DEVELOPING A DECISION-MAKING FRAMEWORK FOR PLAYER RECRUITMENT IN EUROPEAN FOOTBALL CLUBS

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities.

2022

ORBAY ÜNSOY

Alliance Manchester Business School

Management Sciences and Marketing Division

TABLE OF CONTENTS

Word Count: 37,878

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This study sets out to build a conceptual decision-making framework based on the objectives pursued in player recruitment and to investigate player typologies in European club football. The existing literature in talent identification has not dealt with the wide range of sporting, organisational and financial aspects that concern the recruitment of first-team players in European football. Further, conventional playing positions provide a rough idea of the tasks that players are expected to perform. Yet, the increased pace and demands of today’s football call for an enhanced understanding of the technical skill sets and functions of players that go beyond the player positions.

Using value-focused thinking, four fundamental and nine means objectives pursued by key decision-makers of European football clubs in player recruitment are identified and analysed. These objectives are later presented in a means-ends network that visualises the studied decision-making context's multidimensional nature. Following that, the technical performance indicators of 920 players fielded in the top three European football leagues during the 2017/18 season are examined via principal component analysis and agglomerative hierarchical clustering. In the principal component analysis phase, three off-possession and four on-possession skill sets are introduced to characterise the contributions of footballers. Later, six major clusters and nineteen sub-clusters are identified that can be utilised to categorise technical player functions. It is then illustrated how the framework could be beneficial in supporting the recruitment decision-making of football players in a case study.

The significance of this study is that it informs the theoretical understanding of the decision-making context in the recruitment of first-team players in European football. Furthermore, examining player skill sets and functions offers a deeper theoretical perspective beyond the conventional player positions for further research in sports management and sciences.

DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENTS

regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses

ACKNOWLEDGEMENTS

I am truly grateful to Dr Richard Allmendinger and Dr Yu-wang Chen, for their guidance, patience and cooperation throughout the supervision process.

Special thanks go to Rotimi Ogunsakin and Tulin Cengiz for being very supportive and helpful.

And of course, I would also like to thank my friends and family members that have supported me all the way.

1. Introduction

1.1. Background to the Thesis

For professional football clubs, player recruitment decisions account for the most important strategic decisions taken because playing talent is a key source of sustainable competitive advantage (Rossi et al., 2016). The ability of a club to deliver desirable sporting results and generate revenues is highly dependent on the quality of its player recruitment (Baroncelli & Lago, 2016; Szymanski, 2015). Since 2013, clubs have been restricted by the UEFA Financial Fair Play (FFP) Regulations to investing in their playing squad in line with their annual revenues (Franck, 2014; Peeters & Szymanski, 2014; Plumley et al., 2017). This development has benefited clubs in terms of encouraging more responsible and prudent spending on player wages and transfer fees (Plumley et al., 2017). Yet, it also increased the pressure on club decision-makers to improve the effectiveness of their player recruitment activities and processes.

During the financial year 2018, the club revenues within the UEFA territory increased by 5% to reach €21 billion (UEFA, 2020a). Despite the effects of FFP, the total wage costs went up by 9.4% (UEFA, 2020a). Among the cost factors in clubs, player wages amounted to €10.3 billion and absorbed 49% of revenues generated, and thus remained to be the largest cost driver and risk factor of a football club (UEFA, 2020a). Therefore, it is vital for football clubs to build systematic processes and structures to effectively manage player costs. Here, a major challenge for decision-makers in many European clubs is that the European club football competition is dominated by the wealthiest few clubs who generate the largest revenues and afford the best playing talent (Deloitte Sports Business Group, 2021). As a result, between 1996 and 2015, the percentage of the top three teams winning the title was 56, 21 and 8%, respectively, in 52 European leagues (Szymanski, 2015).

Since the technical and physical demands of football have considerably increased (Barnes et al., 2014; Kelly et al., 2021), one of the most vital changes in football has been the emergence of multifunctional players (Wilson, 2018). As pointed out by Michels (2001), all eleven players contribute to defending, moving the ball and attacking in modern football. Hence, it is no more sufficient for a defender to be only competent in defending. Likewise, attackers are expected to help with the defensive side of the game. Therefore, the emergence of multifunctionality further complicates the challenging task of comparing and assessing a large pool of talent to pick from.

In the area of decision-making in player recruitment, academic research has largely focused on talent identification with regards to youth football players (Johnston et al., 2018; Larkin & O'Connor, 2017; Larkin & Reeves, 2018; Río, 2014; Roberts et al., 2019). These studies particularly examine the sporting performance aspects during the very early stages of the careers of players. On the other hand, the sports science and analytics literature on technical performance in football gives an insight into the performance criteria used to assess players (Barnes et al., 2014; Barron et al., 2018; Brechot & Flepp, 2020; Dellal et al., 2011; Fernandez-Navarro et al., 2016; Lago-Peñas et al., 2010; Liu et al., 2016; Rampinini et al., 2009; Yi et al., 2018; Zambom-Ferraresi et al., 2018). Yet, these studies do not reflect the complexities behind the recruitment decisions made concerning the first team players. Therefore, a holistic conceptual framework is much needed regarding the decision-making context behind player recruitment in first teams of clubs and the sporting, financial and organisational factors that concern these crucial decisions.

In addition to possessing competent playing talent, a major source of competitive advantage for a football club is the successful implementation of a football identity (Honigstein, 2015; Michels, 2001; Perarnau, 2016; Wilson, 2018). To be able to play football in a particular style, teams require specific types of players in defence, midfield and attack. Therefore, player recruitment in top European clubs begins with the identification of the desired player characteristics for the playing positions (Ancelotti et al., 2016; Lawrence, 2018; Soriano, 2011). There is a substantial amount of literature produced on the playing styles of football teams (Fernandez-Navarro et al., 2016; Gómez et al., 2018; Kempe et al., 2014; Lago-Peñas et al., 2017; Yi et al., 2019). Yet, there is little research on the player typologies that exist in top-level European football (Aalbers & Van Haaren, 2018; Bekkers & Dabadghao, 2019; Peña & Navarro, 2015). Two of these studies merely focus on the passing behaviour of players (Bekkers & Dabadghao, 2019; Peña & Navarro, 2015) and therefore do not cover the other skill sets that can help to evaluate players.

The following section introduces the research objectives and questions that guide this thesis. After laying out the research design and methods adopted, the contributions and the structure of the thesis are outlined.

1.2. Research Objectives and Questions

The competitive and financial challenges that football clubs face when building a playing squad have resulted in a decision-making context that involves multiple stakeholders such as owners, directors, coaches, talent scouts and data analysts (Ancelotti et al., 2016; Anderson & Sally, 2013; Lawrence, 2018; Parnell, Groom, et al., 2018). According to Kelly and Harris (2010), the increased participation of owners and directors in player recruitment leads to a high level of distrust and hostility between the executives and first-team managers, who used to be individually responsible for identifying and signing players. Hence, there is an increased need to bring together the sporting, organisational and financial aspects that concern the decision-making context in the recruitment of first-team players in European football. Therefore, the first objective of this study is to build a conceptual decision-making framework based on the goals pursued in the recruitment of first-team players.

The second research objective addressed in this study concerns the identification of player typologies in football based on their performances. Regarding the search for new players, Soriano (2011) emphasises that the recruitment efforts start with the clarification of player needs in qualitative terms, such as an attack-minded right-sided defender or a forward competent in aerial duels. Yet, the playing positions of players act as rough conceptualisations and therefore fail to portray a player's technical skill set or playing style.

To achieve the research objectives outlined, the following research questions will be addressed:

- How can a decision-making framework be built to guide player recruitment activities and processes in European football clubs?
- How can the technical skill sets of players competing in European football clubs be conceptualised?
- How can the technical functions of players competing in European football clubs be identified?

In this thesis, a conceptual decision-making framework is built based on a set of sporting and financial objectives that concern player recruitment decisions in European football clubs. Following that, a data analytical model is proposed to derive the technical skill sets and functions of players competing in the top three European leagues: English Premier League, German Bundesliga and Spanish La Liga. This model relates closely to the decision-making framework presented in the third chapter because two of the fundamental objectives presented in the decision-making framework are to maximise sporting results and to effectively use football identity. The fit between a club's playing philosophy and player characteristics is key in building a playing squad (Ancelotti et al., 2016; Keller, 2014).

Evidently, the data-driven categorization of players based on technical skill sets and functions can play a crucial role in recruiting the most suitable players for a given football identity.

1.3. Research Methods

In the third chapter, a decision-making framework is proposed based on the first research question formulated in this study. To be able to do this, the chosen method is value-focused thinking (VFT), which adopts a systematic approach to decision making by conceptualising a decision context based on the fundamental and means objectives (Keeney, 1996a). Later, the dependencies between these objectives are visualised based on a means-ends network (Keeney, 1996a). Here, the concept of value is flexibly used as it can be represented by decision objectives, criteria, preferences and/or objective functions (Keeney, 1999b).

In the fourth chapter, firstly, principal component analysis (PCA) is applied to the player performance data set with a view to deriving technical skill sets. PCA is a widely used dimensionality reduction method that generates a smaller set of variables while aiming to preserve the maximum variance within a data set (Jolliffe, 2002). The produced latent variables, called principal components are linear combinations of the original variables (Jolliffe, 2002). To derive player clusters based on the PCA scores and to build the player categorisation model, agglomerative hierarchical clustering (AHC) is deployed. AHC is a clustering method that helps to group observations based on their similarities and to visualise an underlying hierarchy of groups and sub-groups (Han et al., 2011).

1.4. Research Design

Figure 1.1 summarises the research process that is deployed to tackle the research questions and achieve the research objectives. The first research method utilised, VFT, is a research method that has been deployed in a variety of corporate, government, military and non-profit problem domains characterised by complex and dynamic environments (Parnell et al., 2013).

This method is suitable for the decision-making context of this study for several reasons. The clarification of decision objectives before moving to the decision alternatives enables a larger set of options, more insightful decisions and a long-term consideration of the potential implications surrounding the decision objectives (Selart & Johansen, 2011). In European football, many clubs face organisational and financial problems due to focusing on a limited pool of talent and/or overspending on players to achieve immediate sporting success. Additionally, their recruitment processes may have an overdependence on the player portfolios and lists introduced by player agents (Maguire, 2019; Rossi et al., 2016). Therefore, the application of VFT to player recruitment can help to enhance the quality of organisational decision-making in European club football.

Another major benefit of VFT is that it allows for improved communication and understanding within teams by facilitating the involvement of different viewpoints that may contribute to the formulation of decision objectives, constraints and implications (Barclay, 2014). As stated in the formulation of the research problem, player recruitment in top-level football depends on a collaboration of experts from diverse backgrounds (Ancelotti et al., 2016; Anderson & Sally, 2013; Lawrence, 2018; Parnell, Groom, et al., 2018). Hence, the utilisation of VFT in player recruitment can help to reflect the multidisciplinary nature of this decision-making context.

In the fourth chapter, firstly, the performance data of players is explored to understand the

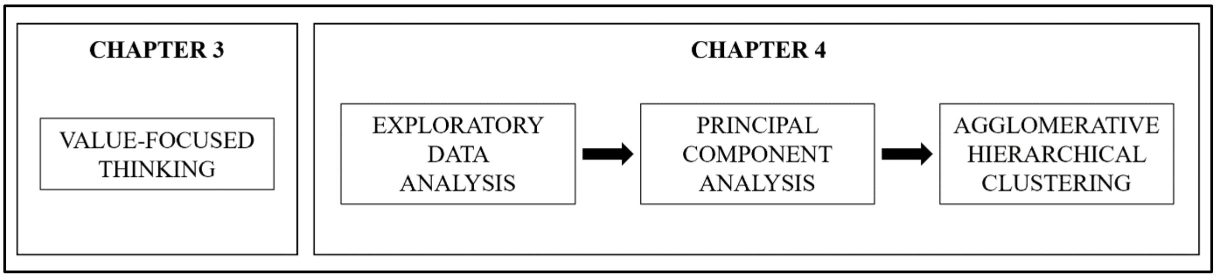

Figure 1.1 Research process framework.

nature of the performance data and to guide the methodological choices concerning which analytical techniques and performance variables to utilise. Next, PCA is applied to the player performance data to derive principal component scores that represent the skill sets of players. In the sports analytics domain, PCA was deployed in identifying playing styles in football teams (Fernando et al., 2015), as well as deriving key performance indicators in tennis (Gløersen et al., 2018), basketball (Bruce, 2016) and rugby (Parmar et al., 2018). The created PCA scores are later used in the AHC phase to discover the technical functions of players. AHC is an effective method for grouping data points together based on their similarities and in visualising a very large number of observations within a hierarchical tree. Hence, it is utilised to develop a player typology framework based on the player performances in the top three European Leagues.

1.5. Contributions of the Thesis

In the third chapter, 'Building a Decision-Making Framework for Player Recruitment', four fundamental and nine means objectives pursued in player recruitment in European football are identified and analysed. These objectives have their roots in the existing literature on football management and performance. Here, the four fundamental objectives identified are: to maximise sporting results, to effectively use football identity, to minimise wage costs and to maximise net transfer fee income. Later, the fundamental and the means objectives are presented in a means-ends network that visualises the multidimensional nature of the studied decision-making context. The assembled decision-making framework combines sporting, organisational and financial aspects of player recruitment in European club football.

In the fourth chapter, 'Identifying Technical Functions of Footballers using Data Analytics', the technical performance indicators of 920 players fielded in the top three European football leagues during the 2017/18 season are analysed. In the first phase of the study, three off-possession and four on-possession skill sets were found to characterise player contributions.

Afterwards, these skill sets are utilised in the creation of six major clusters and nineteen sub-clusters that help to categorise technical player functions. Using the case study of Manchester United, it is then illustrated how this model could support the recruitment decision making of football players and building a football squad.

This study provides both theoretical and practical contributions to the football, sports and operations management literature. The proposed decision-making framework based on VFT advances the understanding of the multidimensional nature of player recruitment in European club football. Also, the derivation of player skill sets and functions offers a deeper theoretical perspective beyond the conventional player positions. Additionally, it is shown that the application of VFT, PCA and hierarchical cluster analysis can be beneficial in analysing the domains of football and sports management. The methodology and rationale behind the third and fourth chapters could readily be adapted to other sports contexts.

In terms of contributions to practice, both models can be utilised by key decision-makers for the effective management of the player-related expenditures, which constitute the biggest cost stream and risk factor in European club football. The decision-making framework proposed in the third chapter presents a holistic picture of the player recruitment context that football clubs operate in based on the outlined financial, organisational and sporting objectives. Furthermore, the constructed data analytical model provides a novel data-driven approach to categorising players on the basis of their technical skill sets and functions. Lastly, the case study of Manchester United showcases how the data-driven taxonomy can help to build a football squad via benchmarking against best practices in England, Germany and Spain.

1.6. Structure of the Thesis

Following the introduction chapter, a literature review will be provided to introduce the elements that characterise the European club football, as well as the player categories that are used to define player typologies. The third chapter will provide the conceptual decision-making framework for player recruitment, whereas the data analytical model on player categorisation will be presented in the fourth chapter. The concluding chapter will provide an overview of research findings, the implications and limitations of the study, and suggestions for future work.

2. Literature Review

2.1. The European Club Football and the Financial Performance of Clubs

Football boasts some of the most popular brands in the world and has managed to capture a very large audience thanks to the technological advancements and continued global marketing efforts surrounding top-level football. Today, the top 20 most revenue-generating clubs have yearly revenue of €9.3 billion, whereas the European football market has a total revenue of €28.9 billion (Deloitte Sports Business Group, 2020, 2021). The football income of a professional club is composed of three main income streams: matchday, broadcasting and commercial (Peeters & Szymanski, 2014). Therefore, a football club's financial position is decided by both the contextual factors such as the population and the economic indicators of the country, as well as the club characteristics such as historical performance and brand equity. In the UEFA territory and the Big Five leagues, which are composed of the highest revenue-generating leagues in England, France, Germany, Italy and Spain, the largest revenue source is the broadcast rights (Deloitte Sports Business Group, 2020; UEFA, 2020a). As argued by Kesenne (2007), there is a high positive correlation between the size of a national economy and the income generated from TV rights of its football leagues. Since 2015, Galatasaray and Zenit St. Petersburg were the only two clubs out of the Big Five leagues that could make it to the Top 20 annual revenues in Europe in 2015 and 2017, respectively (Deloitte Sports Business Group, 2015, 2017). As is the case with the Big Five countries, the size factor is substantial here as Turkey and Russia have two of the largest populations among the UEFA territory.

Table 1 shows the financial revenue streams of the Big Five leagues. Here, we see that the English Premier League (EPL) generates the highest revenues by far, with €3.45 billion, whereas Spanish La Liga is the second highest with €1.83 billion (Deloitte Sports Business Group, 2020). On the other hand, French Ligue 1 has the lowest TV income, with €0.90

billion (Deloitte Sports Business Group, 2020). In terms of reliance on broadcast rights, EPL and Italian Serie A both generate 59% of their revenues from this income stream, whereas German Bundesliga has the lowest ratio of broadcast to total revenue ratio with 44% (Deloitte Sports Business Group, 2020). For the first time in its history, a portion of the EPL games started to be broadcasted on an exclusive online platform, Amazon Prime (Geey, 2020). Since EPL sells the broadcast rights in seven different packages based on the quality and appeal of games, it allows different companies to acquire the relevant rights (Geey, 2020). Since live sports is an area that has interested platforms such as Amazon Prime, Facebook and Youtube (Atkinson, 2018), we might see a rise in exclusively online access to matches in the upcoming years.

