost ball possession (Vogelbein et al. 2014), and the defensive transition plays were defined as the defensive behaviour of the defensive transition team after ball possession was lost. This study used the defensive transition play corresponding to the first two opponent player ball actions as samples, because it was reported that the first two ball actions were crucial moments for increasing the opportunity for goal scoring (Hughes and Lovell, 2019). Moreover, defensive transition at the offensive-midfield was collected as samples (**Figure 1**), because the highest number of ball lost instances were recorded in this zone (Barreira et al., 2014). This study was approved by the Research Ethics Committee at the Faculty of Health and Sport Sciences, University of Tsukuba (Project No. 30-54).

2.2. Measurement of defensive transition play

The previous study constructed 16 criterion-referenced measurement items to measure defensive transition play based on the items of the defensive transition team and the opponent team in soccer games using item response theory (IRT) (Matsuoka et al., 2020b). This study used nine of the 16 items related to approaching (three items), pressing (two items), and the defence line (DFL) control (four items). Success criteria of defensive transition items were the first branch values to classify successful and unsuccessful defensive transition play, estimated by decision tree analysis (**Table 1**). Nine items of defensive transition play items were measured with a binary scale (Achieved =1, Not achieved = 0) based on the success criteria.

2.2.1. Approach items

To measure approaching the ball holder, three items were used: "Distance from the 1st defender to the ball", "Running speed of the 1st defender", and "Angle between the ball and the 1st defender."

Measurement of "Distance from the 1st defender to the ball" was the distance between the 1st defender and the ball. The success criteria were having a distance less than 9.88 meters. If the distance between the closest defender and the ball was less than 9.88 meters, then "Distance from the 1st defender to the ball" was measured as a success (**Figure 2**). "Distance from the 1st defender to the ball" less than 9.88 meters meant that a defender played within 9.88 meters of the ball.

Measurement of "Running speed of the 1st defender" was the running speed of the 1st defender at the opponent player keeping the ball. The success criteria were having a running speed of 14.26km/h or less. If the running speed of the 1st defender was 14.26km/h or less, then "Running speed of the 1st defender" was measured as a success. "Running speed of the 1st defender" of 14.26km/h or less meant that the 1st defender moved at a slow pace.

Measurement of "Angle between the ball and the 1st defender" used the angle between the line connecting the ball and the defence team goal, and the line connecting the ball and the 1st defender. The smaller the value of the angle in degrees, the better positioned the 1st defender was between the goal

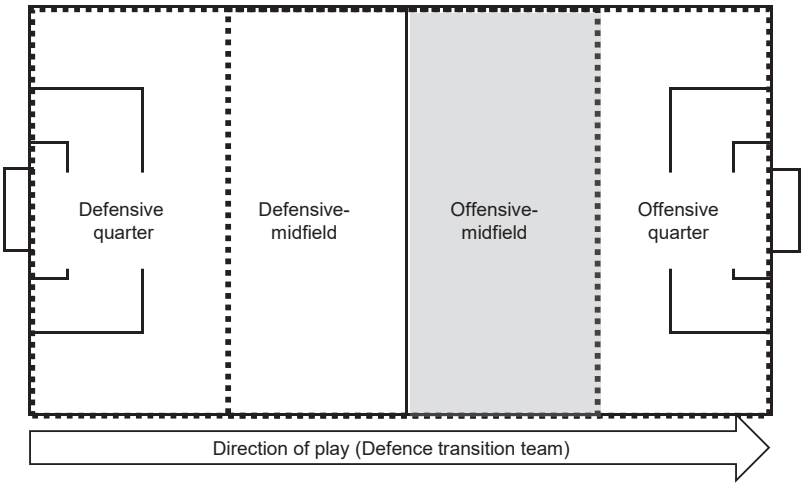
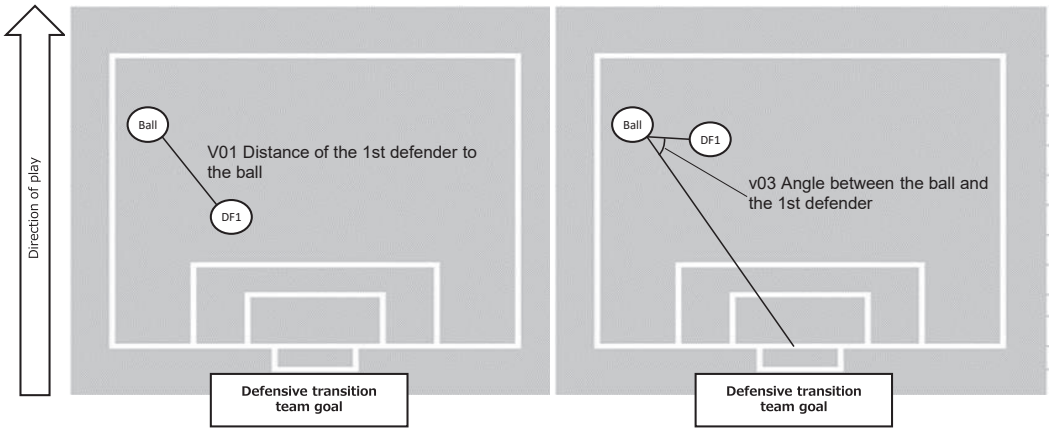


