Review
Identifying Soccer Teams’ Styles of Play: A Scoping and Critical Review

Spyridon Plakias 1,*, Serafeim Moustakidis 2, Christos Kokkotis 3, Themistoklis Tsatalas 1, Marina Papalexi 4, Dionysios Plakias 5, Giannis Giakas 1,* and Dimitrios Tsaopoulos 6

1 Department of Physical Education and Sport Science, University of Thessaly, Karyes, 42100 Trikala, Greece
2 AIDEAS OÜ, Narva mnt 5, 10117 Tallinn, Estonia
3 Department of Physical Education and Sport Science, Democritus University of Thrace, 69100 Komotini, Greece
4 Department of Operations, Technology, Events and Hospitality Management, Manchester Metropolitan University, Oxford Road, Manchester M15 6BH, UK
5 Network Rail, Piccadilly Tower, Manchester M60 7RA, UK
6 Institute for Bio-Economy & Agri-Technology, Center for Research and Technology Hellas, 60361 Volos, Greece
* Correspondence: spyros_plakias@yahoo.gr (S.P.); ggiakas@gmail.com (G.G.)

Abstract: Identifying and measuring soccer playing styles is a very important step toward a more effective performance analysis. Exploring the different game styles that a team can adopt to enable a great performance remains under-researched. To address this challenge and identify new directions in future research in the area, this paper conducted a critical review of 40 research articles that met specific criteria. Following the 22-item Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines, this scoping review searched for literature on Google Scholar, Web of Science, and Pub Med databases. The descriptive and thematic analysis found that the objectives of the identified papers can be classified into three main categories (recognition and effectiveness of playing styles and contextual variables that affect them). Critically reviewing the studies, the paper concluded that: (i) factor analysis seems to be the best technique among inductive statistics; (ii) artificial intelligence (AI) opens new horizons in performance analysis, and (iii) there is a need for further research on the effectiveness of different playing styles, as well as on the impact of contextual variables on them.

Keywords: game style; playing styles; identify; contextual variables; effectiveness

1. Introduction

Performance analysis, i.e., the recording and examination of behavioral events occurring during a competition [1,2], is an essential tool in the hands of coaches. The relevant literature in soccer has traditionally focused on separated variables such as performance indicators to explain teams’ and players’ performance [3–5]. Performance indicators are variables that aim to define some or all aspects of performance. They may concern either a single action (e.g., pass, shot, recovery) or a combination of actions (ball possession, passes per defensive action, etc.) [6]. Recent research, in an attempt to analyze the complexity of the soccer game, has utilized playing styles instead of performance indicators, which may explain their tactical performance in matches and competitions more effectively [7,8].

Playing style is the characteristic pattern demonstrated and repeated by a team in specific situational contexts [9]. The scientific literature around playing styles has grown significantly, with some authors defining game styles based on their subjective perceptions. Notably, various styles have been mentioned such as supported, direct, defensive pressing, attacking [10], very direct and aggressive [11], direct, possession, offensive [12], Dutch [13], positional play [14], FC Barcelona style [15], and more.

The most remarkable shift in football’s evolution is the application of scientific tools supported by scientific data [16]. The use of event statistical data (from new software tools
such as Sportscode, Nacsport, and Longomatch) and tracking data (from optical methods and GPS) has ushered in the Big Data era [17]. The intersection of data science and sport science has tried to unlock the potential of big data to support tactical performance analysis with various methods, such as AI, multivariate statistical techniques and visualizations [18–20]. Researchers have successfully used AI techniques to address soccer analytics tasks such as detecting tactics [21]. In the field of game style identification, the multivariate statistical methods that are mostly used include factor analysis accompanied with principal component analysis (PCA), and k-means clustering. When the research question concerns the effect of contextual variables or the effectiveness of playing styles, then MANOVA and MANCOVA (Multivariate Analysis Of Variance and Covariance) are usually employed. Contextual variables refer to factors that can potentially affect performance, such as match status, match location, opponent level, type of competition, period of the season, playing surface, etc. [22]. Lastly, many studies have employed visualizations, transforming non-visual quantified data into visual forms, to facilitate a better comprehension of the results [23].

