Journal of Physical Education and Sport ® (JPES), Vol. 23 (issue 7), Art 200, pp. 1631 - 1641, July 2023
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

A Multivariate and cluster analysis of diverse playing styles across European Football Leagues

SPYRIDON PLAKIAS[1*], SERAFEIM MOUSTAKIDIS[2], MICHALIS MITROTASIOS[3], CHRISTOS KOKKOTIS[4], THEMISTOKLIS TSATALAS[5], MARINA PAPALEXI[6], GIANNIS GIAKAS[7], DIMITRIOS TSAOPOULOS[8]

[1,5,7]Department of Physical Education and Sport Science, University of Thessaly, 382 21, Trikala, GREECE;
[2]AIDEAS OÜ, Narva mnt 5, 10117, ESTONIA;
[3]Department of Physical Education and Sport Science, National & Kapodistrian University of Athens, GREECE
[4]Department of Physical Education and Sport Science, Democritus University of Thrace, 69100, Komotini, GREECE;
[8]Department of Operations, Technology, Events and Hospitality Management, Manchester Metropolitan University| Oxford Road|Manchester|M15 6BH, UK
[8]Institute for Bio-Economy & Agri-Technology, Center for Research and Technology Hellas, 60361 Volos, GREECE;

Published online: July 31, 2023
(Accepted for publication July 15, 2023)
DOI:10.7752/jpes.2023.07200

Abstract
Performance analysis is a valuable tool for team coaches and has been the subject of extensive study in international research. A significant portion of the scientific literature in the field of football has been devoted to studying playing styles in recent years. The identification of playing styles is now regarded as crucial for conducting an efficient performance analysis. This study aimed to explore the variances in playing styles among eleven distinct European domestic football leagues. A comprehensive sample of 2996 matches, accounting for 5992 observations, was scrutinized. Nineteen latent variables, representing thirty-eight different game styles previously identified in sports science literature, served as the basis for this investigation. Multivariate analysis of variance (MANOVA) revealed significant differences across countries in ten out of nineteen variables. The variables with the highest effect sizes (partial η^2) were transition game, effective game, and defending aggressively, implying that these factors contributed to the most substantial differences among countries. To visualize these disparities, the t-distributed stochastic neighbor embedding (t-SNE) method was employed. Subsequently, k-means clustering was applied to the t-SNE results, grouping the eleven participating countries into five distinct clusters. A unique playing style was discerned in the Scottish league (Cluster 4), setting it apart from all other leagues. Other clusters included Austria, Belgium, and Switzerland (Cluster 1); Spain, Turkey, and Croatia (Cluster 2); Greece and Italy (Cluster 3); and Germany and England (Cluster 5). The findings offer valuable insights for coaches, managers, scouts, and sporting directors, potentially guiding the development of effective game styles and enhancing recruitment strategies for both players and coaches.
Keywords: football, tactics, performance analysis, differences between countries, game styles

Introduction

Performance analysis in football is a rapidly evolving domain, witnessing substantial research in recent years (Kirkendall, 2020; Plakias, Kokkotis, Tsaopoulos, et al., 2023). Factors such as tactics, technique, physical condition, personality and psychology of players can affect the playing styles and performance of teams (Plakias et al., 2022). In recent years, playing styles have been extensively studied by researchers (Plakias, Moustakidis, et al., 2023).

In the research of Sarmento et al. (2013) football experts mentioned in their interviews that distinct ways of playing in the different leagues have been created through historical and cultural factors. They also referred to differences in the philosophy of the coaches and the players' characteristics. Numerous studies dealing with football issues tried to find differences between countries in various areas. In particular, these papers sought to find differences between countries in issues such as: the development activities (early diversification, early specialization, and early engagement) (Ford et al., 2012), the efficiency of soccer betting markets (Oikonomidis et al., 2015), the magnitude of the heading incidence in children's and youth' football (Beaudouin et al., 2020), the grass-root coaches knowledge, understanding, attitude and confidence to deliver injury prevention training (De Ste Croix et al., 2020), the effect of gender equality (Bredtmann et al., 2014), the relative age effect (Doyle & Bottomley, 2018), the proposed soccer result and stock market linkage (Heuvel, 2016), and the players' perceptions of football turf and natural grass surface properties (Roberts et al., 2014).
In performance analysis, several authors have drawn comparisons between national leagues of different countries, using isolated performance indicators. Particularly: i) Clemente et al. (2019) used variables for

Corresponding Author: SPYRIDON PLAKIAS, E-mail: spyros_plakias@yahoo.gr

physical performance such as total distance, sprint distance, number of sprints etc, ii) Oberstone (2011), Yi et al. (2019), Alberti et al. (2013), Sapp et al. (2018) and García-Aliaga et al. (2022) used technical-tactical variables such as shots, passing accuracy, interceptions, fouls, yellow cards etc, iii) Dellal et al. (2011) used both technical-tactical and physical performance variables at the same research, and iv) Pollard and Pollard (2005), Pollard (2006), Pollard and Gómez (2009), Pollard and Gómez (2014) and Hill and Van Yperen (2021) used the effect of a contextual variable (home/away location). The disadvantage of the isolated variables used in all previous studies is that they cannot by themselves explain the interactive and complex systems' aspects of football (Hughes & Bartlett, 2002; Sarmento et al., 2014).

According to Gómez et al. (2018) and Lopez-Valenciano et al. (2022) the combination of multiple performance indicators has provided a more comprehensive picture of the teams' tactical patterns or styles of play, which may interpret better the complexity of the soccer game. Several studies have focused on playing styles in separate countries. For example, Brooks et al. (2016), Castellano and Pic (2019), and Lopez-Valenciano et al. (2022) studied the Spanish league, Lago-Peñas et al. (2017), Zhou et al. (2021), and (Ruan et al., 2022) used data from Chinese league, while others dealt with English (Gollan et al., 2018), Greek (Gómez et al., 2018) or Japanese (Narizuka & Yamazaki, 2019) league.

Furthermore, some papers studied more than one league, but without comparisons between them (Bekkers & Dabadghao, 2019; Fernandez-Navarro et al., 2016; Gyarmati et al., 2014). To the best of our knowledge, there are only two studies that tried to find differences between countries, in playing styles. However, both Mitrotasios et al. (2019) who compared England, Italy, Spain, and Germany as well as Gonzalez-Rodenas et al. (2020) who compared Spain to England, dealt only with ball-possession phase styles of play (Plakias, Moustakidis, et al., 2023).