Figure 1 The offensive-midfield

Table 1 Measurement of defensive transition item and achievement criteria

No	Defensive transitions	Items	Measurement	Unit	Achievement criteria
v01	Approach	Distance of the 1st defender to the ball	Distance between the 1st defender and the ball	meter	< 9.88
v02		Running speed of the 1st defender	Running speed of the 1st defender at the opponent player keeping the ball	km/h	≤ 14.26
v03		Angle between the ball and the 1st defender	Angle between the line connecting the ball and the defence team goal, and the line connecting the ball and the 1st defender	degree	≤ 102.12
v04	Press	Number of defenders (5m)	Number of defenders positioned within 5m from the ball	people	> 0.00
v05		Marking distance to the opponent support player1	Distance from the nearest defender to the support player 1	meter	≤ 3.18
v06	DFL control	Distance moved by defence line	Distance of the vertical movement of the defence line from the current ball action to the next ball action	meter	> -3.59
v07		Speed of defence line	Distance moved by defence line divided by the time between the ball actions	km/h	> -6.41
v08		Width of defence line	Distance of the right sided defence line player and the left sited defence line player	meter	≤ 21.62
v09		Variability of defence line	Standard deviation of vertical distance of defence line players	meter	≤ 1.47

Note. The 1st defender = the closest player to the ball in the defensive transition team; Defenders = the players in the defensive transition team; The support player 1= the closest non-ball holding player to the ball in the opponent team.; The defence line = the closest defender to the defensive transition team goal; The defence line players = the defenders positioned within 6 meter from the defence line.



Note: Left diagram shows an example of successful v01 Distance of the 1st defender to the ball which successful criteria is distance of the 1st defender positioned closer than 9.88m from the ball. Right diagram shows an example of successful v03 Angle between the ball and the 1st defender which successful criteria is angle is narrower than 102.12 degree.

Figure 2 The achievement criteria of v01 Distance of the 1st defender to the ball and v03 Angle between the ball and the 1st defender

and the ball, while a larger value meant that the 1st defender had failed to position themselves properly between the goal and the ball. The success criteria were having an angle of 102.12 degrees or less. If the angle was 102.12 degrees or less, then "Angle between the ball and the 1st defender" was measured as a success (**Figure 2**).

2.2.2. Press items

To measure the pressing of the defensive transition team to the ball, two items were used, "Number of defenders (5m)", and "Marking distance to the opponent's support player 1."

Measurement of "Number of defenders (5m)" was based on the number of defenders positioned within a radius of 5 meters from the ball. The success criteria were having a number of defenders within 5 meters that was more than 1 player. If more than 1 player was positioned within five meters of the ball, then "Number of defenders (5m)" was measured as a success.

Measurement of "Marking distance to the

opponent's support player 1" was the distance from the nearest defender to the closest opposing player other than the opposing player with control of the ball. The success criteria were having a distance of 3.18 meters or less, and if marking distance to the opposing support player was 3.18 meters or less, then "Marking distance to the opponent's support player 1" was measured as a success.

2.2.3. DFL control items

To measure covering of the space behind the defence line (DFL) to prepare for the opposing counter-attack, four items were used, "Distance moved by defence line", "Speed of defence line", "Width of defence line", and "Variability of defence line."

Measurement of "Distance moved by defence line" was based on the distance of the movement of the defence line from the current ball action to the next ball action. If the DFL moved in the direction of the defence team's goal, then the distance moved was shown with a negative sign, and if the DFL moved in the direction of the opposing team's goal, then the distance moved was shown with a plus sign (**Figure 3**). For example, if the DFL moved 3.0 meters toward the defence team's goal, then the measurement showed -3.0 meters. The success criteria of "Distance moved by defence line" was movement that did not exceed -3.59 meters in the direction of the defence team goal. This meant that if the DFL did not move more than 3.59 meters toward the defence team goal between the current opposing ball action and the next opposing ball action, then "Distance moved by defence line" was measured as a success.

Measurement of "Speed of defence line" was based on the distance moved by the defence line divided by the time between the ball actions. Similar to the measurement of "Distance moved by defence line", for DFL speed with movement in the direction of the defence team's goal, the value was shown with a negative sign, and for DFL speed with movement in the direction of the opposing team's goal, the value was shown with a plus sign. The success criteria of "Speed of defence line" was more than -6.41 km/h. If "Speed of defence line" was -6.41km/h or more, it was measured as success.

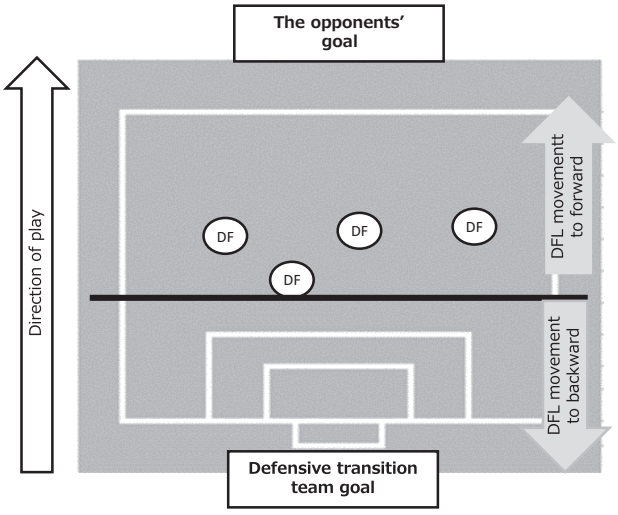
Measurement of "Width of defence line" was distance between the right-most defence line player and the left-most defence line player. The success criteria of "Width of defence line" was having a distance of 21.62 meters or less. If "Width of defence line" was 21.62 meters or less, it was measured as a success.