Thanks to the relative size of the broadcast rights deal in EPL compared to other Big Five countries, seven EPL clubs have made it to the Top 20, and a total of 12 made it to the Top 30 in terms of annual revenues (Deloitte Sports Business Group, 2021). According to UEFA (UEFA, 2020a), English clubs make up 17 of the top 20 clubs in terms of broadcast revenues. A major factor that adds to this domination is the performance-based rationale behind the distribution of TV rights revenues.

Table 1 Revenue streams in the Big Five leagues in the season 2018/19 (Deloitte Sports Business Group, 2020).

Country	Broadcast	Commercial	Matchday	Total
England	3,459	1,616	776	5,851
France	901	415	201	1,902
Germany	1,483	846	520	3,345
Italy	1,460	751	284	2,495
Spain	1,831	1,023	521	3,375

The distribution of these revenues is decided by the domestic leagues or federations that are also responsible for negotiating the deals with the broadcasting candidates. While 50% of these revenues are equally split between clubs competing in EPL, 25% is distributed based on final league position, and another 25% is based on the number of live matches broadcasted in the country within a given season (The Football Association Premier League Limited, 2020). On the other hand, Italian and Spanish leagues additionally consider performance over several seasons and fan attendance in their calculations (Association of the Liga Nacional de Futbol Profesional, 2019; Galardini, 2019). This method allows the top clubs in these countries to have a sizable advantage over others with relatively fewer resources. As a result, the high-to-median ratio in EPL is 1.4, whereas this ratio is 2.6 and 3.1 in Italian Serie A and Spanish La Liga, respectively (UEFA, 2020a). Hence, the methods of distribution can significantly support or hinder the competitive balance in a given league.

The second-largest revenue stream for European competition is the commercial revenues (UEFA, 2020a). Again, EPL is the highest revenue generator here with €1.62 billion, while Spanish La Liga is the second with €1.02 billion (Deloitte Sports Business Group, 2020). From a reliance perspective of the Big Five leagues on commercial revenues, we see that Spanish La Liga and Italian Serie A have percentages of 30%, whereas French Ligue 1 has the lowest ratio of commercial to total revenue ratio with 22%. This shows that the French market might be far from saturated in terms of achieving its commercial potential.

When we examine the Top 20 revenue-generating clubs, we also see that the majority have the broadcast revenues as the highest revenue type (Deloitte Sports Business Group, 2020). However, for the top 5 clubs, Manchester United, Barcelona, Real Madrid, Bayern Munich and Paris Saint-Germain, commercial revenues constitute the largest revenue stream (Deloitte Sports Business Group, 2020). The global appeal of these clubs and their well-known names attract some of the most lucrative sponsorship deals in the world. As argued by Kuper and

Szymanski (2010), the new globalised fan tends to prioritise the European giants over their local domestic clubs. The success and visibility of these clubs in national and international competitions have led to a surge in commercial revenues from the respective leaders of their industries all over the world. For instance, the top two shirt sponsorship deals belong to Manchester United and Real Madrid for €73 and €70 million per year, respectively (Nwokolo, 2021). These deals were made with the car manufacturer Chevrolet from the United States and the airlines company Fly Emirates from the United Arab Emirates, respectively (Nwokolo, 2021). To solidify their position in the international markets and to benefit from the lucrative match appearance fees, clubs also tend to partake in pre-season tours in different regions of Asia and North America (Soriano, 2011). In 2018/19, Real Madrid generated €114 million from international and friendly matches, which amounted to 15% of their total revenues (Maguire, 2019).

In addition to having a first-team kit sponsor, some top clubs such as Manchester United, Chelsea and Liverpool also have training kit sponsors (Geey, 2020). Other prominent commercial deals include deals with sportswear suppliers such as Adidas, Puma and Nike, as well as partnerships from various industries (Szymanski, 2015). These partners and sponsors also have the chance to appear on the backdrops during pre-match and post-match interviews, as well as on the advertising boards surrounding the pitches (Geey, 2020). As the social media channels of clubs are followed by a large audience, they might also be exclusive sponsors for the content published on platforms such as Twitter, Facebook and YouTube. Additionally, the commercial revenues of clubs also benefit from their official merchandise and stadium tours (Szymanski, 2015). Despite that, a good majority of the shirt sale profits go to sportswear companies (Geey, 2020). Today clubs have the opportunity to globally distribute a wide range of merchandise through their online stores.

Lastly, matchday revenues come from the ticket sales of the matches, which are composed of season tickets and single match tickets. The sales of season tickets are vital for clubs as they allow the clubs to generate a substantial amount of income before the season starts. The stadium sizes, the prices of tickets and the ratios of capacity utilisation are key factors that determine the income from matchday operations (Szymanski, 2015). Among the Big Five leagues, EPL has the highest annual revenues with €0.78 billion, followed by Spanish La Liga and German Bundesliga, which recorded €0.52 billion in 2018/19 billion (Deloitte Sports Business Group, 2020). These two leagues have the highest percentage of matchday to total revenue ratio with 16% billion (Deloitte Sports Business Group, 2020). On the other hand, this ratio is 11% for Italian Serie A and French Ligue 1. Before the emergence of broadcast and commercial revenues, this was the main revenue stream for many European clubs, yet this is not the case anymore. For some clubs with a significant global fan base, it can be beneficial to limit the number of season tickets to an extent because football tourist fans tend to pay large sums for single match tickets and club merchandise (Maguire, 2019). For instance, Liverpool sold only 50% of their capacity for season tickets and made £1,589 per seat in 2019/20, whereas West Ham made 90% of its capacity available for season tickets and had an average price of £465 (Maguire, 2019).

2.2. The Relationship between Financial and Sporting Performance in Football

The sporting ambitions of football clubs are largely dependent on their resources. While the wealthiest few clubs tend to compete for the title and UEFA European Champions League qualification regularly, the priority of some others is to survive in the league (Carmichael et al., 2011). This is due to the fact that there is a strong correlation between the wage bill of a club and its sporting performance (Kuper & Szymanski, 2010). As a result, the number one, two and three clubs in wage spending won EPL nine, seven and three times in the space of 20 years between 1996 and 2015 (Szymanski, 2015). In 52 European leagues, the percentage of the top three teams winning the title was 56, 21 and 8%, respectively, whereas 15% of the time, a club outside the top three spenders managed to win the league (Szymanski, 2015). Evidently, there can be cases where a club overperforms its wage budget but the odds of another team winning EPL with the resources that Leicester City had in 2015/16 seem to be quite low.

Between the years 1995 and 2013, Spearman's rank correlation coefficient between league position and wage spending in EPL was 0.76, whereas it was 0.62 between 1958 and 1975 (Szymanski, 2015). The Spearman's rank correlation is a technique that can be utilised to analyse the statistical dependence between two variables, including discrete and continuous types of metrics (Tufféry, 2011). Table 2 shows the correlation coefficient between wage spending and sporting performance in the first two English tiers from 2014/15 to 2018/19. In both leagues, we see that the magnitude of the correlation between wage spending and league position is at its highest. In Championship, there seems to be a huge increase in 2018/19 compared to the previous season. This surge, coupled with the stark difference between the correlations in the two leagues from 2014/15 to 2017/18, shows that the wage spending in Championship was far from efficient during that period. Here, it is possible that the relatively professionalised player recruitment and organisational structures in Premier League were increasingly adopted by the Championship.

Table 2 The Spearman's rank correlation between league ranking and wage spending in the first two English tiers (Deloitte Sports Business Group, 2018, 2020).

League	2014/15	2015/16	2016/17	2017/18	2018/19
EPL	0.74	0.54	0.81	0.75	0.82
English Championship	0.24	0.42	0.50	0.47	0.70

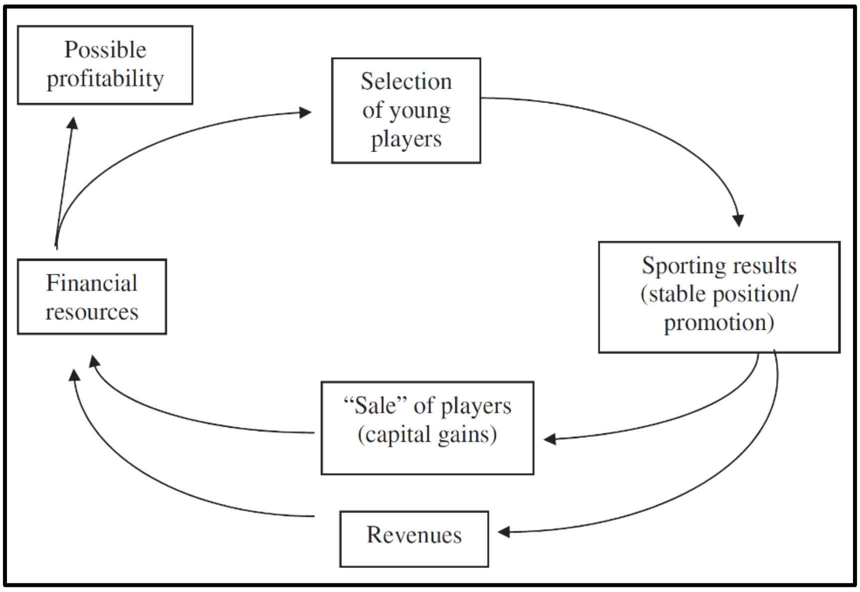

Figure 2.1 The virtuous circle in football. Reprinted from Baroncelli and Lago (2016).

Figure 2.1 shows the virtuous circle model proposed by Baroncelli and Lago (2016), which depicts the linkage between sporting and financial performance. Here, a sizable amount is needed to trigger the cycle and to form a squad that can deliver desirable sporting results and generate player sales income (Baroncelli & Lago, 2016). This figure emphasises the importance of player recruitment and development as both these objectives determine a club's potential to generate football income so that the cycle can be started again with improved resources. This model is especially relevant for clubs that are not among the leading few in a domestic league.

According to another model designed for the leading clubs, income from player sales is not included as they do not need to rely on this type of income (Baroncelli & Lago, 2016). This is because the wealthiest few clubs have the luxury of recruiting the best players available in the market. For instance, the initial transfer fee offer of Barcelona to sign Philippe Coutinho of Liverpool worth £92 million was rejected and followed by a few higher bids until a fee of £142 million was agreed (Maguire, 2019). At the time, Barcelona was the 2nd highest revenue-generating European club as Liverpool was the 9th in this ranking (Deloitte Sports Business Group, 2017). Likewise, Borussia Dortmund did not manage to keep some of their best players that were pursued by wealthier clubs even though they are currently the 12th highest-ranked club in Europe (Deloitte Sports Business Group, 2021). These transfers show that player sales income is crucial for almost every club in the world. Hence, the first model that contains this component was included here ahead of the latter one.

On the transfer fee income acquired by player sales, Szymanski (2015) argues that this type of income can be viewed as the 4th major source of income for a football club. A few major clubs in Europe have turned their competitive advantage in player recruitment and development into a constant source of income. In the 2010s, Benfica, Porto and Ajax recorded a net transfer income of €640, €350 and €300 million, respectively (Tomlinson, 2020). Other major examples that command high transfer fee profits include Borussia Dortmund and Sevilla, thanks to their directors of football, who oversee the player recruitment and sales operations. Michael Zorc and Monchi (Lyttleton, 2017). Nowadays, many clubs rely on the stream of transfer fees that they receive from wealthier clubs. Between 2014 and 2017, in Europe’s Top 15 leagues, the ratio of transfer fee income to total revenue increased from 26% to 38% (Chaudhuri, 2019). For instance, Croatian, Portuguese and Ukrainian top tier clubs had average ratios of transfer income to total revenue of 117%, 76% and 66%, respectively, in 2017 (Chaudhuri, 2019). These ratios underline the importance of enhancing the organisational capability of football clubs in recruiting, developing and selling players to ensure financial stability.

2.3. The Transfer System and the Market Valuation of Players in European Football

According to The FIFA Regulations on the Status and Transfer of Players (FIFA, 2020), a player is free to sign a contract with a club 'if his contract with his present club has expired or is due to expire within six months'. Hence, many players can only move from one club to

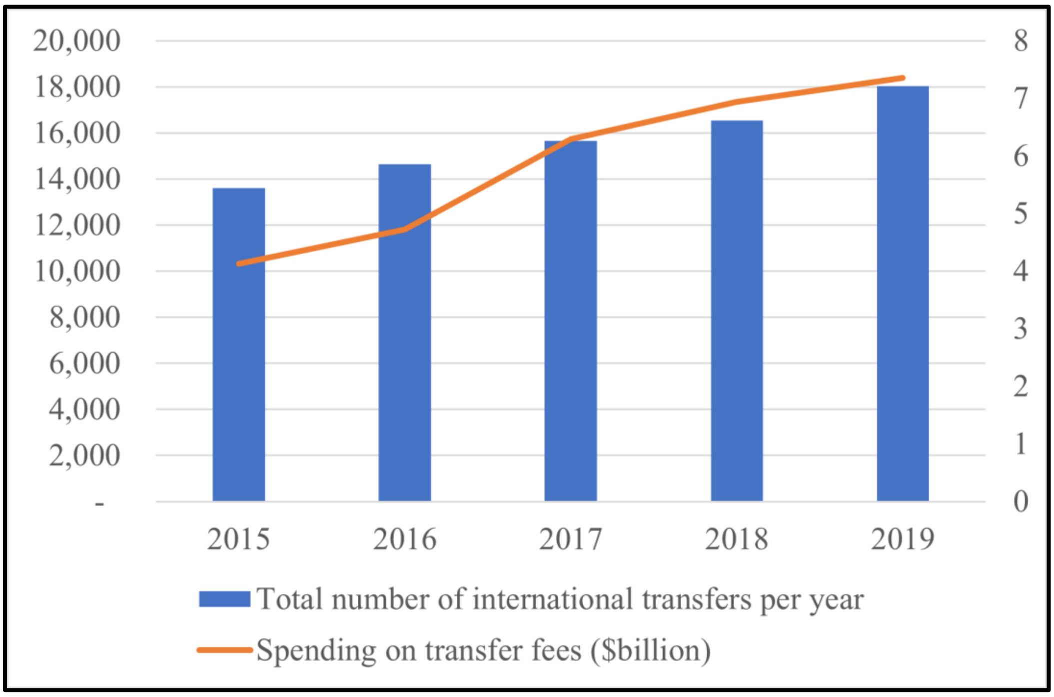

Figure 2.2 International transfer activity and fee spending in 2015-2019 (FIFA, 2021b).

another after a transfer fee is agreed between two clubs. A transfer fee is a financial compensation paid to transfer the economic rights of a contracted player to another club (Dobson & Gerrard, 1999). In 2019, total transfer fees paid amounted to $7.4 billion dollars (FIFA, 2021b). Figure 2.2 shows the positive trend in transfer fee spending and the number of international transfers between 2015 and 2019 (FIFA, 2021b). In 2019, 64.3% of all player transfers occurred when their contracts expired, whereas 11.6% included transfer fees (FIFA, 2021b).

The introduction of the Bosman ruling in 1995 was a crucial turning point in European and world football as it was not possible for a player to change clubs without a transfer fee even when his contract expired (Simmons, 1997). Furthermore, national football federations within the EU were able to limit the number of players recruited from other EU countries. However, the ruling cleared the way for the freedom of movement within the EU territory (Simmons, 1997). The removal of protectionist controls within Europe led many players to change countries and accelerated the movement of labour via cross-border transfers. As a result, there was an all-time record of 18.042 international transfers in 2019 (FIFA, 2021b). To govern the large-scale mobility of labour across the leagues all over the world, FIFA started implementing the mandatory Transfer Matching System (TMS) in 2010 (FIFA, 2021b). Here, the player registration can be transferred in a sequential way between two federations based on an online system that grants international transfer certificates (FIFA, 2010). According to Anderson and Sally (2013), this allowed for a global talent pool to be shared by European football clubs, and the most effective ways of playing football could spread across borders.

Naturally, the emergence of transfer fees triggered a stream of academic research focussing on the determinants of transfer fees paid. In addition to actual transfer fees, two other concepts have been examined in academic literature as a proxy for market valuation: annual salaries published by the German newspaper Kicker (Bryson et al., 2013; Franck & Nüesch, 2011, 2012) and market values published by the web site Transfermarkt (Herm et al., 2014; Muller et al., 2017). Transfermarkt relies on the views of its online community of around 190,000 members based on a hierarchical approach (Bryson et al., 2013). According to this approach, every community member can contribute to the discussion regarding market values, yet several members called 'judges' have the final say on the published figures (Herm et al., 2014).

The academic literature produced on transfer fees, annual salaries and market values shows that the valuation of players is influenced by player characteristics related to age, career games and international games (Carmichael et al., 1999; Dobson & Gerrard, 1999; Feess et al., 2004; Frick, 2007, 2011). These factors positively affect player valuation until they reach their peaks. The quadratic effect of these variables stems from players losing some of their key qualities towards the end of their careers. Another major factor found to affect player remuneration is the playing position of players. Forwards are the highest earners in European competition, whereas midfielders, defenders and goalkeepers follow them respectively

(Bryson et al., 2013; Frick, 2007). As the goal-scoring contributions of attack-minded players receive more attention from the media and fans than other qualities, the distribution of salaries in a squad tends to be highly skewed. In addition, height is a physical attribute that positively affects market value as it leads to higher aerial ability (Bryson et al., 2013; Herm et al., 2014).