Thus, the purpose of this study is to fill this gap by investigating differences in playing styles, encompassing all phases of the game, across eleven different European countries. This study will significantly contribute to the existing literature on performance analysis in football, presenting a comprehensive view of national playing styles.

Methods

Sample: The sample included all matches played in the 1st league division of 11 European countries in the 2021-2022 season. The data was collected using Instatscout platform (https://football.instatscout.com/). For 8 matches, Instatscout had no data or the data were incomplete. According to previous research (Casal et al., 2019; Castellano & Echeazarra, 2019; Gómez et al., 2018) the reliability of the indicators obtained by Instatscout is very high (K values 0.90 to 0.98). Only games from the regular season are featured (without play-offs and play-outs). The number of matches played and matches included from each competition is shown in Table 1. For each match, there were separate observations for both teams. Therefore, the sample included a total of 5992 valid observations.

Table 1. Matches played in each country and matches included in the study

COUNTRY	MATCHES PLAYED	MATCHES INCLUDED
AUSTRIA	132	132
BELGIUM	306	303
CROATIA	180	178
ENGLAND	380	380
GERMANY	306	306
GREECE	182	182
ITALY	380	377
SCOTLAND	198	198
SPAIN	380	380
SWITZERLAND	180	180
TURKYIE	380	380
TOTAL	**3004**	**2996**

Variables-Procedure: The study used factor scores from 19 components extracted from Factor Analysis and Principal Components Analysis that applied in the research of Plakias, Kokkotis, Moustakidis, et al. (2023). Table 2 shows the names given in each latent variable and the names of playing styles produced depending on the sign (positive or negative) of factor scores in each observation.

Sixteen out of 19 latent variables concern a competing team (playing styles of teams), while the rest three (2, 5, 13) relate to the game (styles for the game). The latter resulted from the combination of the behavior of both teams participating in a match.

Table 2. Latent variables and playing styles produced depending on the sign (positive or negative) of factor scores in each observation.

FACTOR	LATENT VARIABLE	POSITIVE SCORES	NEGATIVE SCORES
1	Elaboration of build up phase	Possession style	Direct style
2	Transition game	Many transitions	Few transitions
3	Attacking transition	Counterattack	Positional attack
4	Defensive transition	Opponent's counterattack	Opponent's positional attack
5	Aerial game	Game in the air	Game on the ground
6	Type of attack	Set pieces attack	Open play attack
7	Crossing	Many crosses	Few crosses
8	Type of opponent's attack	Open play defense	Set pieces defense
9	Defensive blocks	Mid block	Low block
10	Press	High press	Deep press
11	Individual defending actions	Many individual defending actions	Few individual defending actions
12	Width of creative phase	Center attack	Wide attack
13	Effective game	More interruptions and duels	More possession from one or the other team
14	Individual attacking actions	Many individual attacking actions	Few individual attacking actions
15	Tendency to create final attempts	Little possession required to generate final attempts (strong tendency)	High possession required to generate final attempts (low tedency)
16	Passing tempo	Low passing tempo	High passing tempo
17	Defending aggressively	Low defensive aggressiveness	High defensive aggressiveness
18	Attacking aggressively	High attacking aggressiveness	Low attacking aggressiveness
19	Offside trap	More frequent adoption of the offside trap	Less frequent adoption of the offside trap

Statistical analysis: Multivariate analysis of variance (MANOVA) was applied after validating normality and homogeneity assumptions. MANOVA was chosen because it reduces Type I error inflation compared with analysis of variance (ANOVA). Additionally, MANOVA can detect statistical differences that many ANOVAs cannot (Clemente et al., 2016).

The assumption of normality of all latent variables in each country was examined using Kolmogorov–Smirnov tests ($p< 0.05$).

The homogeneity of variances was examined with the Box's M test and the Levene's test (Beyene & Bekele, 2016). Bonferroni post hoc multiple comparisons were used to determine where the difference laid. Partial η^2 was used as a measure of effect size, because using large samples (as in this study) creates the issue of guaranteed statistical significance.

Thus, it demands reporting the practical significance by using effect size measures (Khalilzadeh & Tasci, 2017). The following scale was used to classify the effect size (partial η^2) of the test: very small 0-0,01; small 0,011-0,060; moderate 0,061-0,140; large >0,140 (Cohen, 2013). The Statistical Package for Social Science (SPSS, version 25) was used for all analyses.

The t-distributed Stochastic Neighbor Embedding (t-SNE) method was then applied to the mean values of the eleven countries on the 19 latent variables. T-SNE is a machine learning method that uses a dimensionality reduction algorithm allowing the visualization of large dataset into a 2D plot (Ramírez-Arroyo et al., 2022).

Furthermore, k-means Cluster analysis was applied to the values obtained from t-SNE to group countries based on the total of 38 styles adopted in them. K-means Clustering is a tool for finding groups or clusters of observations in multivariate data (Kodinariya & Makwana, 2013; Martínez et al., 2023). K-means was performed with SPSS, while K = 5 was chosen by the authors.

Results

Manova: From the descriptive statistics we obtained Table 3 values for the means and standard deviations of the 11 countries on the 19 latent variables. Table 4 shows the ranking of the 11 countries on the 19 latent variables.

Table 3. Means and standard deviations of the 11 countries on the 19 latent variables.