Measurement of "Variability of defence line" was based on the standard deviation of the positioning of the defence line players. If the defenders who were playing on the DFL were positioned in a straight line, then "Variability of defence line" showed a smaller value. However, if, for example, the centre back player was drawn out from their DFL, then the shape of DFL was no longer a straight line, and "Variability of defence line" showed a bigger value. The success criteria of "Variability of defence line" was having variability of 1.47 meters or less, and if "Variability of defence line" was 1.47 meters or less, it was measured as a success.

2.2.4. Defensive transition skill score (DTS score)

Defensive transition skill score was defined as the t-score of ability (θ) estimated by item response theory with a 2-parameter logistic model. Unidimensionality, item characteristics, and test characteristics were confirmed following the previous studies of Matsuoka et al. (2019; 2020a; 2020b).

The DTS score (ability [θ]) was estimated by the



Note: DFL is the closest defender to the defensive transition team goal. DFL movement is measured by the distance of the movement of the DFL between the opponent ball actions. IF direction of the movement of the DFL is forward (against the opponent's goal), the value of the movement is shown with positive, if the movement of it is backward (to the defensive transition goal), the value is shown with negative.

Figure 3 Measurement of Distance moved by defence line

maximum likelihood estimation method, and an equation of the item characteristic curve (ICC) in the 2PLM model is shown below

$$P(\theta) = \frac{1}{1+\exp(-Da(\theta-b))} \quad (1)$$

Where "P(θ)" is the probability that a defensive transition play (θ) is correct in a certain item, "D" is a scaling factor (1.7), "a" is the item discrimination parameter, and "b" is the item difficulty parameter. From the ICC, the proportion for which a certain defensive transition play(θ) constituted a successful P(θ) or unsuccessful Q(θ) response to a certain defensive transition item was calculated. The ability score (θ) that maximized the likelihood function L (θ|u) was estimated from achievement pattern data of defensive transition play (u), and the proportion of successful responses P(θ) and unsuccessful responses Q(θ). The equation of the likelihood function L (θ|u) is shown below.

$$L(\theta|u) = P(u|\theta) = \prod (P(\theta)^{u} Q(\theta)^{1-u}) \quad (2)$$

Where "L(θ|u)" is the likelihood function, "P(u|θ)" is the conditional probability of achievement pattern data (u) calculated from a certain defensive transition play(θ), "P(θ)" is the proportion of successful responses, and "Q(θ)" is the proportion of unsuccessful responses (1 – P(θ)).

2.3. Measurement of results of defensive transition play

The results of defensive transition plays were measured using the results of opposing penalty area (PA) entry. Hughes and Lovell (2019) reported that the initial two ball actions after ball-gain were crucial for a counter-attack. Thus, if the ball position at the third action after the defensive transition team lost the ball was in the PA of the defensive transition team, and the opposing team successfully played a counter-attack, then this was defined as a failure of the defensive transition plays. All other defensive transition plays were measured as successful defensive transition plays.

2.4. Statistical analysis

Decision tree analysis in which the results of defensive transition plays constituted a dependent variable and each defensive transition item was an independent variable was performed to find the success criteria of the defensive transition items. For the algorithm, the CART was used (Breiman et al., 1984). IBM SPSS ver23 was run.

Latent class analysis was completed to find the hidden classes latent in the defensive transition play data in soccer games. Latent class analysis (LCA) is a statistical method for identifying latent sub-groups from a sample that includes heterogeneous groups. To determine the number of latent classes, the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC) were used, and the model that showed the minimum value of these was selected. Parameters of the latent class were estimated using the EM algorithm with the poLCA package (Linzer and Lewis, 2011) of statistical software R version 3.6.1. To avoid local maxima, this study followed Linzer and Lewis (2011) and performed the analysis by changing the initial parameter value by calling poLCA 10 times for each model.

The Chi-square test and the Fisher-Freeman-Halton exact test were used to analyse the relationship between the characteristics of latent classes and the results of the opposing offence play. Moreover, a one-way analysis of variance (ANOVA) and post-hoc test using the Tukey-Kramer test were performed to compare the difference in the DTS score in the latent classes. For the analysis of the Chi-square test, the Fisher-Freeman-Halton exact test, one-way ANOVA, and post-hoc test, IBM SPSS version 23 statistical software was used. The significance level of hypothesis testing was set at 0.05.

3. Results

Figure 4 shows the BIC and the AIC of the models of two to five latent classes. The lowest BIA (1455.84) and AIC (1367.02) were the latent class model with three classes of defensive transition plays.

Table 2 shows class membership of the samples of the defensive transition plays, the proportion of samples meeting the success criteria, and the conditional response probability for each item.