According to the above: (i) it is a given that the concept of “playing style” is used to better understand the complexity of the game and team tactics, (ii) there is a growing interest from researchers in recent years on playing styles, and (iii) there is not a single way of using the term, while many times it is used without scientific validation. Furthermore, despite researchers’ efforts, there are still several knowledge gaps in relation to expert interviews during qualitative studies, and so far, no attempt has been made in the literature to review the subject comprehensively. Therefore, the objectives of this review are: (i) to critically examine how the term is used in the international literature, how research is conducted on this topic and what are the main objectives of the research, and (ii) to identify and analyze knowledge gaps in order to contribute to future research and a future systematic review. By conducting this comprehensive analysis, we hope to provide valuable insights for researchers, coaches, and practitioners, enhancing their understanding of soccer playing styles and their effectiveness in improving team performance.

2. Materials and Methods

This scoping review followed the 22-item Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) [24]. Furthermore, to conduct the critical review, this paper has critically reviewed the selected articles, following the process adopted by Mackenzie and Cushion [25].

2.1. Literature Search Approach

This review was based on research articles published up until 25 September 2022 (with no restriction on the year of publication) using the search engines PubMed, Web of Science, and Google Scholar. During our search in PubMed and Web of Science, we combined the terms “soccer” and “style” using the Boolean operator AND. In contrast, when searching Google Scholar, due to the large number of initial results, we utilized the expression “soccer style” as a more focused search term. A prerequisite for the inclusion of an article in our study was the presence of the two mentioned terms (in singular or plural) as keywords, either in the title or in the abstract of each article [26].

2.2. Exclusion Criteria

Initially, two of the authors (S.P. and C.K.) checked the titles and abstracts of all retrieved publications. In cases where a disagreement arose between the two reviewers during the article selection process, a third independent reviewer was employed to resolve the discrepancy and reach a consensus. The following categories were excluded: (i) non-English articles, (ii) postgraduate and doctoral dissertations, review articles and books (iii) studies of styles in other sports, (iv) research based on robotic soccer and video games (v) articles related to individual player, coach or referee styles, and (vi) studies that did not focus on playing styles. All the selected articles are peer-reviewed and have been presented
either in journal papers or conferences. Finally, the rest of the authors reviewed again the titles and abstracts to ensure that they met the inclusion criteria.

2.3. Assessed Outcomes

The studies, which are recorded in this article, were divided into three categories, namely: (i) recognition of playing styles, (ii) contextual variables that influence the adoption of each style, and (iii) effectiveness of styles. The grouping was based on the most common objectives of the studies. Articles investigating multiple concepts were included in more than one category [27]. The first category includes all articles with the aim of separating distinct styles, identifying styles’ characteristics and quantifying them. Then, after separating the articles, the following information was extracted from each article: author, year of publication, method (classical inductive statistics/A.I./other), sample (number of matches, competitions), kind of data (P.I./tracking data/other), phases of the game concerning the styles studied, outcome.

3. Results
3.1. Search Results

The initial search identified 1417 titles in the described databases. After duplicates were removed the remaining 864 articles were then screened for relevance based on their title, resulting in another 251 studies being eliminated from the database. The abstract of the remaining 613 articles was then read and another 389 were rejected due to a lack of relevance to the purpose of this study. Finally, only 40 articles remained for the scoping review, when the entire articles were read (Figure 1).

![Figure 1. Flow chart of the methodology used for the article search.](image-url)
3.2. Descriptive Analysis

A total of 40 articles were identified in the review, after applying the proposed inclusion/exclusion criteria. In total, 29 of the 40 articles used classical inductive statistics alone or in combination with other methods (e.g., visualization), 10 used AI methods alone or in combination with other methods (e.g., visualization), and one study combined classical statistics with A.I. and observational methodology (Table 1). Table 1 also shows that 24 surveys used performance indicators from event data, eight used tracking data, while the rest used other types of data (flow motifs, team possessions, questionnaires, etc.). In the same table it can be seen that the sample was usually taken from country leagues (mainly the big four European i.e., English, Spanish, German, Italian) and the World Cup. The map in Figure 2 illustrates the domestic competitions that have been included in the relevant surveys.

![Map of countries whose leagues have been included in relevant surveys](map.png)

Figure 2. Map of countries whose leagues have been included in relevant surveys (the colour of the bubble represents the number of studies in which they have been included).