LATENT VARIABLE	METRICS	COUNTRIES										
		AUSTRIA	BELGIUM	CROATIA	ENGLAND	GERMANY	GREECE	ITALY	SCOTLAND	SPAIN	SWISS	TURKYIE
ELABORATION OF BUILD UP PHASE	MEAN	-0,35	-0,06	-0,07	0,07	0,08	0,10	0,15	-0,41	-0,04	-0,14	0,18
	SD	0,92	0,88	0,88	1,15	0,97	0,93	1,00	1,26	1,05	0,75	0,86
TRANSITION GAME	MEAN	0,88	0,57	0,42	-0,44	0,21	-0,34	-0,40	0,29	-0,17	0,33	-0,27
	SD	0,91	0,88	0,94	0,86	0,89	1,14	0,99	0,93	0,89	0,94	0,84
ATTACKING TRANSITION	MEAN	0,18	0,14	-0,04	-0,07	0,14	0,11	-0,12	-0,40	-0,15	0,05	0,20
	SD	1,03	1,02	0,96	1,00	1,01	0,97	0,93	0,99	0,98	0,98	1,01
DEFENSIVE TRANSITION	MEAN	0,25	0,14	-0,01	-0,04	0,16	-0,03	-0,12	-0,29	-0,15	0,09	0,11
	SD	1,01	1,00	0,93	0,98	1,01	0,95	1,09	0,97	0,94	0,99	0,97
AERIAL GAME	MEAN	0,20	-0,10	-0,06	0,04	0,18	-0,20	0,04	0,80	-0,02	-0,50	-0,25
	SD	0,98	1,04	0,95	0,93	1,03	0,84	1,07	1,04	0,96	0,87	0,82
TYPE OF ATTACK	MEAN	0,15	-0,08	-0,15	0,10	0,19	-0,04	-0,09	0,20	0,02	-0,14	-0,12
	SD	1,13	1,03	1,03	1,00	1,01	0,95	0,92	1,08	1,02	0,92	0,93
CROSSING	MEAN	-0,19	-0,09	0,09	-0,01	-0,29	-0,06	-0,13	0,33	0,17	0,03	0,14
	SD	0,79	0,90	1,00	1,02	0,87	0,99	1,01	1,08	1,06	0,89	1,06
TYPE OF OPPONENT ATTACK	MEAN	-0,12	0,06	0,08	-0,09	-0,17	0,07	0,08	-0,17	-0,03	0,12	0,12
	SD	1,08	0,99	1,02	1,00	0,99	0,94	1,03	1,06	1,02	0,95	0,91
DEFENSIVE BLOCKS	MEAN	-0,06	-0,05	0,24	-0,11	0,15	0,20	0,19	-0,25	0,03	-0,19	-0,16
	SD	0,99	0,95	0,98	0,96	1,02	0,97	1,04	1,07	0,95	0,92	1,00
PRESS	MEAN	-0,07	0,08	-0,12	0,11	0,08	-0,28	0,15	-0,15	-0,09	0,04	-0,01
	SD	0,89	1,03	0,98	1,06	0,98	0,90	1,05	0,88	1,01	0,93	1,00
INDIVIDUAL DEFENSIVE ACTIONS	MEAN	0,12	0,19	0,02	-0,03	0,02	0,07	-0,06	-0,18	-0,12	0,17	-0,03
	SD	1,04	1,12	1,04	1,03	0,96	0,95	0,98	0,90	0,93	1,02	0,97
WIDTH OF CREATIVE PHASE	MEAN	0,07	-0,20	0,13	-0,10	0,21	-0,06	0,07	-0,12	-0,04	0,06	0,03
	SD	0,92	0,97	0,94	1,01	1,07	0,97	0,99	0,95	0,96	1,03	1,04
EFFECTIVE TIME	MEAN	0,55	-0,13	-0,14	-0,01	-0,29	0,42	-0,20	0,34	0,09	0,32	-0,20
	SD	0,98	0,99	0,99	1,01	0,99	0,99	0,93	0,98	0,94	0,92	0,97
INDIVIDUAL ATTACKING ACTIONS	MEAN	0,04	0,19	0,04	-0,09	-0,03	-0,03	0,07	-0,27	-0,14	0,15	0,07
	SD	1,05	1,06	1,01	0,97	1,01	0,94	1,03	0,89	0,95	1,08	0,97
TEDENCY TO CREATE FINAL ATTEMPTS	MEAN	0,03	-0,01	0,03	0,11	0,01	-0,24	0,18	-0,12	-0,14	0,28	-0,13
	SD	0,87	1,00	0,89	1,07	0,95	1,01	1,10	0,98	0,93	0,97	0,96
PASSING TEMPO	MEAN	-0,11	0,21	0,06	-0,25	-0,29	-0,01	-0,30	0,44	0,16	0,17	0,17
	SD	0,97	0,98	0,88	0,99	1,04	1,06	0,99	1,12	0,92	0,86	0,89
DEFENDING AGGRESSIVELY	MEAN	0,23	0,05	-0,09	0,59	0,35	-0,59	-0,28	0,18	-0,28	-0,15	-0,13
	SD	0,95	1,03	0,97	0,90	0,84	0,92	0,96	0,96	0,99	0,91	0,96
ATTACKING AGGRESSIVELY	MEAN	0,06	0,14	-0,18	-0,08	0,07	-0,35	-0,12	0,05	0,14	-0,02	0,11
	SD	1,12	1,04	0,89	0,95	0,96	0,92	1,10	0,96	1,04	0,90	0,94
OFFSIDE TRAP	MEAN	0,05	0,08	-0,07	-0,15	0,04	-0,17	-0,08	0,09	0,04	-0,05	0,17
	SD	1,06	1,02	0,88	0,93	0,95	0,94	1,14	0,99	1,02	0,96	0,95

Table 4. The ranking of the 11 countries on the 19 latent variables.

ELABORATION OF BUILD UP PHASE	TUR	ITA	GRE	GER	UK	SPA	BEL	CRO	SWI	AUS	SCO
TRANSITION GAME	AUS	BEL	CRO	SWI	SCO	GER	SPA	TUR	GRE	ITA	UK
ATTACKING TRANSITION	TUR	AUS	GER	BEL	GRE	SWI	CRO	UK	ITA	SPA	SCO
DEFENSIVE TRANSITION	AUS	GER	BEL	TUR	SWI	CRO	GRE	UK	ITA	SPA	SCO
AERIAL GAME	SCO	AUS	GER	ITA	UK	SPA	CRO	BEL	GRE	TUR	SWI
TYPE OF ATTACK	SCO	GER	AUS	UK	SPA	GRE	BEL	ITA	TUR	SWI	CRO
CROSSING	SCO	SPA	TUR	CRO	SWI	UK	GRE	BEL	ITA	AUS	GER
TYPE OF OPPONENT ATTACK	SWI	TUR	CRO	ITA	GRE	BEL	SPA	UK	AUS	SCO	GER
DEFENSIVE BLOCKS	CRO	GRE	ITA	GER	SPA	BEL	AUS	UK	TUR	SWI	SCO