The proportion of membership in class 1 was the highest (77.2%), and the conditional response probability for all items was relatively high in class 1. The conditional response probability of the items "Distance from the 1st defender to the ball", "Distance moved by defence line", and "Speed of defence line"

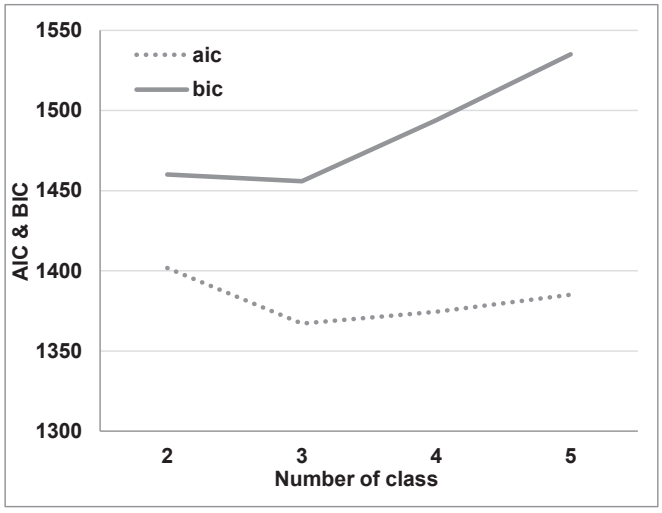


Figure 4 Goodness-of-fit for the latent class models

were 1.00 in this class. Moreover, the conditional response probability of the "Number of defenders (5m)" and "Marking distance to the opponent's support player 1" was higher than in the other two classes. Thus, latent class 1 was interpreted as the class of the High Achieved Transition Play (HA) that showed the highest achievement response of the defensive transition items.

The proportion of membership in class 2 was the lowest (10.8%), and the conditional response probability of the items "Distance moved by defence line" and "Speed of defence line" was 0.00 in this class. Although the conditional response probability of "Distance from the 1st defender to the ball" showed the second highest value in the three classes, the conditional response probability of "Running speed of the 1st defender" was the lowest (0.24) in the three classes. Thus, latent class 2 was interpreted as the class of the Low Achieved Transition Play (LA) that showed the lowest achievement response of the defensive transition items.

The proportion of membership in class 3 was 12.0%, and the conditional response probability for all items was relatively high, especially the conditional response probability of 0.80 in "Distance moved by defence line" and 1.00 in "Speed of defence line." Although the conditional response probability of "Distance from the 1st defender to the ball" was the lowest in the three classes, at 0.60, the conditional response probability of "Running speed of the 1st defender" was 0.80 and the conditional response probability of "Angle between the ball and the 1st defender" was the highest in the three classes. Thus, latent class 3 was interpreted as the class of the Moderate Achieved Transition Play (MA) that showed a moderate achievement response of the defensive transition items.

Table 3 shows descriptive statistics of the DTS scores in each latent class. **Figure 5** shows the difference in DTS score between the three latent classes of the defensive transition plays. The one-way ANOVA results showed that there was a significant effect (F (2, 155) = 218.739, $p < 0.05$) of the latent classes on the DTS score. In the post-hoc test, HA (54.25±5.35) showed significantly higher ($p < 0.05$) than LA (27.85±3.24) and MA (42.61±5.10), and MA showed significantly higher ($p < 0.05$) than LA. **Figure 6** shows examples of the player positions in each latent class.

Table 4 shows the relationship between the

Table 2 Proportion of total sample (N = 158) in the classes and probabilities of meeting criteria, and conditional probabilities on latent class

Items	Criteria	Proportion of samples meeting the criteria	Class 1 High achieved transition play n = 122 (77.2%)	Class 2 Low achieved transition play n = 17 (10.8%)	Class 3 Moderate achieved transition play n = 19 (12.0%)
Approach					
Distance of the 1st defender to the ball	< 9.88 m	0.66	1.00	0.88	0.60
Running speed of the 1st defender	≤14.26 km/h	0.94	0.83	0.24	0.80
Angle between the ball and the 1st defender	≤102.12 degree	0.37	0.65	0.47	0.90
Press					
Number of defenders (5m)	≥ 1 defender	0.73	0.86	0.65	0.00
Marking distance to the opponent support player1	≤ 3.18 m	0.76	0.44	0.12	0.14
DFL control					
Distance moved by defence line	> -3.59 m	0.34	1.00	0.00	0.80
Speed of defence line	> -6.41 km/h	0.52	1.00	0.00	1.00
Width of defence line	≤ 21.62 m	0.89	0.62	0.29	0.07
Variability of defence line	≤ 1.47 m	0.87	0.39	0.18	0.18

Table 3 Descriptive statistics of defensive transition skill score by latent class

	Class 1 High achieved transition play	Class 2 Low achieved transition play	Class 3 Moderate achieved transition play
n	122	17	19
mean	54.25	27.85	42.61
sd	5.35	3.24	5.10
max	66.00	32.60	49.50
min	41.00	20.60	33.60