On the effect of player performances on their transfer fees, many articles demonstrate that goals scored per game affect the market valuation of players positively (Carmichael et al., 1999; Dobson & Gerrard, 1999; Herm et al., 2014; Lucifora & Simmons, 2003; Muller et al., 2017). An important development in the evaluation of player performance is the deployment of detailed performance statistics. As stated by Carling et al. (2007), match analysis systems gather technical, physical and tactical data on individual and team performances. Due to the lack of publicly available performance data, the number of academic studies that examine the effects of various tasks on market valuation is limited. Other than goals, the performance indicators studied in this context include assists (Franck & Nüesch, 2012; Herm et al., 2014; Lucifora & Simmons, 2003; Muller et al., 2017), passes (Herm et al., 2014; Muller et al., 2017), dribbles (Muller et al., 2017), duels (Franck & Nüesch, 2012; Herm et al., 2014) and fouls (Muller et al., 2017). In several articles, expert estimations from newspapers are used as a proxy for player performance (Feess et al., 2004; Garcia-del-Barrio & Pujol, 2007). Yet, this type of data can be highly subjective and prone to cognitive biases based on heuristics.

The existing literature on the determination of transfer fees indicates that buying and selling club characteristics play a major role in the magnitude of fees, in addition to the characteristics of players moving between clubs (Carmichael et al., 1999; Dobson & Gerrard, 1999; Feess et al., 2004; Frick, 2007). This phenomenon is documented based on the financial and sporting performance of clubs prior to the transfers taking place. These performance characteristics include points won, the difference between goals scored and conceded, and average attendance figures.

2.4. Organisational Structures and Player Recruitment in European Football

2.4.1. Management Models for Sporting Operations

The football world can be defined as operating in a turbulent environment in terms of competitive, technological and legal perspectives (Draebye, 2018). The large number of games to be played within the space of nine months, the sporting and financial considerations of the decisions related to the sporting staff and the playing squad, and the asymmetry of information between football clubs regarding the decision contexts make it a necessity to be systematic and effective in their organisational structures and processes. Today’s information and communication technologies allow for increased engagement with the fan bases and media on a domestic and global scale (Maguire, 2019). The data collected on the fans can also lead to more effective marketing and public relations efforts. Yet, this can also lead to more pressure for the coaches and players to perform immediately as the short-term expectations of fans can lead to short-termism within the clubs as well. Another key source of pressure is the club owners, who are focused on protecting their investments and achieving sporting results (Parnell, Widdop, et al., 2018).

A commonly observed consequence of these pressures is the quick dismissal of the first team managers/head coaches. In the season 2019/20, the average lifespan of a manager dropped to an all-time low record of 1.16 years in the top four tiers of English football, down from approximately three years in 1992 (Smith, 2020). Hence it can be said that the first team managers/head coaches are the first to blame in the face of undesirable sporting results. As pointed out by Bridgewater (2016), it is very rare to observe a similar trend in other business sectors. Naturally, this level of job insecurity can potentially lead managers/head coaches to ignore the long-term consequences of their decisions with regards to player recruitment and selection. It also might lead to a high player turnover and increased player acquisition costs since the new manager/head coach may have a different viewpoint on the players he inherits.

Having studied the managerial dismissals in Italian Serie A during 12 seasons, Paola and Scoppa (2011) found out that managerial changes do not have a significant effect on performance within the same season and that they can serve to shift the blame away from other organisational problems within clubs in the face of frustrated stakeholders. Distinguishing between short-term and long-term results, (Hughes et al., 2010) concluded that dismissals may provide a short-term boost yet have a negative impact on long-term team performance. On the manager dismissals in EPL, it was found that the changes for the bottom half clubs were more effective than the top half of clubs (Flint et al., 2016). Yet, it was also suggested that this should not be taken as a direct cause and effect, and further factors impacting performance must be considered before the last resort of sacking the manager/head coach (Flint et al., 2016).

To ensure continuity in the sporting operations, a prominent model deployed in European football is the sporting director model, which is engrained in the football cultures of countries such as Germany, Holland, Italy and Spain (Anderson & Sally, 2013; Parnell, Widdop, et al., 2018). Therefore, this is also called the Southern Europe approach (Jarosz et al., 2015). In this model, the sporting director oversees the various football departments regarding the first team and youth teams and aims to build the most effective strategies and processes in these areas (Parnell, Widdop, et al., 2018). Considering the average period of first-team managers/head coaches at the helm, the sporting director model helps to make recruitment decisions with the next few years in mind (Ancelotti et al., 2016).

The other prominent approach in sporting operations is called the British approach, where the first team manager is the head of the football-related departments, reporting directly to the chairman of the club (Jarosz et al., 2015). In this model, first-team managers are responsible for identifying players and conducting the transfer negotiations, in addition to the coaching and selection of players (Bridgewater, 2016). Despite going through a phase of resistance towards the sporting director model due to the traditional role of the manager overseeing all footballing processes (Parnell, Groom, et al., 2018), 15 EPL clubs employed a sporting director or equivalent in 2019/20 (Slemen, 2020). On this transformation, Anderson and Sally (2013) argue that the era of absolutist managers such as Sir Alex Ferguson, Arsene Wenger and Jose Mourinho in top football is over. While the manager model is known to place a great deal of power in the hands of one decision-maker, the sporting director model allows for sustaining a football identity, the continuity of the sporting staff as well as the links between the first team and the youth teams (Neville, 2015).

2.4.2. Player Recruitment Practices

It can be said that the management of player recruitment practices in European club football has been exposed to several changes since the beginning of the 2000s. Clearly, the differences in financial strength result in a disparity regarding the financial resources that clubs can allocate for talent scouting and recruitment. In accordance with this, large scouting networks started to be set up by the biggest European clubs, such as Arsenal and Manchester United, in the early 2000s (Elberse, 2013; Rivoire, 2011). For instance, Arsène Wenger, the former manager of Arsenal, built an international scouting organisation where 16 scouts spread around the world (Rivoire, 2011). In addition to watching football games, talent scouts work on gathering intelligence on the personalities, characteristics, habits and off-the-pitch behaviours of players (Hughes & Pearce, 2020; Rivoire, 2011). Currently, Manchester United and Manchester City employ 14 and 13 full-time scouts, respectively (Hughes & Pearce, 2020).

One recent development affecting the recruitment of players is the emergence of video scouts, who are responsible for watching football games from leagues all over Europe and the world on paid platforms such as STATS Perform and Wyscout (Stats Perform, 2022; Wyscout, 2022a). Video scouts allow clubs to save logistics costs while covering a large geographical area for player identification. For instance, Borussia Dortmund employ ten video scouts and five regional scouts to identify and assess talent (Lyttleton, 2017).

A critical stakeholder that works with clubs in player recruitment is the player agents and agencies, who consult and represent players in the transfer negotiations, employment contracts and commercial deals (Geey, 2020). As stated by Rossi et al. (2016), agents help a club to access information on potential recruits through their player portfolios and well-developed networks that include other clubs and agents. However, in some cases, the agents might be powerful enough to lead the transfer policy of a club (Maguire, 2019). This practice has the potential to create a serious conflict of interest as the personal relationships and networks of a single player agent can be a decisive factor in recruiting players.

To be able to manage the high risk involved in player recruitment, football clubs have started to benefit from data scouting and analytics in their decision-making processes (Ancelotti et al., 2016; Lawrence, 2018). This development is in line with the argument that successful performance management in sports requires both soft and hard analysis (Gerrard, 2016). Under the management of general manager Billy Beane, whose story was made public with the book *Moneyball* (Lewis, 2004), The Oakland Athletics had the fifth-highest win percentage in his first 15 years despite having the fourth-lowest average payroll among 30 teams (Morris, 2014). This anomaly has triggered a shift in sports such as baseball, basketball and football led by the new type of club owners, who come from disrupting fields where informational efficiencies can be capitalised on (Lewis, 2016).

For instance, John Henry, who runs an investment management firm, acquired Boston Red Sox and Liverpool FC and won league championships with them following a drought of 86 and 30 years, respectively (Markham, 2020). Another prominent example is Matthew Benham, who is the founder of a betting consultancy firm called Smartodds. He became the majority shareholder of English club Brentford in 2009 and Danish club Midtjylland in 2014. FC Midtjylland won three league championships, including their first-ever title (Tippett, 2017). On the other hand, Brentford have promoted to the EPL at the end of the season 2020/21.

On the market inefficiencies in sports, Billy Beane explains that 'there is a tendency of everyone who actually played the game to generalize wildly from his own experience' (Lewis, 2004). The Philadelphia 76ers' president of basketball operations, Daryl Morey, who is one of the data analytics pioneers in basketball, states that Beane's data-driven approach allows a club to identify players with potential that are ignored by the professional scouts because they do not fit their mental typology of a quality player (Lewis, 2016). Similarly, the usage of analytical metrics and a systematic approach to recruitment can help to mitigate cognitive heuristics in football clubs. For instance, Mills et al. (2018) analyse how playing position varies according to skin tone and show that darker skin toned players are statistically more likely to be deployed in external than central positions, which are associated with organizational skills and high technical capacity. Additionally, players with relatively more noticeable physical characteristics can be easier to recall. On this kind of heuristics, Kuper and Szymanski (2010) state that an unnamed club discovered from their scouting reports that blonde players were more likely to be identified as potential recruits.

Regarding the case of the Oakland Athletics and their data-driven approach, Gerrard (2016) explains that they used a different metric to assess the ability of hitters because they found it to be highly correlated with winning percentages. While the traditional scouts only focused on hitting, the club used an on-base percentage based on both hits and walks and capitalised on the relevant market efficiency (Gerrard, 2016). In football, the utilisation of logistic regression led to the conception of expected goals, which quantifies the likelihood of a shot being a goal based on the characteristics of thousands of shots taken (Caley, 2015). Logistic regression is a statistical analysis technique that predicts an outcome based on a set of independent variables that can be continuous or discrete (Tufféry, 2011). Therefore, this method can utilise binary variables such as 0 and 1 as dependent variables, which correspond to non-goal shots and goals. Thanks to this method, expected goals indicate the likelihood of a shot becoming a goal between 0 and 1 based on shot details such as distance to the goal, angle to the goal, as well as the qualitative characteristics such as shots taken by foot or head (Caley, 2015). In assessing team strength, expected goals have been found to be a superior indicator of a team's performance as it measures the quality of chances a team creates and concedes, therefore representing a more solid basis than actual goals that might occur due to randomness in a low-scoring sport like football (Brechot & Flepp, 2020). Accordingly, utilisation of expected goals and expected goals assisted along with actual goals and assists allows for a more reliable analysis of a player's potential contributions in terms of attacking performance.

For instance, Ian Graham, the head of research at Liverpool FC, states that they utilised an expected goal model while assessing Jurgen Klopp's performance as the head coach of Borussia Dortmund (Schoenfeld, 2019). In that season, his club was near the relegation zone, yet the model suggested that they were one of the best teams in the league based on the quality of chances created and conceded Schoenfeld, 2019). In European football, Brentford and Liverpool are two of the few examples that have a specific department that focuses on developing insights based on data analytical models. Also, Arsenal acquired the U.S.-based analytics company StatDNA for £2 million when Arsene Wenger was the manager of the club (Hytner, 2014).

2.5. Player Categorisation in European Club Football

The five major outfield positions in today's competitive football are central defenders, external defenders, central midfielders, external midfielders and central forwards (Bradley et al., 2009; Dellal et al., 2011; Di Salvo et al., 2007). The positional framework by Dellal et al. (2011) further groups central midfielders into defensive and attacking ones. In databases, such as Wyscout (Wyscout, 2022a) and Transfermarkt (Transfermarkt, 2022), we observe three lines in the central midfield: defensive, central and attacking. These location-based player specifications are also referred to as No.6, No.8 and No.10, respectively. However, labelling the players as defensive and attacking based on how close to a team's goal they operate represents a problematic attempt to conceptualize players. In response to these rough simplifications, numerous derivations of shirt numbers have been observed. On his role in Manchester City FC, Kevin de Bruyne explains that he plays as a 'free eight' with a lot of movement (Wilson, 2016). Tottenham Hotspur FC manager Mauricio Pochettino calls two of his players, Dele Alli and Christian Eriksen 'mobile No.10's who interchange positions, driving to areas wide and forward depending on the context (Balague, 2017).

These new attempts can be largely attributed to the increasing fluidity and pace in top-level European club competitions. For instance, sprinting distances and the number of passes increased by 35% and 40%, respectively, in the EPL between 2006-07 and 2012-13 (Barnes et al., 2014). This trend has elevated the importance of playmakers, who act as the passing hubs that construct the game. According to *The Technical Report and Statistics 2006 World Cup* (FIFA, 2006), Andrea Pirlo redefined the role, dictating the attacking play from a deep-lying position. Although players such as Andrea Pirlo, Xavi Hernandez and Paul Scholes are held in high regard as role models today, playmakers who do not score many goals have been underappreciated by the football world. This may be the reason that no such playmaker managed to win the Ballon D'or Prize in the 2000s and 2010s until Luka Modric of Real Madrid CF and the Croatian national team was voted as player of the year in 2018 (Polden, 2018). On the evolution of playmakers, Wilson (2018) argues that Luka Modric represents the new style of playmaker that is physically strong and tactically astute, as opposed to those like Juan Roman Riquelme, who used to play as advanced playmakers who were virtually exempt from defensive responsibilities.

It can be said that the role of a playmaker used to be exclusive to a central midfielder when football was not such a fast-paced game. Yet, central and external defenders can serve as playmakers, too, allowing their teammates in advanced positions to create goal-scoring chances (Cruyff, 2016; Perarnau, 2016). Moreover, external defenders are nowadays more involved in attacking play and shot creation in the opponent's half. In 2015/16, 21 external defenders created more than 20 goal-scoring chances in English Premier League, compared to five in 2011/12 (Wright, 2017).

Another popular function-based term is striker, which is utilized for forwards with high goal-scoring output. According to Technical Report 2018 FIFA World Cup Russia (FIFA, 2018), there are two types of strikers: lone central forwards who stay deep in the opponent's half, such as Oliver Giroud and Miroslav Klose and mobile and versatile strikers like Lionel Messi and Neymar da Silva Santos Júnior. Players mentioned in the first group are also called target men as they are the focal points of attack that hold the ball for others. As for the second group, Eric Cantona and Dennis Bergkamp are classic examples of players that operated in front of the midfield while complementing the target men (Wilson, 2018). There are several alternative labels for this function, such as second striker, trequartista and deep-lying forward (Cruyff, 2016; Gullit, 2016; Wilson, 2018). Yet, the differentiation between such conceptualizations is not well-clarified.

For midfielders, a popular function-based specification is holding midfielder, particularly for central midfielders whose primary function is to protect the area in front of the defence (Wilson, 2018). World Cup-winning coach with the Brazilian national team, Carlos Alberto Parreira, identifies the holding midfielder as a vital role in giving teams balance and allowing both external defenders to partake in offence (FIFA, 2018). According to Wilson (2013), holding midfielders can be divided into deep-lying playmakers, destroyers, ball carriers and players who combine those qualities. Furthermore, the box-to-box midfielder is a commonly used role in football terminology, particularly for dynamic central midfielders who can operate actively in and between both penalty boxes. According to Gullit (2016), many clubs tend to lose games because they get the equation wrong in the central midfield. Hence, the combination of a balanced central midfield in terms of defensive and offensive capabilities is a vital task in building a squad.

The Oakland Athletics General Manager Billy Beane argues that former sports players employed by clubs as recruitment and coaching personnel can be prone to over-generalizations from their own experiences (Lewis, 2004). He also points out that thousand-dollar mistakes have been turning into multi-million ones due to the increasing interest in baseball (Lewis, 2004). This argument is relevant for the European football market and particularly challenging for many clubs that do not possess the financial strength of the few wealthiest clubs. Examining player and team effectiveness in English Premier League, Gerrard (2007) suggests that the Moneyball approach can be transferable to football and clubs with fewer resources could manage to be competitive based on a knowledge-based strategy. On individual effectiveness in terms of physical performance, Weimar and Wicker (2017) find that German Bundesliga clubs undervalue running distances. Accordingly, derivation of technical skill sets can play a vital role in detecting inefficiencies in different football markets and capitalizing on a knowledge-based strategy.

2.6. Player Categorisation via Data Analytics

In the area of sports and football analytics, there has been a growing interest in discovering player categories and styles based on their performances (Aalbers & Van Haaren, 2019; Bekkers & Dabadghao, 2019; Charles, 2016; Green et al., 2016; Gurpinar-Morgan, 2015; Peña & Navarro, 2015; Schulte et al., 2017; Zhang et al., 2017). Before going over the specific applications, the data analytical methods deployed in these studies are introduced below.

2.6.1. Data Analytical Methods

The data analytical methods utilised in these studies can be categorised into two groups: unsupervised and supervised learning. In a supervised learning approach, the class labels for the analysed observations are known in advance, whereas the label information is not pre-specified in unsupervised learning (Han et al., 2011). Among the seven methods introduced below, the first six belong to the unsupervised group. On the other hand, classification is the only technique that utilises pre-determined class labels. Among the seven unsupervised learning methods, four of these are cluster analysis techniques. Cluster analysis helps to assign a set of data observations to previously unknown groups based on their similarities and dissimilarities (Everitt et al., 2011). The similarity and dissimilarity measures are evaluated based on the characteristics of the data observations and may vary depending on the clustering techniques. In the studies that will be reviewed in the next sub-section, the following analytical techniques were utilised:

Principal Component Analysis

PCA is a dimensionality reduction method that creates a smaller set of variables out of a data set with a view to preserving the maximum variance (Jolliffe, 2002). The newly created variables called principal components are linear combinations of the original variables on the basis of the magnitude of the correlations between each other (Jolliffe, 2002).