PRESS	ITA	UK	BEL	GER	SWI	TUR	AUS	SPA	CRO	SCO	GRE
INDIVIDUAL DEFENSIVE ACTIONS	BEL	SWI	AUS	GRE	CRO	GER	UK	TUR	ITA	SPA	SCO
WIDTH OF CREATIVE PHASE	GER	CRO	ITA	AUS	SWI	TUR	SPA	GRE	UK	SCO	BEL
EFFECTIVE GAME	AUS	GRE	SCO	SWI	SPA	UK	BEL	CRO	TUR	ITA	GER
INDIVIDUAL ATTACKING ACTIONS	BEL	SWI	ITA	TUR	AUS	CRO	GER	GRE	UK	SPA	SCO
TEDENCY TO CREATE FINAL ATTEMPTS	SWI	ITA	UK	AUS	CRO	GER	BEL	SCO	TUR	SPA	GRE
PASSING TEMPO	SCO	BEL	SWI	TUR	SPA	CRO	GRE	AUS	UK	GER	ITA
DEFENDING AGGRESSIVELY	UK	GER	AUS	SCO	BEL	CRO	TUR	SWI	ITA	SPA	GRE
ATTACKING AGGRESSIVELY	SPA	BEL	TUR	GER	AUS	SCO	SWI	UK	ITA	CRO	GRE
OFFSIDE TRAP	TUR	SCO	BEL	AUS	SPA	GER	SWI	CRO	ITA	UK	GRE

Pairwise comparisons showed that in 9 of the 19 variables (type of attack, type of opponent attack, defensive blocks, press, individual defending actions, width of creative phase, individual attacking actions, attacking aggressively, and offside trap) although there were statistically significant differences, there were no significant practical differences (partial η^2<0.06). For the remaining variables, the following important practical differences were observed:

ELABORATION OF BUILD UP PHASE: Turkish teams use more possession style than Austrian (η^2 = 0.065) and Scottish (η^2 = 0.071) teams, who prefer direct play.
TRANSITION GAME: The significant differences between the leagues are too many. Table 5 shows the partial η^2 of the pairwise comparisons.

Table 5. Partial η^2 of pairwise comparisons on transition game.

	AUSTRIA	BELGIUM	CROATIA	ENGLAND	GERMANY	GREECE	ITALY	SCOTLAND	SPAIN	SWISS	TURKYIE
AUSTRIA		0,02	0,06	0,3	0,1	0,25	0,25	0,09	0,21	0,08	0,25
BELGIUM			0,01	0,25	0,04	0,17	0,21	0,02	0,15	0,02	0,19
CROATIA				0,17	0,01	0,12	0,13	0	0,08	0	0,12
ENGLAND					0,12	0	0	0,13	0,02	0,14	0,01
GERMANY						0,07	0,09	0	0,04	0	0,07
GREECE							0	0,09	0,01	0,09	0
ITALY								0,1	0,01	0,11	0,01
SCOTLAND									0,06	0	0,09
SPAIN										0,06	0
SWISS											0,09
TURKYIE											

* RED CELLS=LARGE EFFECT SIZE, GREEN CELLS =MEDIUM EFFECT SIZE, YELLOW CELLS= SMALL OR VERY SMALL EFFECT SIZE
** The same applies to Tables 6, 7 and 8

ATTACKING TRANSITION: In the attacking transition phase, the teams of Scotland manage to carry out fewer counterattacks than the teams of Austria (η^2 = 0.073), Belgium (η^2 = 0.063), Germany (η^2 = 0.065), Greece (η^2 = 0.063) and Turkey (η^2 = 0.074).

DEFENSIVE TRANSISTION: Although the initial variables loading this factor are different from the previous one, here too it appears that Austrian teams are counter-attacked more often than Scottish teams (η^2 = 0.067).

AERIAL GAME: Scottish teams make more use of the game in the air than any other country (Table 6). On the contrary, the Swiss teams use it less than the Austrian (η^2=0.125), the English (η^2=0.072), the German (η^2=0.104), the Italian (η^2=0.061) and of course the Scottish (η^2=0.316) teams.

Table 6. Partial η^2 of pairwise comparisons between Scotland and rest countries on aerial game.

	AUSTRIA	BELGIUM	CROATIA	ENGLAND	GERMANY	GREECE	ITALY	SPAIN	SWISS	TURKYIE
SCOTLAND	0,08	0,15	0,16	0,12	0,08	0,22	0,10	0,14	0,32	0,23

CROSSING: Scottish teams adopt a style with many crosses more often than Austrian (η^2=0.065) and German teams (η^2=0.091).

EFFECTIVE GAME: In the games of the German, Italian, Turkish, Belgian and Croatian leagues, the ball is overall more in clear possession by one team or the other than in the games of the Swiss, Austrian, Greek, and Scottish championships (more interruptions and duels, which reduce the net time of possession of the two teams) (Table 7).

Table 7. Partial η^2 of pairwise comparisons between countries with the bigger differences on effective time.

	SWISS	AUSTRIA	GREECE	SCOTLAND
BELGIUM	0,05	0,09	0,07	0,05
GERMANY	0,09	0,13	0,11	0,09
ITALY	0,06	0,11	0,09	0,07
TURKYIE	0,06	0,10	0,08	0,07
CROATIA	0,05	0,11	0,07	0,05

TEDENCY TO CREATE FINAL ATTEMPTS: The Swiss teams show a greater tendency to create final attempts than the Greek ones (η^2=0.065).

PASSING TEMPO: It is lower in the Scottish league than the Italian (.104), German (.100), English (.091) and Austrian (.060). Additionally, it is larger in Italian than Belgian (0.064).

DEFENDING AGGRESSIVELY: Defensive aggressiveness seems to be lower in England and Germany than in Switzerland, Italy, Spain, and Greece (Table 8). Furthermore, in Greece, it seems to be bigger than Austria (η^2 = 0.156), Belgium (η^2 = 0.090), Croatia (η^2 = 0.065), and Scotland (η^2 = 0.144).

Table 8. Partial η^2 of pairwise comparisons between countries with the bigger differences on defending aggressively.