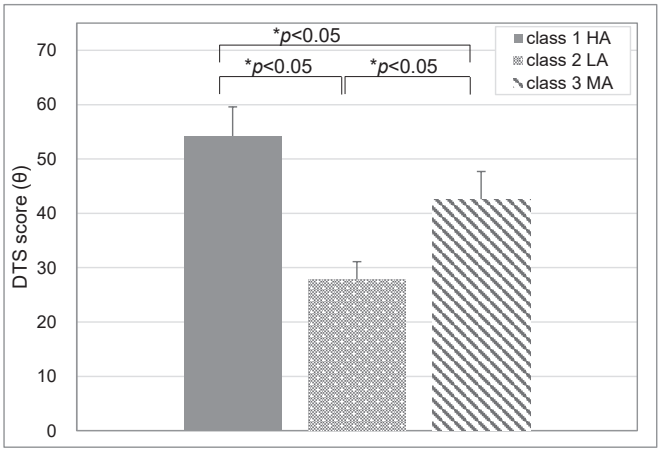
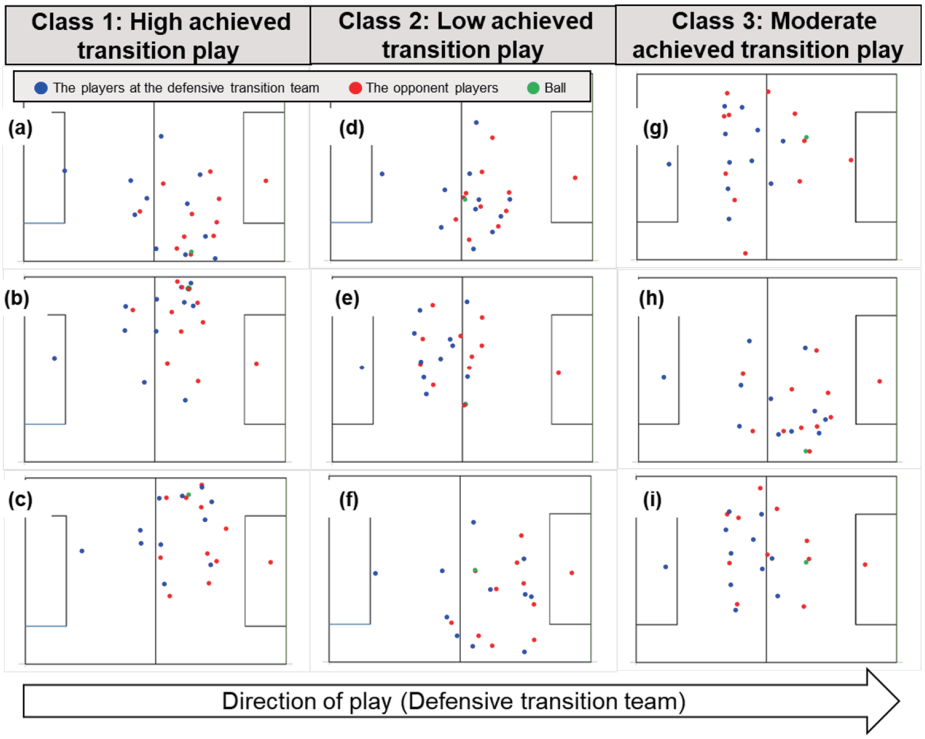


Figure 5 Comparison of defensive transition skill score by latent class

characteristics of the latent classes of the defensive transition plays and the result of the opposing offence play measured by PA entry. There was a significant relationship between the characteristics of the latent classes of the defensive transition plays and the result of the opposing offence play measured by PA entry ($p < 0.05$).

In the PA entry row in the table, the proportion of LA was significantly higher than that of MA or HA, and the proportion of MA was significantly higher than that of HA. In the Not PA entry row in the table, the proportion of HA was significantly higher than that of MA and LA, and the proportion of MA was



Note. This figure demonstrates the examples of player position in defensive transition play representing the three latent classes. The left three figures (a, b, c) show the player position in the high achieved transition play. Players in the defensive transition team (blue circle) positions around the ball (green circle), and pressing the ball holder with marking the opponents players (red circle). The right three figures (g, h, i) show the player position in the moderate achieved transition play. The player in the defensive transition team (blue circle) constructing the defensive block between goal and the ball holder, but the no defender is approaching to the ball holder. The middle three figure (d, e, f) show the player position in the low achieved transition play. There is no player in the defensive transition team approaching the ball holder, and the ball holder has plenty space to penetrate the opposing half.

Figure 6 Visualisation of players positions in three latent classes of the defensive transitions

Table 4 Relationship between latent class of defensive transition plays and opponent offence result

Opponent offence result	Latent class of defensive transition skill						
	Class 1: High achieved transition play		Class 2: Low achieved transition play		Class 3: Moderate achieved transition play		
	n	%	n	%	n	%	p^*
Not PA entry	122	82.4 *a*	8	5.4 *b*	18	12.1 *c*	0.00
PA entry	0	0.0 *c*	9	90.0 *a*	1	10.0 *b*	

Notes: ** the Fisher-Freeman-Halton exact test.*
Subscript characters (*a, b, c* in the Not PA entry row and in the PA entry row) shows the significant difference of the proportions between the cells in each row in the post-hoc test with the bonferroni-adjustment. The subscript character *a* shows the highest proportion in the row, the subscript character *b* shows the second highest proportions in the row, and the subscript character *c* shows the lowest proportions in the row.
For the Not PA entry row, it indicates that the proportion of class 1 is significantly higher than the proportion of class 2 and class 3, and the proportion of class 2 is significantly higher than the class 3.
For the PA entry row, it indicates that the proportion of class 2 is significantly higher than the proportion of class 1 and class 3, and the proportion of class 3 is significantly higher than the class 1.

significantly higher than that of LA.