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Kind of Data</th>
<th>Sample (Number of Matches &amp; Competitions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenga [29]</td>
<td>2003</td>
<td>Chi-square</td>
<td>1 Brazil-Norway</td>
</tr>
<tr>
<td>Andersson [3]</td>
<td>2008</td>
<td>Two-way ANOVA, paired t-test, chi-square</td>
<td>* * Questionaries 10 Male/female Swedish league</td>
</tr>
<tr>
<td>Sporš [31]</td>
<td>2012</td>
<td>Factor-PCA, Cronbach’s alpha</td>
<td>Rating from 0–5 from ten experts</td>
</tr>
<tr>
<td>Basevitch [4]</td>
<td>2013</td>
<td>Independent t-tests. Multiple linear regression</td>
<td>* * Brazilian and Italian matches from all the World Cups, Brazilian and Italian premier leagues from 2003 to 2008</td>
</tr>
<tr>
<td>Lago-Peñas [34]</td>
<td>2017</td>
<td>Factor-PCA</td>
<td>240 Chinese Super League (SL) during the 2016 season</td>
</tr>
<tr>
<td>Santos [35]</td>
<td>2017</td>
<td>Linear regression</td>
<td>13 An elite Spanish team</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Category</td>
<td>Method</td>
</tr>
<tr>
<td>---------------------</td>
<td>------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>Fernandez-Navarro</td>
<td>2018</td>
<td></td>
<td>Linear mixed model (cross-classified multilevel design)</td>
</tr>
<tr>
<td>Yi [38]</td>
<td>2019</td>
<td></td>
<td>K-means clustering, Separate Poisson regression models</td>
</tr>
<tr>
<td>Fernandez-Navarro</td>
<td>2019</td>
<td></td>
<td>Linear mixed model (cross-classified multilevel design)</td>
</tr>
<tr>
<td>Castellano [39]</td>
<td>2019</td>
<td></td>
<td>Factor-PCA, Discriminant analyses, Chi-square</td>
</tr>
<tr>
<td>Mitrotasios [40]</td>
<td>2019</td>
<td></td>
<td>Kruskal-Wallis, Mann-Whitney</td>
</tr>
<tr>
<td>Praça [41]</td>
<td>2019</td>
<td></td>
<td>Social network analysis, one-way ANOVA, two-way ANOVA</td>
</tr>
<tr>
<td>Castellano [42]</td>
<td>2019</td>
<td></td>
<td>Social network analysis, magnitude-based inference and correlation</td>
</tr>
<tr>
<td>Drezezer [43]</td>
<td>2020</td>
<td></td>
<td>Chi-square</td>
</tr>
<tr>
<td>Gonzalez-Rodenas</td>
<td>2020</td>
<td></td>
<td>Multivariate logistic regressions</td>
</tr>
<tr>
<td>Schulze [49]</td>
<td>2021</td>
<td></td>
<td>Factor-PCA, Linear regressions</td>
</tr>
<tr>
<td>Lopez-Valenciano</td>
<td>2022</td>
<td></td>
<td>Pearson’s correlation coefficient tests, Spearman’s rank correlation coefficient test and PCA</td>
</tr>
<tr>
<td>Ruan [50]</td>
<td>2022</td>
<td></td>
<td>Factor-PCA</td>
</tr>
<tr>
<td>Bialkowski [53]</td>
<td>2014</td>
<td></td>
<td>K-means clustering, LDA, k-NN regression</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Category</th>
<th>Method</th>
<th>Kind of Data</th>
<th>Sample (Number of Matches &amp; Competitions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Białkowski [54] 2016</td>
<td></td>
<td></td>
<td>k-means clustering, agglomerative clustering, linear discriminant analysis, k-nearest neighbour</td>
<td>*</td>
<td>374, one league</td>
</tr>
<tr>
<td>Brooks [55] 2016</td>
<td></td>
<td></td>
<td>K-nearest neighbor, L2-regularized support vector machine model</td>
<td>Every pass (with 8 descriptors)</td>
<td>4 seasons (2012/2013 to 2015/2016), 6 different leagues (Dutch, English, Spanish, Italian, French and German first division)</td>
</tr>
<tr>
<td>Bekkers [56] 2019</td>
<td></td>
<td></td>
<td>Mean shift algorithm</td>
<td>Flow motifs</td>
<td>8219</td>
</tr>
<tr>
<td>Decroos [58] 2020</td>
<td></td>
<td></td>
<td>Mixture models</td>
<td>Actions described by their type, location, and direction</td>
<td>760, 2017/18 and 2018/19 seasons of the EPL</td>
</tr>
<tr>
<td>Beernaerts [59] 2020</td>
<td></td>
<td></td>
<td>Qualitative Trajectory Calculus</td>
<td>*</td>
<td>1, 2016–2017 professional soccer competition</td>
</tr>
<tr>
<td>García-Aliaga [60] 2022</td>
<td></td>
<td></td>
<td>t-SNE, dimensionality reduction technique, classification rules with RIPPER</td>
<td>*</td>
<td>EPL, Spanish LaLiga, German Bundesliga, and Italian Serie A from the 2014/2015 to 2018/2019 seasons</td>
</tr>
<tr>
<td>Lee [61] 2022</td>
<td></td>
<td></td>
<td>Deep Neural Networks (DNN) based on Multi-Layer Perceptron (MLP) and feature engineering</td>
<td>*</td>
<td>(a) all Tottenham’s games, (b) 380 (a) 11 seasons (2010/2011–2020/2021) English premier league, (b) 2020-21 EPL</td>
</tr>
</tbody>
</table>