	SWISS	ITALY	SPAIN	GREECE
ENGLAND	0,13	0,18	0,17	0,27
GERMANY	0,07	0,10	0,10	0,21

T-SNE method & k-means Clustering: The t-SNE method extracted 2 dimensions (coordinates) for each country using the country means on the 19 latent variables. Then, using the values from the two coordinates extracted with the t-SNE method, k-means clustering was applied. From this method, the division of the 11 countries into 5 groups was requested. The values in the two dimensions and the cluster of each country are shown in Table 9 and Figure 1.

Table 9. Coordinates extracted by the t-SNE method and clusters extracted by k-means.

COUNTRY	X	Y	CLUSTER
AUSTRIA	-1046,85	98,05	1
BELGIUM	-729,09	277,31	1
CROATIA	-481,61	201,25	2
ENGLAND	136,37	-134,97	5
GERMANY	-190,12	-172,94	5
GREECE	454,63	251,05	3
ITALY	167,84	424,17	3
SCOTLAND	-549,46	-554,04	4
SPAIN	-79,55	245,37	2
SWISS	-691,49	-36,77	1
TURKYIE	-179,37	488,37	2

Discussion

The purpose of the present study was to compare the playing styles of the leagues of 11 different European countries during the 2021-22 season. This research represents a significant step forward in understanding the distinct playing styles across various European football leagues. The application of Multivariate Analysis of Variance (MANOVA), t-Distributed Stochastic Neighbor Embedding (t-SNE), and k-means clustering methods has facilitated a nuanced and comprehensive exploration of this complex subject. The 19 latent variables (representing 38 styles of game) were examined both separately (applying MANOVA) and as a whole (applying t-SNE and clustering methods) for 11 countries.

When the styles of play were looked at as a whole (Figure 1) it appeared that in Scotland there is a different style of play comparing to all other countries. They seem to continue using the traditional British way of playing with lots of long passes, crosses and aerial duels (George et al., 2014). On the contrary, in England the presence of coaches like Guardiola, Sarri, etc. who use many short passes and do not risk losing possession of the ball (Cintia & Pappalardo, 2021; Ellul et al., 2022) has brought tactical evolution of the English Premier League towards a more combinative and passing style of play in recent years, especially driven by the highest ranked clubs (Gonzalez-Rodenas et al., 2020). Perhaps, Scottish nationalism and culture has not yet allowed this development in the domestic league (Jarvie & Reid, 1999; Moorhouse, 1987).

Another example that seems to confirm the opinions of the coaches who consider that the general culture affects the adoption of playing styles in football (Sarmento et al., 2013) is the presence of the Greek and Italian leagues in the same cluster (3). The famous phrase "Una Faccia, Una Razza" expresses the similarity of the two cultures (Koulaouzides & Romano, 2022). Of the countries in the first 4 places of the UEFA ranking, Germany and England are the ones with the highest values in terms of body mass and height of players. This may gives a satisfactory explanation of their common presence in the cluster 5 (Bloomfield et al., 2005). Cluster 1 consists of the only countries in our research that come from Western Europe (Avendano et al., 2006; Harzer & Weber, 2022) without being in the top four of the UEFA ranking. Perhaps their close geographic location explains their common presence in the same cluster.

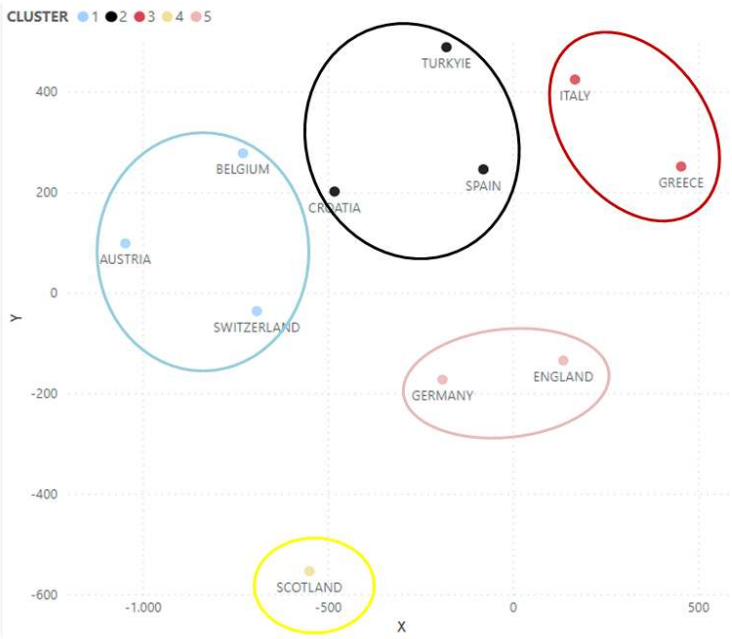

Figure 1. Scatter plot representing the results of the t-SNE method and the Cluster analysis.

When the 19 variables were analyzed separately for the differences between the countries it was seen that in 9 of them (*type of attack, type of opponent attack, defensive blocks, press, individual defending actions, width of creative phase, individual attacking actions, attacking aggressively,* and *offside trap*) there were no significant practical differences. From studies done in the past regarding performance indicators that load the specific 9 variables it was seen that: a) players from Serie A performed lower number of *dribbles per match* than Premier League players (Yi et al., 2019) and b) *attempted tackles per match* was higher in the Bundesliga comparing to English, Italian and Spanish league (Sapp et al., 2018). However, the first study was conducted in team matches in the Champions League and not in domestic competitions, while the second only examines *tackling* and not an overall style of play resulting from *individual defensive actions*, which was the case in the present study.

Regarding the remaining 10 latent variables, the differences between the countries revealed by the pairwise comparisons were too many. Only one of the 10 variables concerned the phase of a team's organized defense and was related to *fouls committed* and *yellow cards* received. Oberstone (2011) showed that in the English league *yellow cards* and *fouls* are less than the Italian and Spanish league and this finding agrees with the findings of our own research. However, in our own research due to the availability of data from more countries, other differences were also seen. In particular, it appeared that the *defensive aggressiveness* was lower in the English league compared to the Swiss and the Greek. It was also smaller in German compared to Swiss, Spanish, Italian and Greek. However, at this point it should be noted that these differences may not only be due to the difference in *defensive aggressiveness* of the players but also to the ease or not with which the referees give a yellow card or a foul. Therefore, perhaps a research comparing the *defensive aggressiveness* of teams from different countries participating in the same competition, such as the Champions League, the Europa League or the Conference League would be useful in the future. No other factor related to organized defense of the teams was found to differ between countries and this is in agreement with the research by Yi et al. (2019) who found no statistically significant differences in defense-related variables at all.