4. Discussion

This study found three latent classes (HA, MA, LA) that were latent in the defensive transition plays through latent class analysis of achievement pattern data of the defensive transition plays.

The defensive transition play in the HA class has a high achievement response of the transition play items, and for the defensive transition play in this class it is assumed that the defensive transition team successfully presses the opposing team after the opposing team gains the ball. The items that showed the highest conditional response probability were "Distance from the 1st defender to the ball", "Number of defenders (5m)", "Marking distance to the opponent's support player 1", "Distance moved by defence line", and "Speed of defence line." This means that not only is the 1st defender approaching the ball, but also the defensive transition team are pressing the opposing non-ball-holding players. As **Figure 6 (a)**, **(b)**, **(c)** shows, it is confirmed that the players in the defensive transition team (blue) are approaching the ball (green) and are marking the opposing players (red) around the ball. Moreover, it is considered that the defensive transition play in this class is the play in which the team maintains their DFL, and are not pushed back toward their goal because of their stable DFL. With regard to the relationship between the latent classes and the result of the opposing offence play, the proportion of the opponent's PA entry is the lowest in the three classes, and this class shows the highest DTS score.

The defensive transition play in the MA class has a moderate achievement response of the transition play items. The conditional response probability of the "Angle between the ball and the 1st defender" is high, so for the defensive transition play in this class it is considered that the defensive transition team successfully defends between their goal and the opposing ball holder, although the defensive transition team fail to press the opponents. As **Figure 6 (g)**, **(h)**, **(i)** shows, it is confirmed that the players in the defensive transition team (blue) are organising a defensive block between the defence team goal and the ball (green). This class shows the second highest DTS score.

The defensive transition play in the LA class has a low achievement response of the transition play items, and for the defensive transition play in this class it is considered that the defensive transition team failed to control the opposing team offence and allowed the opposing team's counter-attack. The conditional response probability of the "Distance moved by defence line" and "Speed of defence line" was 0.0 in this class, indicating that the defensive transition team is forced to drop back to their own goal. In addition, the results show a high conditional response probability in "Distance from the 1st defender to the ball" and low conditional response probability in "Running speed of the 1st defender". This shows that, although the 1st defender is approaching the ball, the play in this class does not satisfy the success criteria of the "Running speed of the 1st defender",

which is running speed of 14.26 km/h or less, and the 1st defender can be sprinting to approach the ball. Moreover, the low conditional response probability of the "Angle between the ball and the 1st defender" shows that the defensive transition team failed to defend the area between their goal and the opposing ball. Thus, for the defensive transition play in LA, it is considered that the approaching of the 1st defender is late, and as a result, the 1st defender has to chase the ball-holder from behind.

As **Figure 6 (d)**, **(e)**, **(f)** shows, it is confirmed that the player in the defensive transition team (blue) failed to press the opposing team, and there is plenty of space for opposing team players (red) and the ball (green) that would allow the opposing team to easily move the ball to the forward direction. This class shows the highest DTS score.

The latent classes like the HA class or the MA class of the defensive transition play investigated in this study are considered as a balanced defence, such as was reported in previous studies. Tenga et al. (2010a; 2010b) reported that the counter-attack against a balanced defence was not effective play. The HA class is where the defensive transition plays in such a way that the 1st defender approaches the ball and the DFL remains stable. The MA class is where the defensive transition plays in such a way that the 1st defender fails to approach the ball but the team defends the area between the defensive transition team's goal and the ball with a stable DFL. What is particularly significant is that the HA class shows significantly fewer opposing PA entry instances, and can thus be considered a balanced defence. By contrast, the defensive transition play in the LA class represents lower achievement of play at the defensive transition moment, which is a crucial moment for the opposing counter-attack after ball-loss, and the plays in this class allow the opposing team PA entry. Hughes and Lovell (2019) analysed the features of offensive transition plays in 29 games in the knockout stages in the 2014-2015 UEFA Champions League, and they reported that forward movement such as a long pass or dribbling in the initial two actions after ball gain is crucial for increasing goal scoring opportunities. Moreover, Tenga et al. (2010a; 2010b) showed that the counter-attack against imbalanced defence was effective. The LA class investigated in this study is considered as an imbalanced defence, as the plays in this class show significantly higher opposing PA entry.

It is expected that the objective classification of these three latent classes from soccer game data using latent class analysis will contribute toward the development of the soccer performance analysis and coaching process. Although the nine criterion-referenced items in the previous study can be used to measure the achievement of the defensive transition (Matsuoka et al., 2020b), it is not possible to classify the characteristics of plays that are latent in the samples. The use of latent class analysis allows identification of the groups of latent plays in samples from qualitative data based on a statistical model. By comparison with cluster analysis that classifies the quantitative data based on the distances among them, latent class analysis determines the number of latent classes with statistical criteria, and it is possible to estimate the class membership of the samples and the conditional response probability of the items.