3.3. Thematic Analysis

Table 2 shows the separation of studies based on the main categories mentioned in the methodology (assessed outcomes), considering that some of the articles belong to more than one category. Figure 3 provides a visualization of the number of articles by year and category.

Figure 3. Number of articles by year and category.
Table 2. Distinction of articles into three categories based on their main purpose.

<table>
<thead>
<tr>
<th>Recognition (28)</th>
<th>Contextual Variables (15)</th>
<th>Effectiveness (7)</th>
</tr>
</thead>
</table>

4. Discussion

4.1. Recognition

The vast majority of related literature has focused on game style recognition. Historically, explanations for the performance of teams and individual players were based on singular events or isolated behaviors. Basevitch, Yang and Tenenbaum [4], and Castellano and Echeazarrá [42] tried to match separate variables with the playing style of teams. Then, they tried to identify variations on the specific variables between the teams. Despite the significance of the information offered by these studies, they cannot cover the complexity of the soccer game, which requires the connection of different performance indicators for recognizing a team’s styles of play or tactical pattern [7,8]. To bridge this gap, Kempe, Vogelbein, Memmert, and Nopp [32], introduced two indexes to differentiate between possession play and direct style. Each index was the sum of z-values of performance indicators considered by the authors to determine game control and offensive behavior. The research was limited by the absence of validation for the two indexes, as the variables employed in their formulation were selected based on the authors’ subjective perceptions. K-means cluster analysis was also used by Gollan, Ferrar, and Norton [36] to recognize playing styles. Three game style clusters were identified: (1) moderately favoring established defense, (2) dominant in transition offense and transition defense, and (3) strong in established offense and set pieces. The disadvantage of this method is that it does neither recognize playing styles, nor it is capable of quantifying them; instead, it categorizes the teams based on the phases in which they excel. Finally, the studies of Dreznner, Lamas, Farias, Barrera and Dantas [43], Fernandes, Camerino, Garganta, Hileno and Barreira [46], and Tenga and Larsen [29] relied on the construction of their own models for game style recognition. The first two divided the field into sections (18 and 14, respectively). After recording the sequences of passes throughout each ball possession (from the beginning to its completion), Chi-Square analysis was performed. The first research found different styles between teams for ball circulation, while the second focused on defensive behavior patterns. Due to the time-consuming aspect of studying all ball possession sequences, it is difficult for the
aforementioned studies to investigate a large sample of matches. Thus, nine matches were analyzed in the first and 12 in the second. In a match between Norway and Brazil, Tenga and Larsen [29] utilized a similar technique involving the development of their own model and Chi-Square analysis. However, they did not divide the field into zones.

Among all studies that applied inductive statistics, those that used factor analysis with PCA (Table 3) were the ones that managed not only to distinguish distinct styles, but also to identify the characteristics of each style (based on the variables that loaded each component). In addition, they were able to quantify playing styles based on the factor scores. Nonetheless, as shown in Table 3 and Figure 4, the styles of play identified in published research have not been able to account for all phases and sub-phases of the game. For example, the absence of defensive set pieces can be seen. Despite the fact that, in the research of Gómez, Mitrotasios, Armatas, and Lago-Peñas [8], two factors have been given the same names and free kicks, it is clear (from observing the variables that load them) that they concern the attacking set-pieces and the attacking free-kicks, respectively. This subjectivity in the naming of latent variables is a disadvantage of studies using factor analysis [64].