For the organized attack of a team, the variables in which differences were found between countries were: *elaboration of build-up, crosses, tendency to create final attempts and passing tempo*. Regarding *elaboration of build-up*, only Turkish teams seemed to use more *possession style* than Austrian and Scottish teams. However, previous research had shown more differences in separate performance indicators related to this variable. Oberstone (2011) had shown that Italian teams are better in *passing accuracy* than English and Spanish teams, while Yi et al. (2019) had shown the complete opposite for Italian teams in the Champions League; Mitrotasios et al. (2019) had shown that the Spanish teams had a greater proportion of *long and combinative attack* and that the Italian ones had the shortest *offensive sequences*, while González-Rodenas et al. (2020), contrary to Mitrotasios et al. (2019), had shown that between England and Spain there are no differences in the odds of implementing *direct attacks* rather than *combinative attacks*; finally, García-Aliaga et al. (2022) had shown that in the English league *long passes* are less frequent than in the other top leagues. As far as *crosses* are concerned, Oberstone (2011) had shown that teams from the Spanish League tend to cross the ball more than Italian teams, but this was not confirmed in our research which, however, was done with data 13 years later. In this one, only the Scottish teams seemed to make more *crosses* than the German and Austrian teams. The research of Yi et al. had shown that German League players take more *shots* than Spanish team players, but this was not confirmed in the present research on domestic competitions, where a difference in the *propensity to create final attempts* was only seen between Swiss and Greek teams. Additionally, in the present research it was shown that the *passing rate* was lower in the Scottish league than the Italian, German, English and Austrian, while at the same time in the Italian league it was higher than the Belgian one. No previous research had studied this factor.

Three variables in which there were differences between countries were associated with transitions. The first two (*attacking transition, defending transition*) referred to whether a team or its opponent, respectively, after recovering the ball, attempts a counterattack or goes into an organized attack, while the third (*transition game*) referred to the number of changes in possession of the ball in total for both teams. Teams from Austria counter-attacked more than teams from Scotland. Also in the games of the Austrian, Belgian, Swiss, Croatian and Scottish leagues there were more transitions than in the games of the Spanish, Turkish, Greek, Italian and English leagues. Mitrotasios et al. (2019) had previously shown that German and Italian teams attempt counter-attacks (instead of other types of attacks) more often than Spanish and English teams.

Apart from transition game in two other variables (*aerial game, effective game*) that referred to the game as a whole (i.e. they related to both teams) there were differences between the games of different countries. The research showed that Scottish teams made more use of the game in the air than any other country. On the contrary, the Swiss teams used it less than the Austrian, the English, the German, the Italian and of course the Scottish teams. Also, in the games of the German, Italian, Turkish, Belgium and Croatian leagues the ball was overall more in clear possession by one team or the other than in the games of the Swiss, Austrian, Greek and Scottish championships (more *interruptions* and *duels*, which reduces the net time of possession of the two teams). There is no previous research with comparisons between leagues of different countries in the specific fields.

The present study, together with those of Mitrotasios et al. (2019) and Gonzalez-Rodenas et al. (2020), aimed to find to find differences in playing styles (and not in separate performance indicators) between leagues of various countries. However, our study presents several novelties in comparison to Mitrotasios et al. (2019) and Gonzalez-Rodenas et al. (2020) because: a) they only dealt with the possession phase and not the other phases of the game, b) the sample of games they used (80 and 40 respectively) was much smaller than ours (2996) and c) they found differences among few leagues (4 and 2 respectively) compared to 11 of our own research.

While this study provides valuable insights into the differences in playing styles across various European leagues, it is not without its limitations. First, the study focuses on eleven leagues and may not fully capture the nuances of playing styles in leagues not included in this analysis. Second, the latent variables

identified, while comprehensive, may not encompass all aspects of playing styles. The complexity of football suggests that there may be other influential factors not considered in this study. Third, the use of specific statistical methods, though robust and effective, may present certain inherent biases or restrictions. For example, k-means clustering assumes spherical clusters and equal cluster sizes, which may not always reflect the true nature of the data. Lastly, the study is cross-sectional, providing a snapshot of the playing styles at a specific point in time. Football strategies and tactics are dynamic and continuously evolving, which may not be fully captured in a cross-sectional study. Future research could consider these limitations and further extend the understanding of playing styles in football.

Conclusion

Our approach has successfully revealed the heterogeneous nature of football playing styles that are influenced by a range of factors, from national culture to coaching philosophies. Notably, our findings indicate significant differences in ten out of nineteen latent variables, with transition game, effective game, and defending aggressively marking the most substantial disparities between countries. These results highlight the multifaceted nature of football, suggesting that each league and country exhibits its unique blend of strategies and tactics, thereby enriching the diversity of the sport. To the best of our knowledge, this is the first study that directly compares soccer playing styles in these eleven competitions. Therefore, the results and conclusions provide important insights and practical applications to coaches, managers, scouters and club sporting directors, because studying the potential similarities and differences between the distinct championships that take place in Europe is essential for an integral management of teams. In particular, the understanding of the different tactical requirements of each competition facilitates important goals, such as the creation of the appropriate tactical model of the team as well as the optimal detection and recruitment of players and coaches.

References

Alberti, G., Iaia, F. M., Arcelli, E., Cavaggioni, L., & Rampinini, E. (2013). Goal scoring patterns in major European soccer leagues. *Sport Sciences for Health*, *9*(3), 151-153.

Avendano, M., Kunst, A. E., Huisman, M., Lenthe, F. V., Bopp, M., Regidor, E., Glickman, M., Costa, G., Spadea, T., & Deboosere, P. (2006). Socioeconomic status and ischaemic heart disease mortality in 10 western European populations during the 1990s. *Heart*, *92*(4), 461-467.

Beaudouin, F., Gioftsidou, A., Larsen, M. N., Lemmink, K., Drust, B., Modena, R., Espinola, J. R., Meiu, M., Vouillamoz, M., & Meyer, T. (2020). The UEFA Heading Study: Heading incidence in children's and youth'football (soccer) in eight European countries. *Scandinavian Journal of Medicine & Science in Sports*, *30*(8), 1506-1517.