The class membership and conditional response probability could be useful for practical fields in soccer. Estimation of the class membership of the samples solves the problem of the selection of video clips of games being dependent on subjective selection by coaches and soccer analysts. O'Donoghue (2006) insisted that the quality of the events (positive play or negative play) selected from a game video depends on the subjective views of analysts, and the quality of the events will differ depending on who is responsible for selecting the events. It is possible to select plays or events without the subjective decision of the analysts by using latent class analysis for achievement pattern data from tracking data in soccer games. Moreover, conditional response probability shows the achievement rate of each item in a particular latent class, and it shows the important items in each class. Coaches and analysts can aggregate the latent classes into which their team's defensive transition play is classified and check the latent classes with the largest number of aggregates. The characteristics of the defensive transition play and the items that need improvement will then be clarified.

The previous study developed nine criterion-referenced measurement items to measure defensive transition based on the item characteristics shown on ICC that are not dependent on the samples, using item response theory (Matsuoka et al., 2020b). By achieving the criterion-referenced items step-by-step from easy ones to difficult ones, sports skills can be learned (Ono et al. 2015), so it is possible to use this

for training. The use of latent class in the latent class analysis adds a provision of an improvement plan for sports skill for each latent class based on nine criterion-referenced measurement items.

For example, the defensive transition plays belonging to latent class 1 (HA) show lower conditional response probabilities in "Width of defence line" and "Variability of defence line." There is a possible problem where a defence player on the DFL might be drawn out from their DFL or a full back on the opposite side stays at a problematic position although the defender is approaching the ball at the defensive transition moment. It is possible to provide an improvement plan to modify a full-back (blue) on the opposite side squeezing the defence inside shown in **Figure 6 (a)** and **(b)**, or where a full-back (blue) needs to position himself where he can cover the space behind the DFL as shown in **Figure 6 (c)**.

By contrast, with latent class 2 (LA), the conditional response probability of four items in DFL control is low. As can be seen in **Figure 6**, the cause of this situation is that there is plenty of space around the ball and the 1st defender is late in approaching the ball. It may be that the defensive transition team lost the ball in the worst situation, but it is possible to provide an improvement plan, with the decision as to who approaches the ball being made by predicting the impending defensive transition before ball-loss. Thus, it is now possible to provide an improvement plan based on the characteristics of plays by latent class analysis with a training plan that utilizes criterion-referenced measurement, and it can be anticipated that this may lead to data-driven coaching.

To generalise the conclusions regarding estimation of latent classes hidden in the sample, it is necessary to consider the limitations imposed on the research by concerns relating to the sample and the measurement items. This study sampled the defensive transition play at the offensive-midfield from the 2016 J-League Championship Finals (Rounds 1 and 2). To generalise the conclusion, it is necessary that the samples need to be selected randomly from all defensive transition plays in all soccer games, but the sample in this study was not randomly selected from the population. In addition, the measurement items in this study are achievement pattern data measured by nine items of approaching, pressing, and covering of DFL in the defensive transition play, following the previous study (Matsuoka et al., 2020b). To generalise the conclusion, measurement items need to represent all populations of measurement items, but the items in this study only satisfy the necessary condition and do not satisfy sufficient conditions. Therefore, it is necessary to consider whether it would be possible to generalize the findings in this study within the limitations imposed by research concerns regarding the samples and the measurement items.

Finally, it is considered that tactical play in soccer games will continue to develop in the future, and it has been noted that it is necessary to take interaction with the opponent into account when constructing measurement items (Tenga et al., 2010a; 2010b; Fernadez-Navarro, 2016: Lago-Ballesteros et al., 2012). Future research needs to focus on the analysis of latent classes of play in achievement pattern data measured by offensive and defensive items in soccer games, and analysis of the causal relationship of the opposing team items that influence the latent class of play.

5. Conclusion

The purpose of this study was to identify latent subgroups with respect to defensive transition plays from achievement pattern data of the defensive transition in professional soccer games. For the purpose, this study completed 1) an analysis of achievement criteria of the defensive transition items using a decision tree analysis, 2) an analysis of latent classes of defensive transitions using latent class analysis (LCA), and reached the following conclusions.

1) The latent sub-groups hidden in the defensive transition play in soccer games are classes of high achieved transition play (HA), moderate achieved transition play (MA), and low achieved transition play (LA).
2) HA is the defensive transition play that has high conditional response probability, MA is the defensive transition play that has moderate conditional response probability, and LA is the defensive transition play that has low conditional response probability.

Acknowledgments

This research was a collaborative research project based on an industry-academia collaborative research agreement between University of Tsukuba and

DataStadium Inc. We appreciate to DataStadium Inc. as they permit us to use soccer game tracking data and ball touch data. This work was supported by JSPS KAKENHI Grant Number 21K19697.

References

Barreira, D., Garganta, J., Guimarães, P., Machado, J., and Anguera, M. T. (2014). Ball recovery patterns as a performance indicator in elite soccer. P. I. Mech. Eng. P-J. Spo., 228, 61-72.

Breiman, L., Friedman, H, J., Olshen, R, A., and Stone, C, J. (1984). Classification and regression trees. Wadsworth.