Network Motifs

Network motifs are a technique that discovers the recurring patterns of interconnections within complex networks (Milo et al., 2002). Here, the identified network motifs combine a small number of links and act as the building blocks in a given complex network (Milo et al., 2002).

K-means Clustering

K-means clustering divides a data set into k number of clusters where k is pre-defined by the user (Everitt et al., 2011). Here, k number of centroids are formed based on the similarity of the observations so that each observation is assigned to the closest centroid (Everitt et al., 2011).

Agglomerative Hierarchical Clustering

In AHC, every data observation initially forms its own cluster and then merges into larger clusters according to the similarity measure selected (Han et al., 2011). At each iteration, the two closest clusters form a bigger cluster, and these successive merges are visualised in a tree-like structure (Han et al., 2011).

Affinity Propagation Clustering

Affinity propagation simultaneously calculates the similarities between pairs of data observations and considers each observation as a potential exemplar (Dueck, 2009). Then, the final clusters are formed with a view to maximising the total similarity between the observations and their exemplars (Dueck, 2009). This requires a message-passing procedure between all the observations until a consensus is reached.

Mean Shift Clustering

In mean shift clustering, every data observation is a potential candidate to be a cluster centroid, and the iterations are made based on the probabilities of the observations being a centroid (Comaniciu & Meer, 2002). Accordingly, the membership of the observations gradually shifts toward the data space, where the probabilities of being a cluster centroid are relatively higher (Comaniciu & Meer, 2002).

Classification

Classification is a supervised learning technique where data observations are grouped into pre-determined class labels based on the characteristics defined as independent or predictor variables (Tufféry, 2011). Here, the dependant variables are the previously known class labels (Tufféry, 2011).

2.6.2. Related Work

The first two studies in this area were presented in the annual industry event, Pro Forum (formerly known as Opta Forum), organized by the data and analytics company Stats Perform (Stats Perform, 2021). In the field of player clustering, Gurpinar-Morgan (2015) presents a study that utilized PCA scores of 16 selected variables in three phases of k-means clustering and discovered five types of defenders, seven types of midfielders and three types of forwards. Further, Charles (2016) deploys AHC on 15 performance indicators with the goal of finding a direct replacement for Andrés Iniesta. While these studies are of interest and related to the research problem that will be tackled in Chapter 4, they lack a sound motivation for the selection of variables and a rigorous experimental analysis of the results obtained.

Based on the network motifs technique (Milo et al., 2002), another line of research focuses on the exploration of player categories based on their passing behaviour (Bekkers & Dabadghao, 2019; Peña & Navarro, 2015). Peña and Navarro (2015) apply affinity propagation on the pass motifs of players and derive 37 clusters. Bekkers and Dabadghao (2019) similarly utilise the derived pass motifs in a mean shift clustering phase and produce 25 clusters. Both papers show how their models could be beneficial to the recruitment decision making of clubs based on a case study of finding an alternative for the former Barcelona midfielder Xavi (Bekkers & Dabadghao, 2019; Peña & Navarro, 2015).

On the other hand, Aalbers and Van Haaren (2019) use a supervised learning technique based on 18 pre-determined player categories for outfield players. Based on the five positional groups utilised in football (Dellal et al., 2011), they assign specific sub-groups for each positional group. For instance, central defenders have three sub-groups, whereas central forwards have four different categories (Aalbers & Van Haaren, 2019). The authors, who work for the data and insight provider Scisports (Scisports, 2021), form these categories and manually label 356 players in collaboration with the Datascouting department of the company (Aalbers & Van Haaren, 2019). This approach does not allow for player typologies that are previously unknown. Additionally, it might be possible that an external defender's performances might place him in the same group as external midfielders. Hence, the existence of positional biases limits the transitivity between five major positional groups.

In basketball, there are several studies that aim to explore player functions via clustering based on individual performance (Green et al., 2016; Lutz, 2012; Zhang et al., 2017). For instance, Green et al. (2016) deploy agglomerative hierarchical cluster analysis on NBA players and discover six main types: ball-dominant guards, defensive specialists, traditional bigs and skilled bigs. Yet, this paper does not demonstrate the clustering hierarchy, nor does it analyse the sub-groups of obtained clusters further. Among other sports, Schulte et al. (2017) performed affinity propagation clustering on hockey players playing in the NHL by analysing their heat maps and discovered eight types of players. Later, they examine two clusters in detail as case studies and present some top players and relatively unknown players from those selected clusters (Schulte et al., 2017). In hockey, all players across roles act in all parts of the field. Therefore, this clustering approach is not suitable for the purpose of this study as it is expected to generate clusters akin to positional specifications in football.

3. Building a Decision-Making Framework for Player Recruitment

This chapter presents the conceptual decision-making framework built to support player recruitment activities in European club football.

3.1. Introduction

Playing talent is the main source of capital that football clubs depend on. The wage spending of a football club is the most important indicator in determining its sporting success (Ferri et al., 2017; Kuper & Szymanski, 2010; Madsen et al., 2018; Morrow, 1999). However, some clubs draw on their capabilities in player recruitment and development and manage to overperform with regards to their budgets. In contrast, others might rank lower than their wage spending suggests. To boost their competitiveness, many European clubs engaged in irresponsible spending on players beyond their means and found themselves in financial trouble (Peeters & Szymanski, 2014). In order to limit such overinvestment, the European governing body UEFA introduced the Financial Fair Play (FFP) regulations in 2013 (Peeters & Szymanski, 2014). According to the FFP regulations, the football income of a club has to be declared on break even with its player wage expenditures and net transfer fee spending over three years (Franck, 2014). Consequently, all European clubs have budgetary restrictions in line with their revenue generation capacities.

Since successful player recruitment leads to desirable sporting results and net transfer profits (Baroncelli & Lago, 2016), the generated income can help fund future transfers or long-term projects in training facilities, stadia and youth academies. On the other hand, recruitment failures risk the financial health of football clubs and can have a negative effect on the fan bases (Szymanski, 2015). Additionally, clubs that mismanage their wage budgets face the risk of relegation to a lower division. To accurately assess the risks involved with recruitment, the decision making processes involve various stakeholders, including the CEO, the sporting director, the head coaches and the personnel tasked with identifying playing talent (Lawrence, 2018). As pointed out by Gerrard (2016), effective decision making in football requires a combination of statistical analysis and expert judgment. In line with this argument, clubs have started to utilise experts in data analytics and modelling to enhance their talent identification processes (Biermann, 2019). Hence, in the player recruitment context, there is a variety of views considered to 'identify the 'right players' at the right price for the club' (Lawrence, 2018). However, there is a lack of theoretical understanding of the dimensions that relate to the recruitment of footballers and the complexity behind a recruitment decision.

In this chapter, a conceptual decision-making framework is built on the concept of VFT and is aimed at clarifying the different dimensions and criteria concerning player recruitment decisions by introducing a set of sporting and financial objectives. VFT adopts a structured approach to decision making via the generation of alternatives in the light of specified fundamental and means objectives (Keeney, 1996b). This is in contrast with alternative-based thinking, which focuses first on alternatives and later addresses the criteria to assess the suitability of a variety of options (Keeney, 1996b). Later, the fundamental and means objectives are interlinked within a means-ends network to visualise the relationships between the various dimensions of the framework.

This study contributes to the field of decision making in sports management in four ways. First, this framework can help to enhance the understanding of the factors that affect player recruitment decisions in football clubs. Second, it is demonstrated how VFT can be utilised to construct a comprehensive decision frame in sports management and human resources management. Third, key decision-makers in clubs can work towards ensuring the effective and strategic management of player costs via the application of the framework. Fourth, in light of the recently produced literature within the domain of football analytics, this framework suggests a set of systematic and consistent criteria to compare and assess players.

In the next section, the concept of VFT and its relevance with regards to European club football is reviewed. In the third section, I use the methodology to analyse the decision-making context, identify objectives, and build a means-ends network of objectives. Finally, I conclude with the framework's practical implications, limitations, and directions for future research.

3.2. VFT-based Methodology

VFT is a decision technique that helps to conceptualise a decision context through the identification and articulation of the values that form the basis of the decision-making process (Keeney, 1996a). This technique was utilised in a variety of settings that include corporate, government, military and non-profit leaders (Parnell et al., 2013). Starting with a value assessment step that guides the decision frame, VFT helps to align the strategic goals of an organisation with the fundamental objectives that a decision-maker wants to accomplish (Keeney, 1996a). Later, the focus is on the means objectives that help to achieve the fundamental objectives. For instance, in the context of car travel, enforcing the law, educating the public about safety and requiring safety features could be viewed as means objectives that can help to achieve the fundamental objective of maximising safety (Keeney, 1996a). Following the clarification of these elements, the relevant criteria are utilised to generate suitable alternatives. This approach argues that the identification of alternatives before specifying the relevant values and objectives is a limited way to define a decision situation (Keeney, 1996b).

According to Selart and Johansen (2011), VFT allows for a broader range of alternatives, the inclusion of more innovative and insightful solutions, and a longer-term consideration of the potential consequences of a decision frame in comparison with alternative-based thinking. Leon (1999) also argues that VFT provides a more comprehensive understanding of a problem by covering more aspects than the alternative-based approach and helping to rethink the potentially avoidable constraints. Additionally, this method offers improved communication within teams and facilitates the involvement of a variety of stakeholders in decision contexts that contain multiple stakeholders and viewpoints (Barclay, 2014).

However, as many decision-makers are used to the alternative-based approach to decision-making, it might be challenging to receive initial buy-in from several stakeholders (Barclay, 2014). On this challenge, Selart and Johansen (2011) point out that VFT is relatively demanding in terms of the information processing capacity required and that it can also increase the handling cost of the ideas generated. Yet, they also argue that the best reward for employees is to contribute to an organisation's success and that VFT can boost their input if their motivation can be fostered (Selart & Johansen, 2011).

The major domains of expertise for VFT applications in academic research include defence, environment & energy, corporate and intelligence (Parnell et al., 2013). In terms of structuring the decision frame, these applications can be categorised into two groups. The first group utilises an interview process with decision-makers when specifying the relevant values and objectives. For instance, VFT was deployed to structure several strategic decision contexts at British Columbia Hydro, including customer service, procurement, environmental impact and health and safety (Keeney & McDaniels, 1992). Based on the viewpoints of the various stakeholders in the construction industry, a decision framework was formed on managing sustainability in the built environment (Alencar et al., 2017). In the context of blockchain technology, VFT was applied to develop frameworks for food subsidy distribution (Pawar et al., 2020) and disintermediation in medical tourism (Parekh et al., 2020).

On the other hand, the second group defines the decision context based on the literature available on the subject matter. For instance, Neiger et al. (2006) constructed a holistic framework by applying VFT-based processes to risk management in enterprise systems.

Manninen and Huiskonen (2019) proposed a value-driven sustainability management framework to connect organisational values with strategic decision-making processes. Based on the growing academic literature on the circular economy, Velte et al. (2018) presented a VFT-based framework that clarified the decision context regarding the concept.

Because of the sensitive nature of player recruitment in football and its essential role in creating a competitive advantage for a football club, the available literature on football management and performance is utilised to specify the decision frame. As pointed out by the Manchester City FC manager Josep Guardiola (Jackson, 2019), the scouting department represents the most important division in a football club and can be credited for 80% of a club's success. This phenomenon is a key factor that limits the creation of primary qualitative research in football on player recruitment activities. Hence, the relevant research gap in football management can be addressed through the application of VFT on the relevant domains of expertise such as finance, strategy, data analytics and sports sciences. Figure 3.1 shows the three phases of VFT-based thinking that will be applied to the research problem and the two steps that are

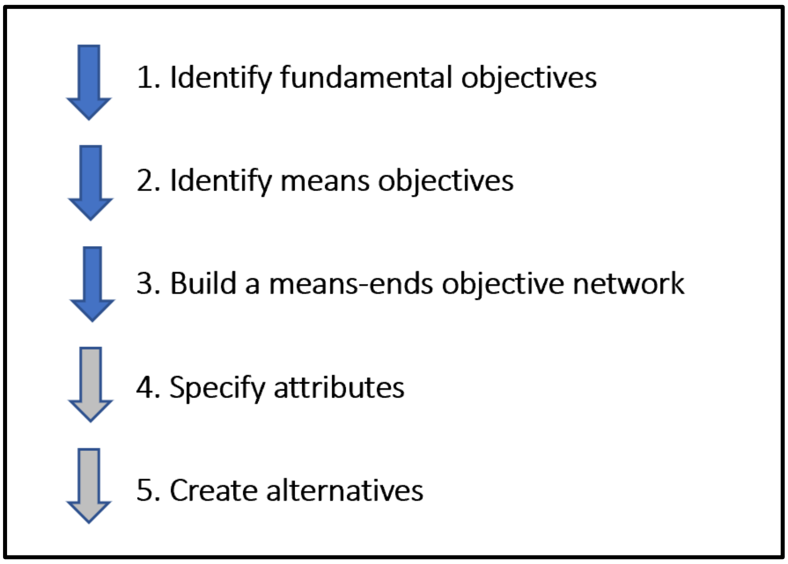

Figure 3.1 Phases of the VFT-based thinking (Keeney, 1996a).

excluded in this study. In the next section, the fundamental and means objectives of player recruitment will be identified. Later, these objectives will be interlinked within a means-ends objective network.

As pointed out by Keeney (1996b), VFT was designed to address complex problems with no clear answers. Since the practice of football club management involves a variety of stakeholders that are responsible for managing the financial and sporting performance of the business, it also typically involves different viewpoints from a player recruitment aspect. Examples of club stakeholders that are part of the decision-making in player recruitment include club owners, CEOs, sporting and recruitment directors, first-team coaches, scouts and analysts (Lawrence, 2018; Parnell, Groom, et al., 2018). As in areas such as sustainability management and circular economy (Manninen & Huiskonen, 2019; Velte et al., 2018), VFT was utilised to bring together a multidisciplinary framework that integrates different aspects of an organisation. Hence, it is suitable for the decision-making context in the recruitment of football players as well.

Another factor that makes VFT an effective tool to examine player recruitment in football is its clear advantages over the alternative-based approach. As VFT offers a longer-term view of a decision context and a larger set of alternatives (Selart & Johansen, 2011), it can play a vital role in meeting the requirements of various stakeholders. Despite the introduction of UEFA FFP rules, over-reliance on short-term results remains to be a major issue in football club management. For instance, English Championship clubs had an average wage/turnover ratio of 107% in 2018/19 (Deloitte Sports Business Group, 2020). Considering the 70% benchmark that UEFA suggests for the European football clubs, it is apparent that there is a habit of overspending in this league in the pursuit of immediate promotion to the English Premier League. Likewise, many ambitious clubs in European football take comparable risks to reach their immediate goals that are known to bring a significant amount of revenues, such as winning the league, making it to the Champions League or avoiding relegation.

Additionally, the effects of short-sightedness in football can be compounded by the over-dependence of clubs on player agents in their recruitment processes. As mentioned in the literature review section, agents help a club to access information on potential recruits through their player portfolios and personal networks (Rossi et al., 2016). In some cases, the relationship can be so strong that the club might only recruit through one agent that acts as a gatekeeper on recruitment (Rossi et al., 2016). Therefore, the alternative-based approach in player recruitment might hinder the effectiveness of decisions where clubs rely too heavily on agent-based networks. On the other hand, the relatively broader plethora of options and the review of the relevant potential consequences introduced by VFT can help to eliminate costly mistakes in the long run.

3.3. Components of the Decision-Making Framework

3.3.1. Identifying Fundamental Objectives for Player Recruitment

The first phase of VFT is the identification of fundamental and means objectives that can be linked through means-ends relationships (Keeney, 1996b). In VFT, fundamental objectives represent the ends that decision-makers wish to achieve, whereas the means objectives are utilised to reach these fundamental objectives (Keeney, 1996b). In football, every recruitment decision has its financial and sporting implications. Hence, the objectives selected are required to capture the vitality of both these aspects so that the values can be accurately articulated and consequently lead to the creation of suitable alternatives.

In order to conceptualise an objective, three characteristics are required: a decision context, an object and a direction of preference that is strived toward (Keeney, 1996b). For instance, in the context of harvesting natural resources, a company can aim to minimise the environmental impact (Keeney, 1996b). In contexts where maximising or minimising the object is not appropriate, the term optimisation is another option that can be utilised (Keeney, 1996b). For an organisation active in information and communications technology, one of the fundamental objectives identified is to effectively use e-communications systems (Drevin et al., 2007). Other examples for directions of preference include the verbs increase, reduce, provide, assure, facilitate, and control (Keeney, 1999b). Figure 3.2 demonstrates the means-ends network of the fundamental and means objectives specified and analysed in this study.

As mentioned in the introduction section, the complexity behind a recruitment decision is characterised by the stakeholders from various areas of expertise contributing to relevant processes. Since clubs need to ensure that their football-related spending translates to sporting success, they rely on a collaboration between a wide range of stakeholders, including CEOs, sporting directors, heads of recruitment, the head coach or the manager, coaches, analysts and scouts (Anderson & Sally, 2013; Lawrence, 2018; Parnell, Groom, et al., 2018).