Bekkers, J., & Dabadghao, S. (2019). Flow motifs in soccer: What can passing behavior tell us? *Journal of Sports Analytics*, *5*(4), 299-311.

Beyene, K., & Bekele, S. A. (2016). Assessing univariate and multivariate homogeneity of variance: A guide for practitioners. *Journal of Mathematical Theory and Modeling*, *6*(5), 13-17.

Bloomfield, J., Polman, R., Butterly, R., & O'Donoghue, P. (2005). Analysis of age, stature, body mass, BMI and quality of elite soccer players from 4 European Leagues. *J Sports Med Phys Fitness*, *45*(1), 58-67.

Bredtmann, J., Crede, C. J., & Otten, S. (2014). The effect of gender equality on international soccer performance. *Ruhr Economic Paper*(501).

Brooks, J., Kerr, M., & Guttag, J. (2016). Using machine learning to draw inferences from pass location data in soccer. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *9*(5), 338-349.

Casal, C. A., Maneiro, R., Ardá, A., & Losada, J. L. (2019). Gender differences in technical-tactical behaviour of Laliga Spanish football teams.

Castellano, J., & Echeazarra, I. (2019). Network-based centrality measures and physical demands in football regarding player position: Is there a connection? A preliminary study. *Journal of sports sciences*, *37*(23), 2631-2638.

Castellano, J., & Pic, M. (2019). Identification and preference of game styles in LaLiga associated with match outcomes. *International Journal of Environmental Research and Public Health*, *16*(24), 5090.

Cintia, P., & Pappalardo, L. (2021). Coach2vec: autoencoding the playing style of soccer coaches. *arXiv preprint arXiv:2106.15444*.

Clemente, F. M., Owen, A., Serra-Olivares, J., Nikolaidis, P. T., Van Der Linden, C. M., & Mendes, B. (2019). Characterization of the weekly external load profile of professional soccer teams from Portugal and the Netherlands. *Journal of Human kinetics*, *66*(1), 155-164.

Clemente, F. M., Silva, F., Martins, F. M. L., Kalamaras, D., & Mendes, R. S. (2016). Performance Analysis Tool for network analysis on team sports: A case study of FIFA Soccer World Cup 2014. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, *230*(3), 158-170.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

De Ste Croix, M., Ayala, F., Sanchez, S. H., Lehnert, M., & Hughes, J. (2020). Grass-root coaches knowledge, understanding, attitude and confidence to deliver injury prevention training in youth soccer: a comparison of coaches in three EU countries. *Journal of Science in Sport and Exercise*, *2*(4), 367-374.

Dellal, A., Chamari, K., Wong, d. P., Ahmaidi, S., Keller, D., Barros, R., Bisciotti, G. N., & Carling, C. (2011). Comparison of physical and technical performance in European soccer match-play: FA Premier League and La Liga. *European Journal of Sport Science*, *11*(1), 51-59.

Doyle, J. R., & Bottomley, P. A. (2018). Relative age effect in elite soccer: More early-born players, but no better valued, and no paragon clubs or countries. *PloS one*, *13*(2), e0192209.

Ellul, S., Spiteri, J. V., Delice, M. E., & Grima, S. (2022). Managerial Decisions on the Technical Performance Efficiency: The Impact on Elite English Premier League Clubs. *Journal of Sports Economics & Management*, *12*(1), 23-40.

Fernandez-Navarro, J., Fradua, L., Zubillaga, A., Ford, P. R., & McRobert, A. P. (2016). Attacking and defensive styles of play in soccer: analysis of Spanish and English elite teams. *Journal of sports sciences*, *34*(24), 2195-2204.

Ford, P. R., Carling, C., Garces, M., Marques, M., Miguel, C., Farrant, A., Stenling, A., Moreno, J., Le Gall, F., & Holmström, S. (2012). The developmental activities of elite soccer players aged under-16 years from Brazil, England, France, Ghana, Mexico, Portugal and Sweden. *Journal of sports sciences*, *30*(15), 1653-1663.

García-Aliaga, A., Marquina Nieto, M., Coterón, J., Rodríguez-González, A., Gil Ares, J., & Refoyo Román, I. (2022). A Longitudinal Study on the Evolution of the Four Main Football Leagues Using Artificial Intelligence: Analysis of the Differences in English Premier League Teams. *Research Quarterly for Exercise and Sport*, 1-9.

George, S., Ionel, M., & Cristian, P. (2014). A comparative Study on the Evolution of the Parameters in Professional Soccer Matches. *Procedia-Social and Behavioral Sciences*, *127*, 63-67.

Gollan, S., Ferrar, K., & Norton, K. (2018). Characterising game styles in the English Premier League using the "moments of play" framework. *International Journal of Performance Analysis in Sport*, *18*(6), 998-1009.

Gómez, M.-Á., Mitrotasios, M., Armatas, V., & Lago-Peñas, C. (2018). Analysis of playing styles according to team quality and match location in Greek professional soccer. *International Journal of Performance Analysis in Sport*, *18*(6), 986-997.

González-Rodenas, J., Aranda-Malaves, R., Tudela-Desantes, A., Nieto, F., Usó, F., & Aranda, R. (2020). Playing tactics, contextual variables and offensive effectiveness in English Premier League soccer matches. A multilevel analysis. *PloS one*, *15*(2), e0226978.

Gonzalez-Rodenas, J., Aranda, R., & Aranda-Malaves, R. (2020). The effect of contextual variables on the attacking style of play in professional soccer.

Gyarmati, L., Kwak, H., & Rodriguez, P. (2014). Searching for a unique style in soccer. *arXiv preprint arXiv:1409.0308*.

Harzer, C., & Weber, M. (2022). Region of Western Europe (Austria, Belgium, France, Germany, the Netherlands, and Switzerland). In *The International Handbook of Positive Psychology* (pp. 185-221). Springer.

Heuvel, R. M. (2016). The Soccer Stock Market Anomaly.

Hill, Y., & Van Yperen, N. W. (2021). Losing the home field advantage when playing behind closed doors during COVID-19: change or chance? *Frontiers in psychology*, *12*, 658452.

Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *Journal of sports sciences*, *20*(10), 739-754.

Jarvie, G., & Reid, I. A. (1999). Sport, nationalism and culture in Scotland. *Sports Historian*, *19*(1), 97-124.