Factors may not have been correctly labeled, and it should not be assumed that two factors with the same name are the same thing (jingle fallacy) [65]. For example, in the research of Lago-Peñas, Gómez-Ruano, and Yang [34], factors 4 and 5 are given the general name transitional play. However, observing the variables that load the factors, we find that the variables of factor four (lost balls) can only appear in situations of defensive transition, while the variables of factor five (picking up free balls) can only appear in situations of offensive transition. Similarly, Ruan, Ge, Gómez, Shen, Gong, and Cui [50] give factors 6 and 8 the same name (defense of goalkeeper) even though, as the authors report, different teams have higher values in each of them. Moreover, in the study by Gómez, Mitrotasios, Armatas and Lago-Peñas [8], factor four was named counter-attack despite the fact that one of the loaded variables is lost balls, which cannot result in a counter-attack. So, either it should be clarified that the style concerns the game as a whole (for both teams together) or if the authors wish to adopt it as a single team’s style, a more general label (such as transitional play) should be given, which can cover both offensive and defensive transitions. In the same study, factor six is exclusively related to the variable actions in own fourth (they had divided the field into fourths along its length and by own fourth they meant the one near the team’s goal). However, from what has come to our attention, it is unclear why this particular factor was called transitional play and more explanation should be provided by the authors.

Table 3. Factors extracted in the studies employing factor analysis with PCA.

<table>
<thead>
<tr>
<th>Article</th>
<th>Factors’ Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>[28]</td>
<td>Possession style, Crosses, High press</td>
</tr>
<tr>
<td>[31]</td>
<td>Finishing efficiency, Ball possession performance, Counter-attack efficiency, Type of defense (man to man to man/mixed), Redirection of the opposing team’s attack build-up</td>
</tr>
<tr>
<td>[33]</td>
<td>Possession directness, Width of ball regain, Use of crosses, Possession width, Defensive ball pressure, Progression of the attack</td>
</tr>
<tr>
<td>[34]</td>
<td>Possession style, Set pieces attack, Counterattacking play, Transitional play (2) *</td>
</tr>
<tr>
<td>[34]</td>
<td>Ball possession, Ending actions, Individual challenges, Counter attack, Set pieces, Transitional play, Fouling actions, Free-kick</td>
</tr>
<tr>
<td>[8]</td>
<td>High press, type of attack</td>
</tr>
<tr>
<td>[47]</td>
<td>High intensity play, Possession and passing, Offensive actions, Defensive actions, Individual challenges, Serious fouls, Attacking aggressively Defense close to the own goal, High intensity confrontation, Mid positioning defense with pressure, Error, Defense in advanced zones, Receiving a dangerous situation, Defense of goalkeeper (2) *</td>
</tr>
<tr>
<td>[50]</td>
<td>Constant, Receiving a dangerous situation, Defense closed to the own goal, Error, Keeper claim, High intensity confrontation, Mid-positioning defense with pressure, Defense in advanced zones, Keeper smother</td>
</tr>
</tbody>
</table>

* Two different factors have been given the same name.
which is a very common tactical tool of coaches [66–69], the passing tempo that gives very
A(t)) the formation of a team at time t. Then, using hierarchical clustering, they obtained

Then they applied T-pattern detection to identify patterns of play in possession. Finally,

They recorded the origin and destination of each pass in each distinct sequence of ball

The field has also been divided into zones (20) by Amatria, Maneiro and Anguera [62].

possession. At the same time, technical actions (ball control, dribbling) were recorded.

Thus, the various placements of the team’s in the plots reflect the disparity between

consequently, the various placements of the team’s in the plots reflect the disparity between

variables and the styles emerging from technical-tactical variables. Lastly, despite the fact

was fitted to the directions of the actions in that component. Using the learned mixture

models, each action was encoded as a weight vector. It was concluded that using specific

were fitted to the locations

In the first stage, for each action type, a mixture model was fitted to the locations

actions. In the first stage, for each action type, a mixture model was fitted to the locations

resultantly the various placements of the team’s in the plots reflect the disparity between

which can process large multivariate datasets and visualize them into a 2D plot [73]. Con-

volutive clustering was also utilized [54]. Brooks, Kerr, and Guttag [55] divided the field into

16 zones and counted the number of passes that originated from each zone. The classifi-

fication task was then accomplished using K-nearest neighbor (K-NN) classifier and a

heatmap of pass origins. The field has also been divided into zones (20) by Amatria, Maneiro and Anguera [62].