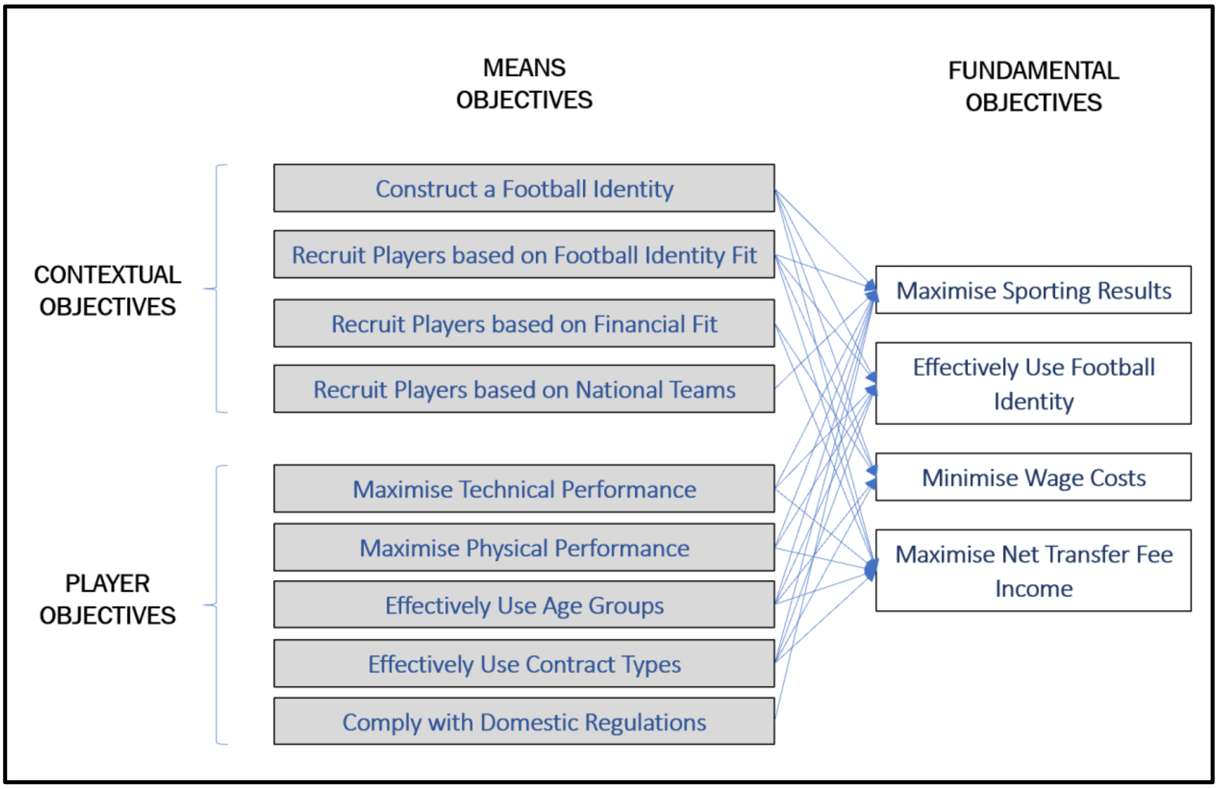

Figure 3.2 The means-ends network of the proposed decision-making framework.

In countries such as Spain, Germany and Holland, sporting directors have been the norm as the heads of football operations who are responsible for leading the youth development, performance analysis and player recruitment activities (Anderson & Sally, 2013; Parnell, Groom, et al., 2018). Therefore, they act as the conduit between the executive boards and the coaching staff in systemising player recruitment as well. Although English clubs have a tradition of employing managers who combine the duties of coaching the team as well as recruiting the team (Bridgewater, 2016), recent years have seen many EPL clubs adopting this model to allow for a sustainable football management approach (Parnell, Groom, et al., 2018).

In this decision-making framework, values and objectives considering a player recruitment decision are examined from the perspective of a head of football, who is in control of the sporting budget and therefore accountable to the executive board. Therefore, this framework would be applicable for both sporting director and manager models in managing football operations. In both cases, the key decision-maker is responsible for the strategic direction of the club and accountable for the consequences of the sporting decisions within the next few years.

However, from a club owner or chief executive officer's perspective, the framework would need to include a variety of financial metrics such as club revenues, analysis of key revenue streams and the wage costs regarding the whole club operations. In addition, fan engagement would be a key aspect in analysing the relationship of the club with the fan bases.

Moreover, the financial situation of the club is vital to the selection of the objectives. For the top revenue-generating clubs in Europe, the maximisation of net transfer fee income may not be relevant. On the other hand, for many clubs that have limited resources, obtaining short-term results might be very pressing, and the construction of a football identity may seem like a luxury to have. For instance, if a club's major aim is to avoid relegation to a lower tier during the next season, the means might justify the ends.

Table 3 summarises the mentioned objectives and the relevant literature available on player recruitment in football. In the following two sub-sections, the fundamental and means objectives of the framework are presented.

Table 3 Fundamental and means objectives for player recruitment.

Objectives	References
Fundamental Objectives	
To Maximise Sporting Results	Bridgewater (2016)
To Effectively Use Football Identity	Ancelotti et al. (2016), Michels (2001)
To Minimise Wage Costs	UEFA (2020a), (Deloitte Sports Business Group, 2021)
To Maximise Net Transfer Fee Income	(Szymanski, 2015)
Means Objectives: Contextual	
To Construct a Football Identity	Ancelotti et al. (2016), Soriano (2011)
To Recruit Players based on Football Identity Fit	Ancelotti et al. (2016), Soriano (2011)
To Recruit Players based on Financial Fit	Soriano (2011), Tippett (2017)
To Recruit Players based on National Teams	Dobson and Gerrard (1999), Frick (2007)
Means Objectives: Player	

To Maximise Technical Performance	Castellano et al. (2012), Rampinini et al. (2009)
To Maximise Physical Performance	Barnes et al. (2014), Bradley and Ade (2018)
To Effectively Use Age Groups	Anderson and Sally (2013), Elberse (2013)
To Effectively Use Contract Types	Carmichael et al. (1999), Feess et al. (2004), Frick (2007)
To Comply with Domestic Regulations	(Geey, 2020), (UEFA, 2020a)

(1) To Maximise Sporting Results

The first fundamental objective selected is the maximisation of points won in the domestic league. For the wealthiest few clubs in Europe, this objective can also include sporting success in all European and domestic cup competitions. Yet, we consider domestic league points as a common denominator that can be viewed as relevant for every club in Europe. As pointed out by Bridgewater (2016), clubs might have different league targets, such as winning the league, qualifying for European competitions and avoiding relegation. Here, an important factor to be considered is the time perspective. As stated by Manninen and Huiskonen (2019), companies may expect to reach their sustainability goals as a result of a process that takes several years. Likewise, setting goals and expectations on point maximisation can also be defined based on a few years' time as opposed to the upcoming season. Since the expectation of immediate results in sports can lead to irresponsible spending, the clarification of the time dimension can be vital in an ecosystem where decision-makers find it hard to balance the short-term and long-term interests of clubs.

As pointed out by Barros (2006), European clubs need to utilise their resources meticulously to ensure the sustainability of their existence. Therefore, the level of competitiveness achieved in one season should not jeopardize the long-term health of the clubs. Consequently, VFT can help decision-makers to structure their recruitment objectives in a sustainable way. The case of Borussia Dortmund is a good example of a club that suffered from spending large amounts on player-related expenditures based on the assumption that European Champions League qualification would be guaranteed every year. Yet, this strategy brought them to the brink of bankruptcy after failing to qualify for this competition in two consecutive years, and they had to accept a €2 million loan from their major rivals Bayern Munich to cover the monthly player wages (Honigstein, 2020).

(2) To Effectively Use Football Identity

The second fundamental objective selected is the effective use of a football identity. Three-times Champions League winner Carlo Ancelotti states that the football identity of a team is related to how a team defends and attacks and that it is ‘the key to everything on the field’ (Ancelotti et al., 2016). According to the FIFA Coach of the Century Rinus Michels, a specific football style can be functional as a team-building factor and a motivator (Michels, 2001). In the context of player recruitment, a well-defined football identity can be very useful in driving and focusing recruitment efforts. After the appointment of Jürgen Klopp at Liverpool FC, it is noted that the expectations from a Liverpool full back or midfielder became very clear, and the likelihood of making a mistake was considerably reduced (Hughes & Pearce, 2020). Likewise, the scouts working at Manchester City work based on very specific requirements that correspond to ‘Guardiola-type players’ (Taylor, 2020).

Although top names such as Klopp and Guardiola have very clear approaches to playing football, Ancelotti et al. (2016) point out that the football identity of a team also depends on other key factors such as the characteristics of players in the club and, more importantly, what is specifically asked by the key decision-makers based on the history and the tradition of the club. A prominent example of a lack of fit between the football identity of a head coach and a club was Fabio Capello at Real Madrid CF since he was sacked after winning La Liga on the grounds that he wasn't deploying an attacking brand of football (Ancelotti et al., 2016). On the other hand, Norwich City decided to continue with the head coach Daniel Farke despite finishing 14th in the English Championship because he was viewed as capable of establishing a style of football defined by the club and developing young players (Southwell, 2019). In his second year, they got promoted to the EPL(Southwell, 2019). This example also shows how the time dimension of the fundamental objectives can be crucial in organisational decision-making.

(3) To Minimise Wage Costs

The third fundamental objective of the framework is the minimisation of the player wage expenditures. As mentioned earlier, wage spending is the most important factor in determining sporting success (Ferri et al., 2017; Kuper & Szymanski, 2010; Madsen et al., 2018; Morrow, 1999). The European governing body UEFA suggests a 70% wage to revenue ratio in the FFP regulations for healthy management of club costs (Plumley et al., 2017). Despite this, clubs competing in the first tiers of France, Portugal and Turkey have an average wages-to-revenue ratio of 76, 75 and 79%, respectively (UEFA, 2020a). As the player wages constitute the biggest cost factor of a European football club (UEFA, 2020a), it is imperative that clubs minimise their wage expenditures and develop a realistic projection of what their player costs would amount to during the upcoming few years.

Among the Big Five leagues, the highest average wage to revenue ratio is that of the French Ligue 1 with 73%, whereas Italian Serie A follows them with 70% (Deloitte Sports Business

Group, 2020). On the other hand, the lowest wage to revenue ratio belongs to German Bundesliga with 54% (Deloitte Sports Business Group, 2020). The CEO of Borussia Dortmund, who was appointed when the club was close to insolvency in 2005, declared: ‘we would never again go into debt in the pursuit of sporting success, now we only spend what we have earned’ ("How Borussia went from bust to boom," 2013). The financial performance of German first tier clubs demonstrates that it is possible to leave a considerable margin between the wage to revenue ratio of a club and the suggested 70% ratio by UEFA.

Due to the Covid-19 pandemic, European football clubs had a steep decrease in matchday and commercial revenues, and it is estimated that the top 20 revenue-generating European clubs lost a total revenue of €2 billion during the seasons 2019/20 and 2020/21 (Deloitte Sports Business Group, 2021). Even though the total effects might not be fully realised for several years, it would be fair to say that the sustainable and prudent management of the player wage expenses is currently more relevant than ever. For instance, Borussia Mönchengladbach did not have to part ways with their best players thanks to building equity of €100 million since 2003 (Nielson, 2021a). Likewise, the two highest revenue-generating German clubs Bayern Munchen and Borussia Dortmund had equity of €497 and €355 million in 2019, respectively (Nielson, 2021a).

(4) To Maximise Net Transfer Fee Income

The fourth fundamental objective is the maximisation of net income from player sales. As stated by Szymanski (2015), income from player sales can be viewed as the fourth major revenue stream, in addition to the three main components of matchday, TV and commercial revenues. This is usually done through investing in players with potential that are yet to reach their peaks. Therefore, there is always a significant risk factor attached to this type of recruitment decisions. Yet, there are many European clubs who have made a habit of benefiting significantly from investing in pre-peak players and selling their contractual rights to more wealthy clubs.

For instance, Portuguese club Benfica and Dutch club Ajax gained a net income of €640 and €300 million in the 2010s, respectively (Tomlinson, 2020). In the years 2017/18 and 2018/19, French Ligue 1 and Dutch Eredivisie clubs managed to have a net transfer income to total revenue ratio of 43 and 32%, respectively (Weatherspoon, 2021). Naturally, the majority of the buying clubs for these players have been the relatively wealthier clubs within the Big Five leagues. This flow of talent demonstrates how player recruitment can be crucial in the long-term competitiveness of a football club.

To boost their matchday and commercial revenues, some clubs turn to build brand-new stadiums (Szymanski, 2015). To be able to finance these costly projects, they find themselves in a position where the prudent management of wage and transfer fee costs is essential for the club finances. For instance, London teams Arsenal FC and Tottenham Hotspur FC had similar periods prior to their stadium openings in 2006 and 2019, respectively (Rivoire, 2011; Sunderland, 2018). Under the manager Arsene Wenger, Arsenal FC had to let some of their best playing talents leave the club to help finance the stadium project (Rivoire, 2011). Likewise. Tottenham Hotspur FC under Mauricio Pochettino agreed not to sign any new players in the summer of 2018 as their project amounted to £1 billion due to added costs of Brexit (Sunderland, 2018).

As is the case with the Covid-19 pandemic, these examples show that extreme situations might radically alter how recruitment decisions are made. Such developments can lead to more opportunities for the pre-peak players and to more creative methods in organisational problem-solving. Yet, it is not an easy task to be competitive while continually minimising the player costs and maximising the transfer fee profits. If a club continuously parts ways with its best players over a few seasons, there will be a detrimental effect on the team performances as the new players might take time to fit into a new environment (Szymanski, 2015). The combination of these point to a trade-off between desirable sporting results and wage and transfer fee expenditures.

3.3.2. Identifying Means Objectives for Player Recruitment

In a sports organisation, the results are dependent on the performance of the players (Brewster & Cerdin, 2018). Yet, the means objectives within the decision context of player recruitment can be characterised in two main sections: player characteristics and contextual characteristics. While the first group is related to the key objectives related to player attributes, the latter focuses on the club context and where the players might potentially be sourced from. As every club has finite resources to allocate with regards to the player recruitment activities, having a well-targeted recruitment scope is highly critical to the success of decision-making. In the next section, we firstly review contextual characteristics, which set the basis for the creation of suitable alternatives.

Contextual Objectives

(1) To Construct a Football Identity

As pointed out by Ancelotti et al. (2016), the football identity of a club is very important in specifying the types of players to be signed for each playing position. Once the player types are clarified, it becomes easier to target a certain part of the talent pool available (Evans, 2020). In addition, the negotiations with several candidates can be carried out simultaneously (Soriano, 2011). This tends to enhance the bargaining power of a key decision-maker when conducting contract negotiations with selling clubs and player representatives. Therefore, conceptualising and developing a playing identity is vital to the development of a player recruitment framework. Based on the previous lack of football identity in Norwich City FC, sporting director Stuart Webber explains:

‘We went from different types of managers and different types of players with no real explanation. It looked a bit scattergun, and I think supporters lost belief in the club. Where are we going? What's our identity?’ (Bate, 2018).

Based on a clear football identity, decision-makers can also benefit from its pull power for potential recruits. According to Herzberg (2005), characteristics related to employee satisfaction can be analysed in two groups: hygiene and motivation factors. Here, two main hygiene factors are financial compensation and job security, whereas motivators include achievement, opportunities for personal growth, responsibility, and recognition (Herzberg, 2005). Apart from the financial considerations, other key reasons for players to change clubs are to play for better teams and to become better players. The national team director of the German Football Association, Oliver Bierhoff, describes contemporary football players as independent entrepreneurs that actively have to think about their career paths (Austin, 2017a). Hence, when the financial benefits of two offers are comparable, a club with a smaller budget might be able to beat the offer of a relatively larger budget based on a well-defined sporting proposition.

As mentioned in the previous section, the preferred playing style of the head coach/manager is very influential in shaping the football identity, yet there should be an organisational fit with the history of the club and the playing talent at hand (Ancelotti et al., 2016). Therefore, the successful implementation of a suitable football identity for a football club would depend on the involvement and communication of multiple stakeholders within a club. The qualitative description of a playing identity can pave the way for increased collaboration between the executive and sporting staff. As stated by Lawrence (2018), clarification of its goals and priorities in written form can enhance the communication processes and the effectiveness of a sports organisation. In addition, research on elite academies shows that the lack of clarity and communication between the first team and academy branches could have a negative impact on the organisational performance (Relvas et al., 2010). Hence, a clear football identity would also benefit the decision-making processes regarding youth recruitment.

Figure 3.3 demonstrates the four main playing styles that are derived based on a combination of event and tracking data in Spanish La Liga (Castellano & Pic, 2019). This conceptualisation brings together attacking and defending dimensions based on the PCA of the data set (Castellano & Pic, 2019). In terms of style in attack, this model has two options: elaborate and direct. This is in line with the literature produced on offensive team styles, which are most commonly called possession and direct play (González-Rodenas et al., 2020; Kempe et al., 2014; Yi et al., 2019). Possession style represents teams that rely on the passing quality of their players with many horizontal passes made to disrupt the shape of the opposition (Kempe et al., 2014). On the other hand, teams with a direct style largely depend on the pace of their attacks based on a smaller number of vertical passes (Kempe et al., 2014). In terms of the styles out of possession of the ball, two major styles are deep defending and high-pressure defending. These two styles separate teams that defend close to their own goal and those who are proactive in applying defensive pressures high up the pitch.

As pointed out by (Vogelbein et al., 2014), pressuring the opposition in its own half and winning the ball there can lead to better shot-taking opportunities. High-pressure defending teams try to capitalise on this by blocking the passing lanes of the opposition in a coordinated way. For many teams, this is viewed as a risky style as teams who manage to pass through the pressure can find easy chances in the space left behind high defensive blocks. Yet, the recent success of top head coaches/managers such as Marcelo Bielsa. Jurgen Klopp, Ralf Rangnick and Josep Guardiola demonstrate that a football identity based on high-pressure defending can be a key source of competitive advantage (Honigstein, 2015; Michels, 2001; Perarnau, 2016; Wilson, 2018).