Casal, C. A., Andujar, M. Á., Losada, J. L., Ardá, T., and Maneiro, R. (2016). Identification of defensive performance factors in the 2010 FIFA World Cup South Africa. Sports, 4, 54.

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.

Hobbs, J., Power, P., Sha, L., and Lucey, P. (2018). Quantifying the value of transitions in soccer via spatiotemporal trajectory clustering. In proceedings of the 12th annual MIT Sloan sports analytics conference, Boston, MA, USA.

Hughes, M., and Lovell, T. (2019). Transition to attack in elite soccer. J. Hum. Sport Exerc., 14, 236-253.

Huh, J., Riggs, N. R., Spruijt-Metz, D., Chou, C. P., Huang, Z., and Pentz, M. (2011). Identifying patterns of eating and physical activity in children: a latent class analysis of obesity risk. Obesity, 19, 652-658.

Kato, K. (2016). Soccer ni okeru data bunseki to team kyouka [Data analysis and team reinforcement in soccer]. IEICE communications society magazine, 10, 29-34. (in Japanese).

Lago-Ballesteros, J., Lago-Peñas, C., and Rey, E. (2012). The effect of playing tactics and situational variables on achieving score-box possessions in a professional soccer team. J. Sports Sci., 30, 1455-1461.

Lee, P.G., Ha, J., Blaum, C.S., and Gretebeck, L. (2018) Patterns of physical activity in sedentary older individuals with type 2 diabetes. Clin. Diabetes Endocrinol., 4, 7.

Linzer, D. A., and Lewis, J. (2011). poLCA: Polytomous variable latent class analysis version 1. 4. J. Stat. Softw., 42, 1-29.

Matsuoka, H., Tahara, Y., Ando, K., and Nishijima, T. (2020a). Development of criterion-referenced measurement items for soccer defensive tactical play from tracking data. Football Science. 17, 29-40.

Matsuoka, H., Tahara, Y., Ando, K., and Nishijima, T. (2020b). Tracking data karano senjutsu akushon koumoku no kaihatsu [Development of tactical action items from tracking data]. Research on Sports Data Analysis: Theory, Methodology, and Applications. 7, 71-76. (in Japanese).

Matsuoka, H., Tahara, Y., Ando, K., and Nishijima, T. (2019). Shinsou gakushu no tamenno soccer no tracking data karano transition play no sokutei [Measurement of transition play from tracking data of soccer games for deep learning]. JSSF 17th congress program. 38. (in Japanese).

Morrow Jr. J. R., Jackson, A. W., Disch, J. G., and Mood,D. P. (1995). Measurement and Evaluation in Human Performance. Human Kinetics.

O'Donoghue, P. (2006). The use of feedback videos in sport. Int. J. Perform. Anal. Sport., 6, 1-14.

Ono, M., Jo, T., Oyama, B, K., and Nishijima, T. (2015). Enban nage dousa ginou no hyouka kijun [Achievement evaluation standard for men's discus throw]. Japan Journal of Test and Measurement in Health and Physical Education, 14, 1-10. (in Japanese).

Popham, W. J., and Husek, T. R. (1969). Implications of criterion-referenced measurement. J. Educ. Meas., 6, 1-9.

Suzuki, K., and Nishijima, T. (2004). Validity of a soccer defending skill scale (SDSS) using game performances. Int. J. Sport. Stud. Hlth., 2, 34-49.

Tenga, A., Holme, I., Ronglan, L. T., and Bahr, R. (2010a). Effect of playing tactics on goal scoring in Norwegian professional soccer. J. Sports Sci., 28, 237-244.

Tenga, A., Holme, I., Ronglan, L. T., and Bahr, R. (2010b). Effect of playing tactics on achieving score-box possessions in a random series of team possessions from Norwegian professional soccer matches. J. Sports Sci., 28, 245-255.

Tenga, A., Kanstad, D., Ronglan, L,T., and Bahr, R. (2009). Developing a new method for team match performance analysis in professional soccer and testing its reliability. Int. J. Perform. Anal. Sport., 9, 8-25.

Vogelbein, M., Nopp, S., and Hökelmann, A. (2014). Defensive transition in soccer–are prompt possession regains a measure of success? A quantitative analysis of German Fußball-Bundesliga 2010/2011. J. Sports Sci., 32, 1076-1083.

Name:
Hiroki Matsuoka

Affiliation:
Faculty of Health and Sport Sciences, University of Tsukuba

Address:
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8574 Japan

Brief Biographical History:

2015-2017 Master's Program in Health and Sport Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba

2017-2020 Doctoral Program in Physical Education, Health and Sport Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba

2020-present Faculty of Health and Sport Sciences, University ofTsukuba

Main Works:

- Matsuoka, H., Tahara, Y., Ando, K., and Nishijima, T. (2020). Development of criterion-referenced measurement items for soccer defensive tactical play from tracking data. Football Science, 17: 29-40.
- Matsuoka, H., Tahara, Y., Ando, K., and Nishijima, T. (2020). Development of defence and offence play items for deep learning model of offence play analysis in soccer game. Football Science, 17: 69-85.

Membership in Learned Societies:

- Japan Society of Physical Education, Health and Sport Sciences
- Japanese Society of Science and Football