Khalilzadeh, J., & Tasci, A. D. (2017). Large sample size, significance level, and the effect size: Solutions to perils of using big data for academic research. *Tourism Management*, *62*, 89-96.

Kirkendall, D. T. (2020). Evolution of soccer as a research topic. *Progress in Cardiovascular Diseases*, *63*(6), 723-729.

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, *1*(6), 90-95.

Koulaouzides, G., & Romano, A. (2022). Una Faccia, Una Razza: Similarities, Differences, and Parallels in Adult Education Policy Development in Greece and Italy. *International Journal of Education Policy and Leadership*, *18*(1), 1-16.

Lago-Peñas, C., Gómez-Ruano, M., & Yang, G. (2017). Styles of play in professional soccer: an approach of the Chinese Soccer Super League. *International Journal of Performance Analysis in Sport*, *17*(6), 1073-1084.

Lopez-Valenciano, A., Garcia-Gómez, J. A., López-Del Campo, R., Resta, R., Moreno-Perez, V., Blanco-Pita, H., Valés-Vázquez, Á., & Del Coso, J. (2022). Association between offensive and defensive playing

style variables and ranking position in a national football league. *Journal of sports sciences*, *40*(1), 50-58.

Martínez, E. L., García, G. M. G., & Molina-Martín, J. J. (2023). Quantification of the competition load of the sets in high-level volleyball in the year 2021. *Journal of Physical Education and Sport*, *23*(1), 134-142.

Mitrotasios, M., Gonzalez-Rodenas, J., Armatas, V., & Aranda, R. (2019). The creation of goal scoring opportunities in professional soccer. tactical differences between spanish la liga, english premier league, german bundesliga and italian serie A. *International Journal of Performance Analysis in Sport*, *19*(3), 452-465.

Moorhouse, H. F. (1987). Scotland against England: Football and popular culture. *The international journal of the history of Sport*, *4*(2), 189-202.

Narizuka, T., & Yamazaki, Y. (2019). Clustering algorithm for formations in football games. *Scientific reports*, *9*(1), 1-8.

Oberstone, J. (2011). Comparing team performance of the English premier league, Serie A, and La Liga for the 2008-2009 season. *Journal of Quantitative Analysis in Sports*, *7*(1).

Oikonomidis, A., Bruce, A. C., & Johnson, J. E. (2015). Does transparency imply efficiency? The case of the European soccer betting market. *Economics Letters*, *128*, 59-61.

Plakias, S., Kokkotis, C., Moustakidis, S., Tsatalas, T., Papalexi, M., Kasioura, C., Giakas, G., & Tsaopoulos, D. (2023). Identifying playing styles of european soccer teams during the key moments of the game. *Journal of Physical Education and Sport*, *23*(4), pp. 878 - 890.

Plakias, S., Kokkotis, C., Tsaopoulos, D., Moustakidis, S., Papalexi, M., Giakas, G., & Tsatalas, T. (2023). The effectiveness of direct corners in high level soccer depending on the type and the zone of delivery. *Journal of Physical Education and Sport*, *23*(2), 449-456.

Plakias, S., Mandroukas, A., Kokkotis, C., Michailidis, Y., Mavromatis, G., & Metaxas, T. (2022). The correlation of the penetrative pass on offensive third with the possession of the ball in high level soccer. *Gazzetta Medica Italiana-Archivio per le Scienze Mediche*, *181*(9), 633-638.

Plakias, S., Moustakidis, S., Kokkotis, C., Tsatalas, T., Papalexi, M., Plakias, D., Giakas, G., & Tsaopoulos, D. (2023). Identifying soccer teams' styles of play: a scoping and critical review. *Journal of Functional Morphology and Kinesiology*, *8*(2), 39.

Pollard, R. (2006). Worldwide regional variations in home advantage in association football. *Journal of sports sciences*, *24*(3), 231-240.

Pollard, R., & Gómez, M. A. (2009). Home advantage in football in South-West Europe: Long-term trends, regional variation, and team differences. *European Journal of Sport Science*, *9*(6), 341-352.

Pollard, R., & Gómez, M. A. (2014). Components of home advantage in 157 national soccer leagues worldwide. *International journal of sport and exercise psychology*, *12*(3), 218-233.

Pollard, R., & Pollard, G. (2005). Home advantage in soccer: A review of its existence and causes.

Ramírez-Arroyo, A., García, L., Alex-Amor, A., & Valenzuela-Valdés, J. F. (2022). Artificial Intelligence and Dimensionality Reduction: Tools for approaching future communications. *IEEE Open Journal of the Communications Society*, *3*, 475-492.

Roberts, J., Osei-Owusu, P., Harland, A., Owen, A., & Smith, A. (2014). Elite football players' perceptions of football turf and natural grass surface properties. *Procedia Engineering*, *72*, 907-912.

Ruan, L., Ge, H., Gómez, M.-Á., Shen, Y., Gong, B., & Cui, Y. (2022). Analysis of defensive playing styles in the professional Chinese Football Super League. *Science and Medicine in Football*, 1-9.

Sapp, R. M., Spangenburg, E. E., & Hagberg, J. M. (2018). Trends in aggressive play and refereeing among the top five European soccer leagues. *Journal of sports sciences*, *36*(12), 1346-1354.

Sarmento, H., Marcelino, R., Anguera, M. T., CampaniÇo, J., Matos, N., & LeitÃo, J. C. (2014). Match analysis in football: a systematic review. *Journal of sports sciences*, *32*(20), 1831-1843.

Sarmento, H., Pereira, A., Matos, N., Campaniço, J., Anguera, T. M., & Leitão, J. (2013). English premier league, spaińs la liga and italýs seriés a–What's different? *International Journal of Performance Analysis in Sport*, *13*(3), 773-789.

Yi, Q., Groom, R., Dai, C., Liu, H., & Gómez Ruano, M. Á. (2019). Differences in technical performance of players from 'the big five'European football leagues in the UEFA Champions League. *Frontiers in psychology*, *10*, 2738.

Zhou, C., Lago-Peñas, C., Lorenzo, A., & Gómez, M.-Á. (2021). Long-term trend analysis of playing styles in the Chinese soccer super league. *Journal of Human kinetics*, *79*(1), 237-247.