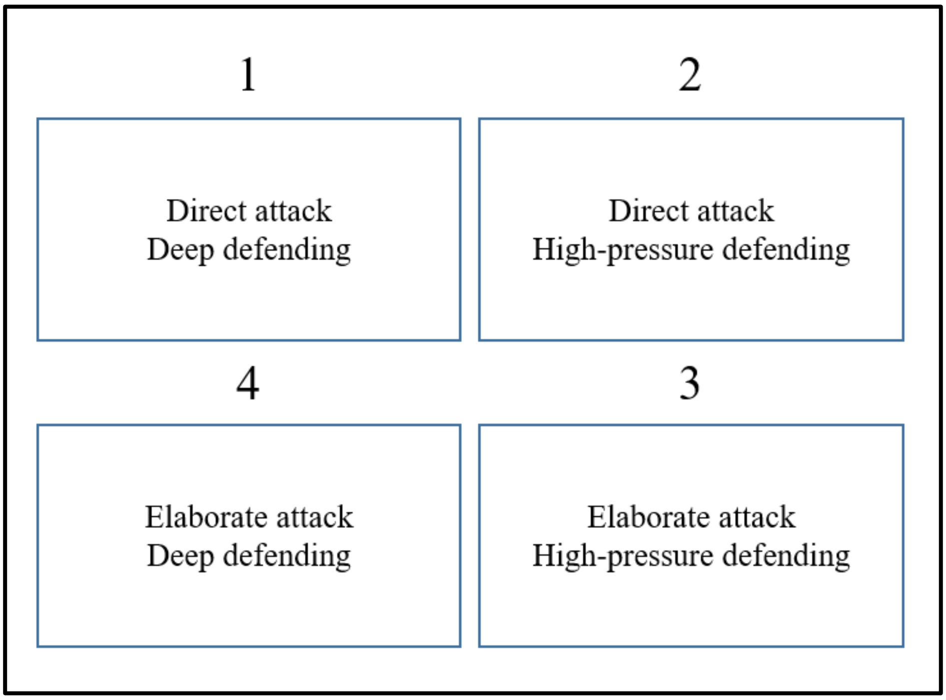

Figure 3.3 Four major playing styles derived from event and tracking data (Castellano & Pic, 2019).

However, it would be misleading to think that a team can organise itself on the pitch based on one playing style during a whole season. Research in football analytics suggests that team styles change according to the quality of the opposition and the match state, which is characterised by whether teams are winning, drawing or losing (Castellano & Pic, 2019; Fernandez-Navarro et al., 2018). When a team is one or two goals up in a match, it might switch to deep defending to hold on to the match state, and the losing team might have to utilise an elaborate attack style, regardless of the starting styles (Fernandez-Navarro et al., 2018). Likewise, there might be little use in high-pressure defending if the opponent prefers to deploy a direct style. This phenomenon requires decision-makers to consider stylistic flexibility in the construction of a football identity. If a team has one of the largest wage budgets in a league, it can continuously rely on an elaborate attack-based identity. Yet, for the other teams, it would be useful to consider a preferred style within the identity and outline how alternative styles could be utilised when necessary. For instance, the president of Barnsley FC, Paul Conway, explains that they focus on physically fit players and a direct high-pressure-based identity because they cannot afford the most talented players in the market (Nielson, 2021b).

(2) To Recruit Players based on Football Identity Fit

A major benefit of having a clear football identity is to be able to source players from clubs with similar identities. For instance, this phenomenon resulted in constant transfer flows from Ajax and Arsenal to Barcelona (Atkins, 2014; McNicholas & Crafton, 2021). The possession-based game of Arsenal under Arsene Wenger helped Barcelona sign suitable players for their football identity (McNicholas & Crafton, 2021). Also, the influence of Johan Cruyff in crafting the football philosophies of both Ajax and Barcelona has led to a natural talent pathway from Ajax to Barcelona (Gangoso, 2019). According to ex-head of player identification at Tottenham Hotspur, Rob Mackenzie, the fact that Bayer Leverkusen had a similar defensive style to theirs was a key factor in recruiting star forward Heung-Min Son (Bate & Mackenzie, 2016). Here, we see that the demanding nature of today's football dictate that the attack-minded players should be suitable for the defensive style as well. Likewise, it would be fair to argue that the attacking style of teams would be equally crucial for defenders.

As mentioned in the previous sub-section, football clubs tend to benefit from a degree of stylistic flexibility within a football identity. Therefore, conceptualising team styles would also help to recruit from clubs within different styles. For instance, a team whose identity is based on an elaborate attack style can substantially benefit from a wide forward coming from a club adopting a direct style. This is because, in the last minutes of some games where they are leading the score, this player can capitalise on the space if the other team plays based on an elaborate attack-high defending combination. Also, there might be cases where there is a misfit between the team identity and the capabilities of a player. The sporting director of VfB Stuttgart and ex-head of recruitment of Borussia Dortmund, Sven Mislintat, explain that Julien Weigl was playing for a direct style team before Borussia Dortmund, and this did not stop them from signing him to play for their elaborate attack-based identity (Biermann, 2019). Due to the lack of his suitability to the team identity, he defines Weigl as ‘a player having the misfortune of playing in the wrong team’ (Biermann, 2019).

(3) To Recruit Players based on Financial Fit

One of the main challenges for key decision-makers in football is to adhere to the financial regulations provided by the domestic federations and the UEFA. When clubs overspend with regards to their resources for immediate achievements, they also risk the livelihood of their clubs and the financial and social value they bring to the communities. A major advantage of value-based thinking over alternative-based thinking is a broader range of alternatives and a longer-term consideration of potential decisions (Selart & Johansen, 2011). In the context of player recruitment in football, it can be said that some of the most counterproductive transfer decisions occur when a club overinvests time and energy on one single player instead of simultaneously pursuing a few players that can serve the same role. As stated by Soriano (2011), having no alternative than a potential recruit makes it hard to negotiate a fair price for the buying club. Therefore, setting a broad enough domestic and international recruitment scope and generation of a variety of options in accordance with a given football identity can be a key differentiator in building a cost-effective team.

Naturally, the priority of many clubs is to recruit players from the leagues that they compete in. The familiarity of players with the league style makes them attractive options. However, the competition for these players might mean that the wealthiest clubs can outmuscle others in any given league. Therefore, the lower domestic leagues can be one of the alternative places to source affordable players. The accessibility of live broadcast of top European leagues and international club and country competitions has given sportspeople the opportunity to follow players from different football styles and cultures (Anderson & Sally, 2013). Since the 2000s, players from all over the world can be followed by clubs through paid platforms such as STATS Perform and Wyscout (Stats Perform, 2022; Wyscout, 2022a). Also, clubs that are relatively wealthy can afford to employ scouts based in various countries, who are responsible for setting up a network of contacts in their respective geographical areas with a view to unearthing playing talent. For instance, Manchester United and Manchester City employ 14 and 13 full-time scouts, respectively (Hughes & Pearce, 2020).

The combination of this recent development with substantial differences in club revenues between countries allows for opportunities for arbitrage across leagues. This opportunity is even more prevalent when the leagues compared are of similar strength. A major example in this aspect is the flow of talent from the German Bundesliga to the EPL in recent years. These two leagues were the No.1 and No.3 highest revenue-generating leagues with annual revenues of €3.50 and €5.85 billion in 2018/19, respectively. For instance, Manchester City signed Kevin de Bruyne, Leroy Sané and Ilkay Gündoğan and for €76, €52 and €27 million, respectively (Manchester City - All Transfers, 2021). Also, Chelsea signed German international players Timo Werner and Kai Havertz for €80 and 53 million, respectively (Chelsea FC - All Transfers, 2021). Here, we see that the financial disparity and the similarity of competition strength between two leagues led to a series of lucrative transfers.

Likewise, club decision-makers can analyse the wages spent in a given European league and specify a geographical scope to focus club resources and activities on. UEFA publishes its benchmarking report every year and covers the average wage budget of every UEFA country (UEFA, 2020a). Therefore, it would be possible to have an idea about the financial situation in these leagues. For instance, Brentford FC is a prominent example among English clubs that have taken advantage of arbitrage opportunities by recruiting from other European countries, including Denmark, France, Holland, Germany and Spain (Tippett, 2017). Through the success of several signings, these clubs also benefited from the generation of net transfer fee income, which is one of the four fundamental objectives of the framework. Hence, it can be argued that openness to recruiting from abroad in their approach enabled these clubs to increase their organisational competitiveness.

Additionally, financial and sporting indicators can be combined at a club level to analyse which teams are overperforming with regards to their budgets. Research in transfer fee valuation of footballers suggests that transfer fees are positively affected by the sporting achievements and financial strength of the selling club (Dobson & Gerrard, 1999; Feess et al., 2004). However, desirable sporting performance does not always lead to sporting results. As stated by Brechot and Flepp (2020), there can be a significant gap between a team's performance in terms of finding and conceding quality chances and the results in a low-scoring sport such as football. As mentioned in the literature review section, this is a phenomenon that can be captured through the expected goals framework. By combining this type of data with financial data regarding clubs, it can be possible to specify the teams that potentially field players that are undervalued by the market.

(4) To Recruit Players based on National Teams

As is the case with club competition in Europe, competition of national teams can help to generate recruitment alternatives. Within the domain of the world's football governing body, FIFA, there are six confederations that represent Asia, Africa, Europe, South America, Oceania, and North, Central America, and the Caribbean. In addition to the matches that decide the qualifiers of the FIFA World Cup, these clubs partake in their own major competitions such as European Championships and Copa America and their respective qualifiers. Additionally, they organise friendly matches where they can get ready for competitive matches. These games bring together the most capable players from each country and give these players the platform to accumulate experience in a competitive environment. Evidently, international caps have a positive impact on the market valuation of players (Dobson & Gerrard, 1999; Frick, 2007).

The competitions above are also referred to as A teams of nations all over the world. In addition to these, nations compete in similar organisations in a wide range of age levels starting from under-21s and under-19s. Currently, there are restrictions in Europe with regards to the movement of players under the age of 18 (FIFA, 2020). Therefore, it can be said that the combination of A, under-21 and under-19 national teams could characterise an appealing way to identify quality players on a comprehensive scale. FIFA utilises a points-based system to rank national A teams and publishes FIFA World Rankings on its web site every month (FIFA, 2021a). The procedure for this system considers the ranking of the opposing team in weighting the results and can be used to group national A teams based on strength.

Player Objectives

The growing technical and physical demands of the game have certainly led to more performance pressure for football players. The stakeholders of the football ecosystem, the owners, the executive directors, the fanbases and the media have different degrees of effects on the environment that the staff and players are expected to thrive in. The combination of these factors makes it a necessity to develop a structured approach to assessing performance. In football, performance can be analysed through technical, tactical and physical measures, while psychological factors can also be taken into account (O'donoghue, 2009). Since the tactical assessment of players is highly context-specific based on the tactical roles and instructions given by the coaching staff, we next focus on the two more quantifiable aspects of player performance: technical and physical.

(1) To Maximise Technical Performance

Technical performance is the most important indicator of success in football performance analysis (Castellano et al., 2012; Rampinini et al., 2009). As it is the most commonly available type of data, the majority of sports science literature on football performance is based on technical performance. Table 4 summarises the event variables studied in a selection of studies that analyse team performance (Barnes et al., 2014; Dellal et al., 2011; Fernandez-Navarro et al., 2016; Lago-Peñas et al., 2010; Liu et al., 2016; Rampinini et al., 2009; Yi et al., 2018).

Based on these variables, it can be argued that there are seven major tasks that players are expected to perform. These variables can be further grouped into two sub-groups: in possession and out of possession. According to Jones et al. (2004), ball possession can be described as having sufficient control of the ball to manipulate its direction purposefully. Here, crosses, dribbles, passes and shots constitute the variables in possession of the ball, whereas ball recoveries and interceptions characterise the tasks out of possession. Furthermore, duels contain both types of variables since offensive duels are carried out in possession of the ball whereas players partake in defensive, loose ball and aerial duels when out of possession.

Table 4 Key technical performance indicators in football.

Tasks	References	Related variables
Ball recovery	3, 5, 7	Ball recovery, dangerous opponent half recovery, opponent half recovery
Cross	3, 4, 5, 6, 7	Successful cross
Dribble	1, 5, 6, 7	Long dribble, successful dribble
Duel	1, 2, 5, 6	Aerial duel won, defensive duel won, loose ball duel won, offensive duel won
Interception	5, 7	Interception
Pass	1, 3, 4, 5, 6, 7	Assist, expected assist, successful back pass, successful forward pass, successful key pass, successful linkup play, successful long pass, successful pass, successful pass to final third, successful penetrative pass, successful through pass, successful vertical pass
Shot	1, 3, 4, 5, 6, 7	Expected goal, goal, head shot on target, shot on target

References: (1) Barnes et al. (2014); (2) Dellal et al. (2011); (3) (Fernandez-Navarro et al., 2016); (4) (Lago-Peñas et al., 2010); (5) (Liu et al., 2016); (6) (Rampinini et al., 2009); (7) (Yi et al., 2018).

To gain an insight into players' capabilities, the success rates of these actions, their locations on the pitch and the relevant game context are analysed based on the data collected by companies such as Wyscout, Stats Perform and Statsbomb (Stats Perform, 2022; Statsbomb, 2022; Wyscout, 2022a).

According to Michels (2001), a combination of factors such as durations of matches, team sizes, continuous action, and the necessity to execute quick decisions determine the complexity of football. Recent years have seen the game being played at a much faster pace compared to previous decades thanks to the development and professionalisation of sports sciences and talent development. Nowadays, all players, including defenders, are required to display excellence on the ball to be able to play a high-tempo game. In parallel with this development, attackers are expected to help the defensive organisation of teams. As pointed out by former Dutch national team, Ajax and Real Madrid head coach, Leo Beenhakker, there is no place in football for players who do not perform defensive tasks since ‘the opposition has so much quality they'll always find a teammate in space’ (Barbier, 2012). The combination of these factors increases the risk involved in player assessment activities since different aspects of technical performance are relevant for all positional groups.

A key advantage for a player who is comfortable in carrying tasks in and out of possession is to be able to cover different positions on the pitch. As stated by Lyttleton (2017), a central defender who can play as a wide defender or, a wide midfielder who can also play in the centre, has the potential to be highly coveted by football clubs. Since the match state and the quality of the opposing teams can lead to stylistic switches during games, this versatility could be highly beneficial for a football club. Also, decision-makers can choose to have fewer players in their squads than the previous year if they manage to recruit versatile players that can cover multiple positions. This could represent a useful means to minimise the wage budget of a club.

In recent years, the utilisation of data analytics in football leads to the development of new metrics to help assess technical performance. One model that became public knowledge is the expected goal framework introduced in the literature review section. Thanks to this framework, it is possible to examine the quality of goal scoring opportunities that players are involved in via a logistic regression model based on the characteristics of thousands of shots taken (Brechot & Flepp, 2020; Caley, 2015; McHale, 2018). Accordingly, a player might find good goal scoring chances yet not score, or vice versa. Hence, the addition of expected goals and assists to variables of actual goals and assists can lead to a more reliable analysis of attacking performance (McHale, 2018). For instance, if a player records a good number of expected assists but a low number of assists in a season, this shows that he managed to provide scoring opportunities for his teammates and that these chances were not converted to goals. Therefore, he can be relatively more competent in chance creation than a player that has recorded the same number of assists based on a much lower number of expected assists. This example shows that chance creation skill can be captured well by a combination of assists and expected assists. Additionally, it is possible to analyse how many goals a goalkeeper was expected to concede and compare this with the goals he conceded (Yam, 2019). In such cases, club decision-makers can benefit from the decision insights derived from data analytical models.

Another emerging aspect of performance analysed in detail is ball progression. Recently, analytical models have been developed to credit actions performed on the ball depending on the value they bring in the attacking sequences of teams (Austin, 2019; Decroos & Davis, 2020; Singh, 2019). For instance, the concept of expected threat examines how a pass changes the likelihood of a sequence leading to a goal in the next five seconds (Singh, 2019). Likewise, Liverpool FC uses a goal probability added model where every action on the pitch is translated into a percentage of a goal opportunity based on tracking data that involves the locations of actions, the ball and 22 players (Austin, 2019). Here, if a goal chance is worth 0.50 expected goals, then this value is allocated to the number of players partaking in a passing sequence. This type of analysis gives decision-makers the chance to pinpoint the most valuable passers and ball carriers, therefore allowing for comprehensive analysis of technical ability that goes beyond passes, goals and assists.

(2) To Maximise Physical Performance

It can be argued that football is one of the most physically demanding team sports in the world. Generally, a player runs between 9 and 14 km and carry out approximately 1330 activities during a match (Sarmento et al., 2014). As football seasons tend to last around nine months and involve 30 to 50 games per season, players are required to be physically fit for their competition levels in addition to the technical demands of the game. Research on EPL shows that sprint distances and the number of passes went up by 35 and 40% respectively between 2006-07 and 2012-13 (Barnes et al., 2014). However, unlike the technical performance data, reliable physical performance data is not commonly available in football. Since a few top domestic competitions such as EPL and German Bundesliga have league-wide deals with tracking data partners, access to this kind of data can be provided to the clubs in these competitions. However, for all the other leagues all over the world, live scouting and video scouting remain to be the major ways to assess physical performance.