The studies of Bialkowski, Lucey, Carr, Yue, Sridharan, and Matthews [53], employed K-means clustering and Linear Discriminant Analysis (LDA) and distinguished five different styles adopted by teams during matches. Two years later, the same author published a similar article in which agglomerative clustering was also utilized [54]. Brooks, Kerr, and Guttag [55] divided the field into 16 zones and counted the number of passes that originated from each zone. The classification task was then accomplished using K-nearest neighbor (K-NN) classifier and a heatmap of pass origins. The field has also been divided into zones (20) by Amatria, Maneiro and Anguera [62]. They recorded the origin and destination of each pass in each distinct sequence of ball possession. At the same time, technical actions (ball control, dribbling) were recorded.

Then they applied T-pattern detection to identify patterns of play in possession. Finally, Bekkers and Dabadghao [56] used a mean shift algorithm to cluster teams based on their passing motifs, where four clusters emerged.

Narizuka and Yamazaki [57] adopted the Delaunay network to get (as adjacency matrix
A(t)) the formation of a team at time t. Then, using hierarchical clustering, they obtained

Figure 4. Classification of playing styles identified by Factor-PCA in each of the game phases.

Therefore, although factor analysis is an excellent solution for grouping variables to

Figure 4. Classification of playing styles identified by Factor-PCA in each of the game phases.
not only the average formation (i.e., “442”, “4141”, “433”, “541” or “343”) for each team in the match but also the positional exchange of players within the match formations. Decroos, Roy, and Davis [58] used mixture models to achieve a representation of soccer actions. In the first stage, for each action type, a mixture model was fitted to the locations (x, y). Then, for each component of each mixture model in stage 1, a new mixture model was fitted to the directions of the actions in that component. Using the learned mixture models, each action was encoded as a weight vector. It was concluded that using specific vectors, game styles could be identified based on each type of action. Finally, García-Aliaga, Marquina Nieto, Coterón, Rodríguez-González, Gil Ares, and Refoyo Román [60] employed the t-distributed Stochastic Neighbor Embedding (t-SNE) clustering algorithm, which can process large multivariate datasets and visualize them into a 2D plot [73]. Consequently, the various placements of the teams in the plots reflect the disparity between their playing styles.

As demonstrated above, a small number of studies focused solely on the possession phase of the ball [52,55,56,62]. The studies of Bialkowski, Lucey, Carr, Yue, Sridharan, and Matthews [53], Bialkowski, Lucey, Carr, Matthews, Sridharan, and Fookes [54], and Narizuka and Yamazaki [57] reached only conclusions regarding team formations. Bialkowski, Lucey, Carr, Yue, Sridharan, and Matthews [53] and Bialkowski, Lucey, Carr, Matthews, Sridharan, and Fookes [54] distinguished five different styles of play, however without information about the characteristics of each style, they are of limited practical value to coaches. After classifying the teams, García-Aliaga, Marquina Nieto, Coterón, Rodríguez-González, Gil Ares, and Refoyo Román [60], used the RIPPER method to find the variables that contribute most to the categorization of the teams. This has practical value, as opposed to simply grouping things together. The vectors of the actions which were created in the research of Decroos, Roy, and Davis [58] gave more information than simple numbers, but they also refer to separate variables. However, the ability to examine them simultaneously can provide valuable information to the coaching staff.

4.2. Contextual Variables

Contextual variables (specific match factors) have an impact on competitive demands and influence performance metrics [74,75]. In recent years, there has been a strong tendency to study the effect of contextual variables on the teams’ adoption of different playing styles. As shown in Table 2 and Figure 3, twelve of the thirteen relevant studies were conducted from 2018 onwards. The Treemap in Figure 5 shows that the majority of studies have explored the effect of match location. Particularly, Lago [30], Santos, Lago-Peñas, and García-García [35], Gómez, Mitrotasios, Armatas, and Lago-Peñas [8], Fernandez-Navarro, Fradua, Zubillaga, and McRobert [37], Bekkers and Dabadghao [56], Gollan, Bellenger, and Norton [44], and Gonzalez-Rodenas, Aranda, and Aranda-Malaves [45], dealt with the impact of the variable match location on the playing styles adopted by the teams. From all the above studies it seemed that home teams tried to build up from the back, rather than adopt direct play, resulting in a more combinational, possession-based style of play. They also seemed to try to speed up the match and press their opponents high, whereas in general, their game was more offensive.