Recently, companies such as Sportlogiq and Skillcorner have developed computer vision-based solutions to this problem through broadcast tracking that generates data from match videos (Skillcorner, 2022; Sportlogiq, 2022). Comparing the physical output in the Big Five leagues, Skillcorner (2020) found out that Spanish La Liga and Italian Serie A record the highest total distance; however, EPL has the highest high-intensity distance, which is characterised by running above the speed of 5.5 metres per second. As is the case with these three leagues of comparable strength, the comparison of physical activity levels in different leagues can be useful in assessing the recruitment candidates coming from other leagues. According to Bradley and Ade (2018), playing positions of players and the game states are highly influential in determining their physical output. Therefore, contextualising physical performance could be crucial in getting the most out of physical performance data.

(3) To Effectively Use Age Groups

Age is one of the key factors to consider when building a playing squad and recruiting players. As is the case with products and organisations, athletes go through different life cycles. The physical and technical capabilities of players within a squad are expected to be affected by whether players are at the earlier, peak or later parts of their careers. The literature on market valuation of players suggests that age has a positive correlation with transfer fees and player wages until a certain point, and after that, this becomes negative for the rest of their careers, which results in a quadratic relationship (Carmichael et al., 1999; Dobson & Gerrard, 1999; Frick, 2007). Therefore, many club decision-makers need to focus relatively more on pre-peak and post-peak players to recruit the best playing talent they can afford. As explained in the fundamental objective of maximisation of transfer fee profits, pre-peak players help smaller clubs to take advantage of the transfer fee spending of bigger clubs. Naturally, this phenomenon does not stop the wealthiest clubs from investing in pre-peak players. Since there is intense competition to recruit the players with the potential to be the best, decision-makers in top clubs also face the necessity to recruit pre-peak players before any other club does. For instance, Manchester United FC ex-manager Sir Alex Ferguson signed potential stars of the day, Wayne Rooney and Cristiano Ronaldo, at the age of 18. Recruiting players in the early stages of their careers allow these clubs to avoid paying relatively higher transfer fees in the more advanced stages of their careers, whereas the risk involved in the decision could be much lower.

According to Anderson and Sally (2013), the peak age in the EPL is 26. Yet, since the playing positions have a substantial impact on the output of players, recent studies have focused on position-based samples. Having studied the four highest revenue-generating domestic leagues in Europe, Dendir (2016) found out that defenders peak at the age of 27, midfielders at 25-27 and forwards at 25. The relatively higher peak age of defenders can be attributed to the fact that accumulated experience in this position is more likely to compensate for the age-induced decrease in physical output (Dendir, 2016). In a study focusing on the players competing in the top European competition UEFA Champions League, Kalén et al. (2019) analysed the Transfermarkt values of players in four different age groups: 21-25, 26-30, 31-35 and 35+ years. According to this study, age has the highest positive influence on market valuation in 26-30 followed by 21-25, whereas 31-35 and 35+ groups are negatively affected by age (Kalén et al., 2019). Based on the effect of age on physical and technical performance as well as market valuation, club decision makers can specify ideal age intervals for different playing positions in accordance with their player expenditure projections. These specifications would help the recruitment alternatives to be generated based on well-informed value objectives.

(4) To Effectively Use Contract Types

A football player is free to negotiate contractual terms with other clubs in the last six months of their contracts and to leave his former employer when his contract expires (FIFA, 2020). This has become possible after the introduction of Bosman regulations as transfer fees were mandatory even after the expiration of employment contracts (Simmons, 1997). In 2020, 62.5% of new recruits in the world were out-of-contract ones that did not include a transfer fee, and only 11.6% required a fee (FIFA, 2021b). On the other hand, loan deals constituted 16.2% of all deals, whereas 9.7% returned to their parent clubs from loan periods (FIFA, 2021b). This composition demonstrates that many clubs refrain from paying transfer fees and rely on short-term loans and out-of-contract signings. Existing literature on transfer fee valuation shows that contract expiry left in a player’s contract has a positive effect on the transfer fee received (Carmichael et al., 1999; Feess et al., 2004; Frick, 2007) As a result, there can be a substantial difference between the fee exchanged between the same clubs and the same player in two scenarios with differing contract years left on the contract. Considering that the maximum duration allowed in a contract is five years (FIFA, 2020), for instance, three years give the selling club a much greater bargaining power, as opposed to the scenario where there is only a year left on a player's contract. Therefore, from a buying club perspective, a maximum contract expiry value can be specified as a means objective, in addition to pursuing the options of out-of-contract and loan players.

Since out-of-contract players are very much in demand, players with one year left on their contracts could lead to attractive deals for both parties. For instance, German international player Ilkay Gundogan was bought by Manchester City from Borussia Dortmund for €27 million in 2016 at the age of 25 (Ilkay Gundogan, 2021). As his contract was to run out in 2017, he could have left for free a year later. With this deal, Borussia Dortmund benefited from the spending power of Manchester City, whereas Manchester City could get a world-class player entering his peak years for a relatively smaller fee than their other signings. In comparison, Manchester City signed four defenders for a combined amount of €205.5 million in 2017 (Manchester City, 2021). Likewise, a club could have the opportunity to recruit a player for much less than his worth in a similar decision context.

Another important type of contract is done through recruitment on loan, where a player is temporarily signed from his parent club. As mentioned, signing players at an early age gives clubs the chance to pay smaller transfer fees and avoid the competition later in the careers of players. A potential setback to this is that there might be a development gap for these players between the ages of 18 and early 20s as they may not get first-team football during this period (Tremlett, 2017). In such cases, these deals can allow players to get the first-team experience in a club from a lower sporting level. For the loaning club, this tends to represent a short-term solution with no or very little transfer fee, thereby limiting the risk involved. Another commonly observed case is that a player might not manage to fulfil the expectations that his club has, and therefore he might be loaned out to another club. In both cases, there is the possibility of the parent club and the loaning club sharing the wage costs, so this might enable the loaning club to pay only a percentage of the total wages in his contract with the parent club. To benefit from the contributions of loan deals, some of the richest clubs in the world, such as Chelsea, have long been utilising a recruitment strategy that involves loaning out their pre-peak players to specific clubs that they have been cooperating with (Austin, 2017b).

(5) To Comply with Domestic Regulations

All football clubs are required to comply with the restrictions and quotas specified by their national football associations. Within the UEFA territory, some EU-member federations require work permit applications for non-EU citizens (UEFA, 2020a). On the other hand, some countries specify a maximum quota regarding the recruitment of non-EU players (UEFA, 2020a). For instance, Greece allows clubs to have a maximum number of eight non-EU players (The Football Association, 2020; UEFA, 2020a). Prior to the United Kingdom leaving the EU, a points-based system was introduced by the English football federation. In this system, players need to get 15 points from several criteria related to their club and national team appearances, as well as the sporting performance of the selling club (The Football Association, 2020; UEFA, 2020a). Another important rule that affects player recruitment is the UEFA Homegrown Player Rule, which states that four players should come from a club’s own youth academy and four others are required to have trained in the same country in a playing squad (Freeburn, 2009). Additionally, national associations are able to implement their own homegrown player regulations (Geey, 2020).

3.4. Building a Means-Ends Network of Objectives

In this section, the fundamental and means objectives are linked to show which objectives help others to be achieved (Keeney, 1996b). This network is built to reflect the relationships between fundamental and means objectives and the possible alignment of various objectives

(Keeney, 1996b). Among the contextual objectives, we see that constructing a football identity and recruiting players based on football identity fit are linked to all four fundamental objectives. These relationships underline how crucial it is to have a sporting concept to guide the recruitment efforts. Additionally, recruiting players based on financial fit is linked to the two fundamental objectives of minimising wage costs and maximising net transfer fee income. Therefore, this would have a substantial impact on the financial sustainability of clubs. Lastly, recruiting based on national teams is merely related to the maximisation of sporting results.

Within the player objectives, effectively using age groups and contract types of players are also related to all four fundamental objectives. As these two means objectives have consequences in terms of both financial and sporting aspects, they would be central to any decision-making context in recruitment. The other two means objectives, maximising technical and physical performance, are linked to three fundamental objectives as they can lead to desirable sporting results, deployment of a football identity and generation of net transfer fee income. Finally, complying with domestic regulations in terms of player quotas and restrictions allows clubs to compete in their relevant leagues so it can be conceptualised as an objective related to desirable sporting performance.

3.5. Managerial Implications and Conclusions

In this chapter, a decision-making framework on player recruitment was built via the application of VFT. While doing so, the available literature on football management and performance was utilised to bring together fundamental and means objectives with regards to the decision context in player recruitment. Finally, the objectives that enabled the realisation of other objectives were linked with a means-ends objective network. This network helps to combine sporting and financial performance aspects of player recruitment and to visualise the relationships between various objectives in recruitment decision-making.

This framework contributes to the research in sports management as it enhances the understanding of the multidimensional nature of the decision-making context regarding player recruitment in football. It also contributes to the area of operations research by demonstrating that VFT can be utilised to structure a holistic decision space in the domains of sports management and human resources management. The methodology behind this framework can be applied to other decision-making contexts in these two areas via identification of the relevant decision objectives and constraints.

From a practical perspective, this framework can be utilised by key decision-makers in football clubs as a conceptual basis for the effective and sustainable management of player costs. Since both the financial and sporting aspects of player recruitment decisions are covered, it can help to promote a wider understanding between multiple stakeholders who come from different areas of expertise. Additionally, the time dimension of the framework allows for setting recruitment objectives based on a longer timeframe than a year-long football season. In a competitive setting such as football, where clubs tend to suffer from short-term practices that target immediate sporting success, this framework provides a solid foundation for strategic planning and execution. Finally, key performance indicators related to the sporting performance of players and teams are presented based on the emerging literature in football analytics. These indicators can be influential in the creation of player alternatives and the comparison of these players.

4. Identifying Technical Functions of Footballers using Data Analytics

This chapter presents the player categorisation model that aims to derive the technical skill sets and functions of players.

4.1. Introduction

Managing the player cycle is a critical element for the competitiveness of football clubs. As players constitute the major resource that can lead to sustainable competitive advantage, recruitment of playing talent involves the most important strategic decision to be made in football clubs (Rossi et al., 2016). In the financial year 2018, player wages amounted to €10.3 billion and absorbed 49% of revenues generated, thus remaining to be the largest cost driver and risk factor of a football club (UEFA, 2020a). Furthermore, success and failure in player recruitment determine the sporting results that in turn translate to the football income of a club, which is composed of broadcasting rights, merchandising, sponsorship agreements and ticket sales (Szymanski, 2015). Based on their sporting performance in a given season, teams end up having fewer or more financial resources to build playing squads for the following seasons (Baroncelli & Lago, 2016). Another viable source of revenue for many clubs is income from player sales. Clubs that discover players early in their careers and sell their contractual rights to other clubs for higher transfer fees manage to benefit from this type of revenue. The share of transfer fee income in revenues generated by European clubs rose to a record level of 38% in 2017, up from 24% in 2014 (Chaudhuri, 2019). The combination of these points underlines how the overall success of a professional football club is highly dependent on the quality of its player investment activities and processes.

A common trend in top-level European football is that player recruitment efforts start with the identification of desired player characteristics needed for specific positions (Ancelotti et al., 2016; Lawrence, 2018; Soriano, 2011). When planning to strengthen a football team, Soriano (2011) explains that the weaknesses in the squad are identified, and profiles of potential recruits are clarified, such as a forward competent in aerial duels or an attacking right-sided defender. Consequently, chief scouts, sporting directors and chief executive officers collaborate with first-team staff on fulfilling the specified requirements (Anderson & Sally, 2013; Lawrence, 2018; Soriano, 2011). Carlo Ancelotti (Ancelotti et al., 2016), who is one of three managers to have won the UEFA Champions League three times, explains that his main responsibility in recruitment is to say, ‘we need this type of player for this position and that type of player for that position’. He also points out that a first-team manager should not be dictating player investment activities as the expected duration to work in that position is less than two years on average (Ancelotti et al., 2016). This recruitment approach allows the club decision-makers to have several alternatives and prevent them from entering bidding wars with other clubs (Soriano, 2011). Thus, data-driven categorization of player skill sets and functions is of vital interest to club decision-makers that are responsible for bringing together footballers with complementary skill sets and ensuring effective succession planning.

The evolution of top-level football has also elevated the importance of examining player functions. Tactical developments such as collective pressing and zonal positioning, introduced by pioneers such as Valery Lobanovskyi, Rinus Michels, Johan Cruyff and Arrigo Sacchi, have provoked a radical transformation in player functions and skill sets (Wilson, 2018). According to *The Technical Report 2018 FIFA World Cup Russia* (FIFA, 2018), the distance between the deepest and highest positioned players has been reduced to only 26 metres, with rapid ball movement becoming a necessity to create space against compact units. Hence, the requirement of high passing ability for every player to enable a fluid ball movement is now adopted by many teams in football (Balague, 2017; Cruyff, 2016; Honigstein, 2015). In accordance with the rising demands in constructing the play, the defensive workload for attacking midfielders and forwards has increased to cope with the fluidity of opponent attacks. This recent paradigm shift towards multi-functionality in European football is a further reason to analyse the skill sets of players and explore player functionalities.

The purpose of this chapter is to develop an analytical model that conceptualizes player functions based on their performance indicators. By analysing individual technical performance in the English Premier League, German Bundesliga I and Spanish La Liga during the season 2017/18, I propose a model for the creation of a taxonomy of football player functions. Those three domestic leagues constitute the unit of analysis because their clubs have been dominating the top European competition UEFA Champions League, with 17 of 20 finalists over the last ten years (UEFA, 2020b). I employ data analytical methods (including PCA and clustering) on technical performance indicators retrieved from the global data provider, Wyscout (Wyscout, 2022a).

With this study, I make four specific contributions to the literature. First, by building a taxonomy of footballers based on technical performance, I advance the understanding of player skill sets and functions, which can be further investigated in relation to talent management and market valuation. Second, I demonstrate how the framework can be used to support decision making regarding player recruitment. Third, I show that data mining and analytical methods can facilitate data-driven decision making to enhance organizational performance in the football industry. Even though some first team managers and head coaches demand total control of recruitment in football (Kelly, 2008), the majority of top-level clubs currently rely on a collaboration of experts overseen by the chief executive and the sporting director (Lawrence, 2018; Parnell, Groom, et al., 2018). A major responsibility of a sporting director is to employ a sustainable strategy for ‘buying low and selling high’ with a view to creating a competitive advantage (Lawrence, 2018). Moreover, clubs face the risk of having too many players with a similar function (Cruyff, 2016; Gullit, 2016) and the necessity to have suitable backups for their key players in case of injuries (Ancelotti et al., 2016). When building organizational decision-making processes in player identification, recruitment, development and selection, club directors and coaches can benefit from the derived skill sets and categories.

The fourth contribution of this study is to demonstrate that player positions are helpful in characterizing players, yet they may serve as a priori assumptions that cause cognitive biases in decision making, especially when coupled with heuristics based on physical attributes. Since team tactical changes over the course of matches and seasons lead to positional changes for players, Liverpool FC manager Jürgen Klopp warns about assigning players fixed positions: ‘Who even decides the positions? If you are a left back, maybe you started as a left winger. A central defender might have started as a No.6’ (Smith, 2017). Furthermore, Mills et al. (2018) examine how playing position varies according to skin tone and demonstrate that darker skin toned players are more likely to be used in external positions than central positions, which are associated with organizational skills and high technical capacity. The findings can be beneficial to tackling such decision-making biases caused by heuristics. The study is organized as follows. In the next section, I describe the data and methods utilised, followed by results and discussion. Finally, I conclude with the framework's practical implications, limitations and directions for future research.

4.2. Data and Methods

4.2.1. Data Sources and Preparation

The performance indicators of players are retrieved from the database of Wyscout (Wyscout, 2022a), a well-known player scouting platform that collects and provides match event data and videos to professional football clubs (Kidd, 2019). This platform is used by more than

1,000 soccer clubs in 100 countries, including Manchester United and Barcelona, in addition to international bodies such as UEFA and FIFA (Kidd, 2019). The operational definitions can be found in their events manual (Wyscout, 2022b). In addition to football clubs and federations, football players also use the database to learn about the strengths and weaknesses of the opposing teams (Thomas, 2018).

The data set used in this study consists of all 1,066 matches played in the 2017/18 season in the top football tiers of England, Germany and Spain. These leagues constitute the highest revenue-generating domestic football competitions, so they are able to attract the best players (Szymanski, 2015). As pointed out by Felker (2012), knowledge workers are motivated by higher levels of intellectual activities and professional opportunities. This argument applies to the vast majority of people working in the global football industry as well. Even though the player functions are expected to differ across leagues, this sample can be viewed as the most suitable for setting an international benchmark for players and decision-makers alike.

Access to the database of the company for the selected leagues was achieved thanks to the company’s API service. Here, a Python script was utilised to download player data via competition, season, team and player IDs. This script helped to retrieve the individual statistics of players during the analysed 2017/18 season. In the database, performance indicators of players within a domestic season are provided in total numbers throughout the season as well as in output per 90 minutes. I consider the latter in this study because the total numbers can be misleading in overstating the output of players who are fielded most frequently. To present a realistic picture of the contributions of players in the sample, a minimum threshold of minutes is set. Consequently, players fielded for less than 900 minutes per season are excluded from the data set, as proposed by Kharrat et al. (2020).

Table 5 Mean values and maximum absolute standard deviation from the mean (%) of key match events per 90 minutes.