Team’s ranking, opponent’s ranking, and match status are also variables whose effect on playing styles has been studied extensively. The research of Gonzalez-Rodenas, Aranda, and Aranda-Malaves [45], showed that high-ranked teams were less likely to use direct attacks than combinative attacks in comparison with low-ranked teams. García-Aliaga, Marquina Nieto, Coterón, Rodríguez-González, Gil Ares, and Refoyo Román [60] found that the top-ranked teams in the EPL differentiated from the rest of the teams in the same league by making less long passes, playing more vertically, even at the expense of accuracy, and dribbling. In [44], using EPL data of the 2015–2016 season in which Leicester shocked the league by winning it, it was demonstrated that established defense was prevalent among lower-ranked teams, while the champion was dominant in transitions and the rest of the high-ranked teams dominated established attack and set pieces. In this case, there appears to be a disadvantage associated with studies that collect their sample data.
from a single competition or, even worse, from a small number of matches. In such cases, generalizing the outcomes is inherently risky, as teams like Leicester, for instance, rarely win the league. As shown in Table 1, only 13 of the 40 studies processed data from more than one competition. As demonstrated by Lago [30], Santos, Lago-Peñas, and García-García [35], Fernandes, Camerino, Garganta, Hileno, and Barreira [46], Gollan, Bellenger, and Norton [44], and Gonzalez-Rodenas, Aranda, and Aranda-Malaves [45], the quality of the opposition exhibited similar trends with match location. Finally, match status appears to influence the adoption of different playing styles significantly. In particular, winning teams present higher probabilities of attacking by means of counterattacks and direct attacks than combinative attacks compared to losing teams. In addition, they apply high pressure less frequently [37,41,45,46].

![Figure 5. Proportion of articles based on contextual variables and game phases in which their effect on teams' playing styles was studied.](image)

However, there are some contextual variables whose effect on playing styles has received little or no attention. Comparisons between different competitions were performed amongst the top four leagues in Europe according to the UEFA ranking (English, Spanish, German, and Italian). Although Gonzalez-Rodenas, Aranda, and Aranda-Malaves [45] found no statistically significant differences between the English and Spanish leagues during the 2017–2018 season, Mitrotasios, Gonzalez-Rodenas, Armatas, and Aranda [40] discovered that Spain La Liga had a higher proportion of long and combinative attacks, the English Premier League had a higher tendency of progressing through fast and direct attacks, the German Bundesliga had the most counter-attacks, and the Italian Serie A had the shortest offensive sequences as well as a higher proportion of counter-attacks and direct attacks than combinative and fast attacks. García-Aliaga, Marquina Nieto, Coteron, Rodriguez-González, Gil Ares, and Refoyo Román [60] found a distinction between the English teams and the rest of the teams in the other leagues, determined by fewer free kicks, fewer long passes but more vertical, more errors in ball control, and greater success in dribbling, using data from the specific four countries from five different seasons (from 2014/2015 to 2018/2019).

The effect of match half time (first or second) was examined in two studies. One for the attacking phase [45] and one for the defensive phase [46]. The first found that progressing by counterattacks was less likely in the first half than progressing with combinative attacks in the second half. In the latter, the national teams of Germany and Argentina displayed defensive tactics in the second half that did not exist in full matches. The effect of five contextual variables on the playing styles (type of grass, stage of the competition, coaches, changes along years, and opponent’s style) was explored in just one study for each variable. Andersson, Ekblom, and Krustrup [3] observed fewer sliding tackles and more short
passes during games on artificial turf as compared to natural grass. According to Yi, Gómez, Wang, Huang, Zhang, and Liu [63], teams performed better in passing, pass accuracy, and delivery into the attacking third when playing against direct-play teams than when playing against possession-play teams. Bekkers and Dabadghao [56] found that when teams change managers, their flow motifs change, adopting the style of the specific manager’s prior teams. Fernandes, Camerino, Garganta, Hileno, and Barreira [46] found that some teams may change defensive patterns along stages. Finally, Zhou, Lago-Peñas, Lorenzo, and Gómez [47] found that high-intensity play and offensive actions of Chinese SL increased substantially with time (from 2012 to 2017). It is impressive that the effect of formation on the playing style of teams has not yet been studied, despite that fact that it is an important element of their tactics [76], while the scientific literature has already shown that it affects physical and technical performance indicators [77,78]. The effect of other contextual variables could also be studied, such as the period of the season [79] or the market value of the teams [80].

4.3. Effectiveness

Seven papers studied the effectiveness of playing styles. Of particular interest is the distinct way in which each study expressed efficacy. Four of the studies that measured effectiveness solely addressed styles during the attacking phase. Fernandez-Navarro, Fradua, Zubillaga, and McRobert [20], used the possession effectiveness index which is a combination of the variables Expected Goals (shot location and shot type were the variables considered to calculate this metric) and Ball Movement Points that measures each ball move in a possession. Ball moves were assessed based on the risk they pose to the opponent. The results suggested that the effectiveness of game styles changes under specific circumstances and that not all contextual variables affect them in the same way. Bekkers and Dabadghao [56] measured the effectiveness of styles of play by introducing an Expected Goal Motifs model (a sequence of at least one pass that leads to a goal-scoring opportunity with a certain expectation of being converted). Drezner, Lamas, Farias, Barrera, and Dantas [43] defined two main classes-incomplete and complete penetration dynamics. Incomplete penetration dynamics were defined as those that do not reach the opponent’s last defensive line. As a result, the effectiveness was calculated according to the degree of success of the ball circulation that leads to a penetration. Finally, in Schulze, Julian, and Meyer [49] study, goals and goal scoring opportunities (GSOs) were used to measure the effectiveness of attacks. In the two papers that examined all phases of the game, Castellano and Pic [39] based their study on the final result (win/lose/draw), while Lopez-Valenciano, García-Gómez, López-Del Campo, Resta, Moreno-Perez, Blanco-Pita, Valés-Vázquez, and Del Coso [7] based their study on the number of points obtained and the final ranking position. Only one study has dealt with the effectiveness of defensive playing styles, for which the expected goals of the opponent were used to calculate [51]. A positive aspect of the studies of Castellano and Pic [39] and Lopez-Valenciano, García-Gómez, López-Del Campo, Resta, Moreno-Perez, Blanco-Pita, Valés-Vázquez, and Del Coso [7] is that they did not merely use the data of the provider companies (e.g., InStat, Tracab, Mediacoach), but instead they created new variables (from the already existing ones), taking into account the interaction between the two teams or by applying normalizations of the data based on the time of possession, the number of attacks, etc.

5. Conclusions

This article provides a critical analysis of the research that relates to the playing styles of soccer teams and meets specific criteria. It was found that the data from a single competition is insufficient to generalize the outcomes of the studies. Researchers should place a greater emphasis on the practicability of their findings for team coaches. Visualization of conclusions is a method that helps in this direction. In addition, the utilization of extra factors and, in particular, the invention of new ones (with more tactical significance) will assist in identifying alternative play styles and evaluating their performance. AI and factor
analysis can provide useful information in the identification of playing styles, but the latter requires attention to the interpretation of the components. More research on playing styles that focuses on the actual application of findings on football fields is therefore required.

**Author Contributions:** Conceptualization, S.P., S.M., C.K. and D.T.; methodology, S.P., S.M., C.K., T.T. and D.T.; validation, S.M., G.G and D.T.; writing—original draft preparation, S.P., C.K., M.P. and D.P.; writing—review and editing, S.M., TT, G.G. and D.T.; visualization, S.P. and D.P.; supervision, S.M., G.G. and D.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


8. Gómez, M.-Á.; Mitrotasios, M.; Armatas, V.; Lago-Peñas, C. Analysis of playing styles according to team quality and match location in Greek professional soccer. *Int. J. Perform. Anal. Sport* 2018, 18, 986–997. [CrossRef]


15. Chassy, P. Team play in football: How science supports FC Barcelona’s training strategy. *Psychology* 2013, 4, 7. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.