Match events	England	Germany	Spain	Mean	Max. absolute std. dev. (%)
Crosses	15.1	14.1	15.8	15.0	6%
Dribbles	27.2	28.0	29.3	28.2	4%
Goals	1.2	1.3	1.3	1.3	5%
Head shots	1.6	1.8	1.7	1.7	6%
Passes	404.2	406.7	400.2	403.7	1%
Passes to final third	58.0	56.5	54.7	56.4	3%
Shots	10.9	11.4	10.7	11.0	4%

In the data exploration and preparation phase, firstly, the variation of key events between the three leagues in England, Germany and Spain was analysed. The purpose of this was to ascertain if there are large differences across leagues in player output. Table 5 shows that the variation in the key events such as goals, shots, passes and dribbles is within 6% of the mean values in the selected leagues during the season 2017/18. As this can be viewed as an acceptable level of variation, all players from the three leagues are analysed in one sample.

Secondly, correlation analysis is applied to the data set with a view to discarding redundant metrics. As match events such as passes and duels are represented by two variables each in the data set as ‘total’ and ‘successful’, this can lead to a high level of multicollinearity between variables. Here, nineteen event pairs such as ‘key passes’ and ‘successful key passes’ have a positive correlation higher than 0.8. Therefore, the variables representing total attempts among these pairs are removed from the sample to shed light on events that occur less frequently and possess more explanatory power.

Table 6 Major types of tasks in football and event variables used in this chapter to characterise these tasks.

Tasks	References	Related variables
Ball recovery	3, 5, 7	Ball recovery, dangerous opponent half recovery, opponent half recovery
Cross	3, 4, 5, 6, 7	Successful cross
Dribble	1, 5, 6, 7	Long dribble, successful dribble
Duel	1, 2, 5, 6	Aerial duel won, defensive duel won, loose ball duel won, offensive duel won
Interception	5, 7	Interception
Pass	1, 3, 4, 5, 6, 7	Assist, expected assist, successful back pass, successful forward pass, successful key pass, successful linkup play, successful long pass, successful pass, successful pass to final third, successful penetrative pass, successful through pass, successful vertical pass
Shot	1, 3, 4, 5, 6, 7	Expected goal, goal, head shot on target, shot on target

References: (1) Barnes et al. (2014); (2) Dellal et al. (2011); (3) (Fernandez-Navarro et al., 2016); (4) (Lago-Peñas et al., 2010); (5) (Liu et al., 2016); (6) (Rampinini et al., 2009); (7) (Yi et al., 2018).

Based on previous research on technical performance and playing styles in football (Barnes et al., 2014; Dellal et al., 2011; Fernandez-Navarro et al., 2016; Lago-Peñas et al., 2010; Liu et al., 2016; Rampinini et al., 2009; Yi et al., 2018), seven major tasks which can be considered as most relevant to technical player functions are shown in Table 6. The table also provides the event variables used in this study to characterize the various tasks.

Given the likelihood of players to perform more attacking and less defensive events when their teams keep the ball for longer periods, values of nine events are adjusted to account for variability in team possession. Thus, events that correlate more than 0.80 with average team possession percentages, such as successful passes and shots on target, are adjusted on the basis of a 50% possession rate. For instance, the number of successful forward passes made by Fernandinho Luiz Roza of Manchester City FC, which has an average possession percentage of 64.2%, is reduced to 16.29 from 20.92.

Lastly, to ensure that the event names comply with existing literature in sports sciences, I check the performance indicator descriptions and compare them to the references in Table 6. Consequently, I change the events accelerations and smart passes to long dribbles and penetrative passes, respectively. Wyscout (Wyscout, 2022b) defines acceleration as a dribble with the ball longer than 10 metres at an increasing pace, yet this could be confused with the acceleration of any kind with or without the ball. Additionally, a smart pass is described as a creative pass that outplays at least 2 or 3 players leading to a preferable attacking position (Wyscout, 2022b). This type of pass was conceptualized as a penetrative pass by Tenga et al. (2010) and was found to be more effective than non-penetrative passes.

When interpreting the findings, I utilize individual wages from the 2018 version of Football Manager (FM) (Football Manager 2018, 2017) in order to compare market valuations of players and identify top earners within clusters and club squads. The FM database is brought together by around 1,300 researchers and 100 head researchers who gather intelligence on player attributes and contracts (Bleaney, 2014; Skandalis et al., 2015). As individual player wages are rarely made public by professional football clubs, I use FM 2018 wages as a proxy for player remuneration.

4.2.2. *Data Modelling*

Dimensionality Reduction

In this phase, I aim to explore the patterns in the player performance data with a view to discovering skill sets that can be later deployed in the identification of player functions. In the decision-making process related to player recruitment and development, decision-makers need to have an optimal set of key performance indicators (Lawrence, 2018; O'Donoghue, 2008). Moreover, the presence of too many variables would negatively affect the cluster analysis as a high number of dimensions would lead to distances being relatively uniform (Beyer et al., 1999). For this reason, I deploy PCA, a widely used method that creates an alternative smaller set of variables while aiming to preserve the maximum variance (Jolliffe, 2002). The latent variables called principal components (PC) are linear combinations of the original variables that correlate highly with one another (Jolliffe, 2002). This technique has been used in identifying the playing styles of teams in football (Fernando et al., 2015), as well as producing key performance indicators in tennis (Gløersen et al., 2018), basketball (Bruce, 2016) and rugby (Parmar et al., 2018). Similarly, it can enable the creation of a small number of variables that preserve the most information and summarize the relationships within the data set.

Among the components created, the first PC explains the most variation in the original variables, and each successive component contributes less to the total variance explained. As there are as many components created as the number of variables with decreasing importance, one of the challenges in utilizing PCA is to pick a small number of components to be analysed. As a minimum level of 70\% is viewed as reasonable (Jolliffe, 2002), I pick the number of principal components according to this requirement. The interpretation of a specific component is carried out through analysing the correlations of the component with each variable, which are called component loadings of variables (Jolliffe, 2002). As each PC is a different linear combination of original variables, the loadings from -1 to +1 demonstrate how important the variables are for each component (Han et al., 2011). These linear combinations provide the PC scores for each data object, namely the players in this study. In the clustering phase, the individual principal component scores of players are utilized to form and analyse clusters. As proposed by Osborne et al. (2008), I merely consider the variables with absolute component loading values over 0.3 when analysing the components. This cut-off value is set to establish the variables that are influential in the formation of a component. To evaluate the suitability of the sample for data reduction via PCA, I carry out Kaiser-Meyer-Olkin (KMO) and Bartlett's tests. A KMO value of over 0.5 is required to confirm the adequacy of the sample whereas Barlett's test verifies the significance of results ($p < .001$) (Tabachnick et al., 2007).

Cluster Analysis

Cluster analysis is a statistical technique that combines data objects or observations that share similar characteristics in previously unknown groups (Han et al., 2011). Unlike data classification, which groups observations into pre-determined categories, cluster analysis enables the user to discover new concepts in a domain (Everitt et al., 2011). Accordingly, clustering is known as an unsupervised learning method, whereas classification utilizes a supervised learning procedure. A supervised approach in football would require the input of a sample of top-level club decision-makers that could be fully representative of the selected leagues. Due to the highly competitive nature of football, gathering such data might pose a big challenge in the area of sports management research. Moreover, this approach would not allow for the generation of new concepts regarding player functions. Thus, I pick cluster analysis over classification and utilize existing football concepts in the interpretation of the findings.

The choice of clustering type in this study, AHC, creates a tree-like structure where each data object starts as a single cluster that successively merges with other objects in each iteration (Everitt et al., 2011). The clustering procedure in AHC is summarized and presented by a dendrogram, which portrays the similarities of merged clusters as well as the hierarchical formation of clusters (Everitt et al., 2011). As AHC helps to visualize a hierarchy of underlying groups and sub-groups among a given set of data objects (Han et al., 2011), it can be employed to establish a taxonomy of technical player functions, which is one of the contributions of this study.

4.2.3. Empirical Analysis

As mentioned previously, PCA has been deployed in sports-related research to reduce the dimensionality of performance data. A potential problem here is that newly created components can be difficult to interpret (Kane & Shiri, 2016). In the context of performance analysis, the conceptual meaning of PCs might be hard to describe when variables corresponding to different areas of specialization correlate together. Moreover, since 12 of 19 on-possession performance variables are related to passing, a high amount of variance is expected to be explained by the component that brings together passing-related variables. This imbalance can cause the other components related to shooting, dribbling and crossing events to seem unsubstantial whilst examining the range and importance of skill sets available. I therefore, attempt to categorize variables into major task-based groups before utilising PCA. Figure 4.1 demonstrates the variable flowchart that provides the variable types utilised in different phases of the study.

Fernandez-Navarro et al. (2016) and (Castellano et al., 2012) use defensive and attacking play as two main groups of tasks. (Lago-Peñas et al., 2010) further categorise variables related to attacking play as offence and goal-scoring, with actions such as passes, crosses and assists represented in the offence group.

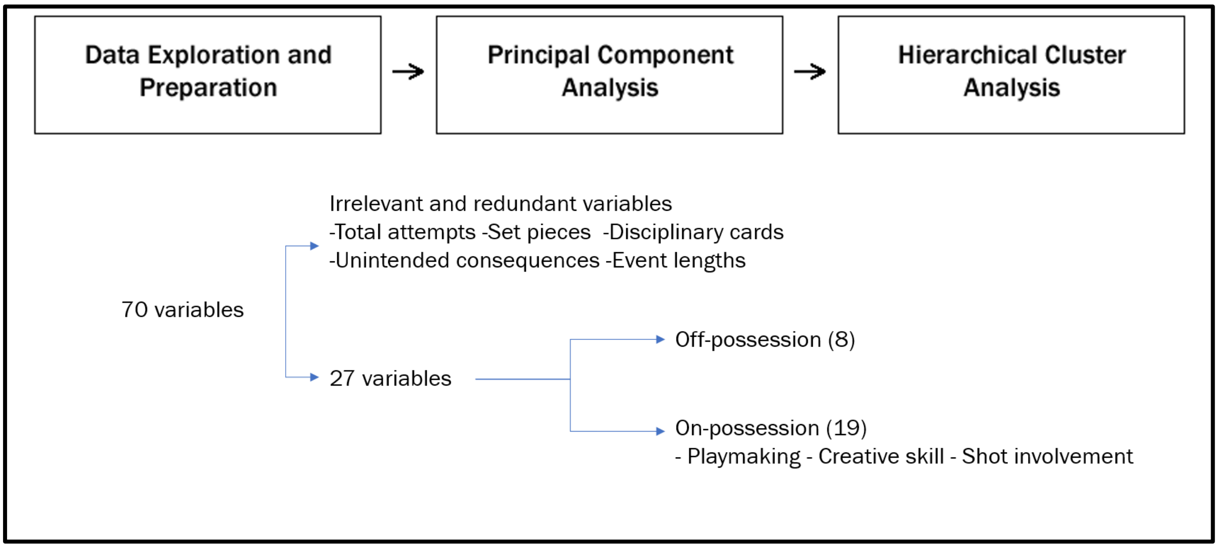

Figure 4.1 The phases of the analysis and the variable flowchart.

According to the FIFA Coach of the Century, Rinus Michels (Michels, 2001), major tasks in modern football are defence, build-up and offence. Yet, events such as loose ball duels and aerial duels are hard to pinpoint as defensive or attacking, as these events occur when no player is in control of the ball. Therefore, I first specify the variables as off-possession and on-possession. According to Jones et al. (2004), ball possession is defined as ‘sufficient control over the ball to enable a deliberate influence on its direction’. Table 7 gives an overview of 27 variables that are grouped into eight off-possession and 19 on-possession variables in line with this principle. The off-possession group includes three ball recovery events, three duel events, interceptions and head shots on target, whereas the remaining variables are categorized as on-possession. Among the on-possession variables, taking shots and making assists for other players can both be categorized as offence (Fernandez-Navarro et al., 2016). However, as proposed by (Lago-Peñas et al., 2010), these two events could represent different skills in their own right.

Likewise, tasks such as through passes, passes to final third, and dribbles are not easy to distinguish as build-up or offence variables. Given that AHC is effective in measuring the correlations between a large number of variables and grouping variables that are similar to each other (Linoff & Berry, 2011), I then utilise AHC in bringing together similar on-possession performance variables.

Table 7 Off-possession and on-possession variables used in the principal component analysis phase

Variables	Mean	Std. Dev.	Min.	Max.
Off-possession variables per 90 minutes				
Ball recoveries	4.74	2.06	0.63	13.50
Dangerous opponent half recoveries	0.26	0.15	0.00	0.95
Defensive duels won	1.32	0.62	0.05	3.14
Field aerial duels won	1.79	1.40	0.00	13.10
Head shots on target	0.06	0.09	0.00	0.59
Interceptions	4.10	1.72	0.74	8.51
Loose ball duels won	1.94	0.62	0.27	5.05
Opponent half recoveries	1.48	0.63	0.21	4.97
On-possession variables per 90 minutes				
Assists	0.08	0.09	0.00	0.58
Expected assists	0.09	0.08	0.00	0.48
Expected goals	0.13	0.15	0.00	1.10
Goals	0.13	0.17	0.00	1.16
Long dribbles	0.62	0.57	0.00	3.11
Offensive duels won	3.14	2.25	0.00	13.76
Shots on target	0.41	0.39	0.00	2.60
Successful back passes	5.51	2.25	0.93	12.66

Successful crosses	0.48	0.47	0.00	2.83
Successful dribbles	1.82	1.58	0.00	10.25
Successful forward passes	9.21	4.56	0.82	24.68
Successful key passes	0.22	0.20	0.00	1.25
Successful linkup plays	1.03	1.70	0.00	9.80
Successful long passes	1.83	1.26	0.00	6.32
Successful passes	31.08	11.46	7.49	67.08
Successful passes to final third	3.65	2.00	0.29	12.73
Successful penetrative passes	0.29	0.28	0.00	2.03
Successful through passes	0.21	0.19	0.00	1.29
Successful vertical passes	11.61	6.19	1.79	32.65

In all types of cluster analysis, the calculation of similarities between players is decided by the similarity measure and linkage method selected (Everitt et al., 2011). To explore how on-possession variables move in relation to one another, I choose the Pearson correlation coefficient as the similarity measure in variable clustering. For clustering based on this measure, three main linkage options are available: single, complete and average (Everitt et al., 2011). When calculating the similarity between two clusters, single and complete linkage considers the shortest and longest distances of cluster members, respectively (Han et al., 2011). The average linkage method defines the average distances between data objects in different clusters and therefore is relatively more robust than single and complete linkage options (Everitt et al., 2011). Thus, it is selected as the similarity measure for this phase. To visualize the relationships between variables and differentiate between variable groups, I created a heat plot based on Pearson correlations alongside the clustering dendrogram.

In the AHC phase for player clustering, the linkage method selected is Ward's method. Ward's method is a popular clustering method that seeks to combine clusters that minimize the within-cluster variance (Han et al., 2011). Thus, it can be useful in supporting the utilised constructive clustering concept as it tends to penalize large distances within clusters heavily and to create relatively compact clusters (Everitt et al., 2011). According to Hennig (2015), a constructive concept is used for pragmatic reasons such as structuring of entities and complexity reduction, as opposed to a realist one, where a real existing structure is pursued. Ward's method utilises Euclidean distance as a dissimilarity measure, and this is suitable for the nature of the data set composed of merely continuous variables. As pointed out by (Sarstedt & Mooi, 2014), Euclidean distance is the most commonly used dissimilarity measure for continuous data. In line with the clustering aims, utilizing Euclidean distances can help to cluster players based on the relative magnitude of each feature across players. In Ward's method, the Euclidean distance of a new cluster $I \cup J$, formed by merging two clusters I and J to give a new cluster K is calculated as

$$d(I \cup J, K) = \frac{(n_I + n_K)\mathrm{d}(I, K) + \left(n_J + n_K\right)\mathrm{d}(J, K) - n_K \mathrm{d}(I, J)}{n_I + n_J + n_K}$$

To assess cluster validity, the average Silhouette coefficient of formed clusters is calculated. The silhouette coefficient is an index that evaluates the compactness and separation of clusters at the same time (Rousseeuw & Kaufman, 1990). It varies from -1 to +1, where a higher value refers to a more desirable clustering quality (Han et al., 2011). Negative values in the Silhouette coefficient indicate a clustering that has either too many or too few clusters created. Specifically, a measure between 0.2 and 0.5 is considered a reasonable solution (Sarstedt & Mooi, 2014). The Silhouette index has proven to be a robust and reliable validity metric to measure clustering performance for unsupervised data clustering (Bolshakova &

Azuaje, 2003; Handl et al., 2005; Vendramin et al., 2010). The Silhouette index can be calculated as

$$S(x) = \frac{b(x) - a(x)}{\max[b(x), a(x)]}$$

where *x* is a member of a cluster, *a(x)* is the average Euclidean distance between *x* and other members in the cluster, whereas *b(x)* is the minimum of the average distances between *x* and members of the neighbouring cluster. As a result, the Silhouette coefficient of a cluster is the average *S(x)* values of its elements.

An advantage of Ward's AHC is that the number of clusters is not required to be decided by the user beforehand (Han et al., 2011). However, a cut-off point has to be defined based on the number of hierarchical levels desired. To be able to do this, I visualise Euclidean distances of two clusters merged within the last 20 cluster merges in an agglomeration schedule. Two clusters with a relatively high Euclidean distance in a merge indicate that dissimilar clusters are being combined together and that the subsequent agglomeration stages after a certain threshold shall not be taken into consideration. Additionally, a local maximum in the Silhouette coefficient would be desirable as it would provide the ideal balance of separation and compactness of formed clusters.

After the clustering solution is decided, the main clusters are conceptualized by differential labelling based on their average principal component scores. As stated by Manning et al. (2008), differential cluster labelling assigns categories to clusters based on a comparison with the characteristics of other generated clusters. Accordingly, the PC scores of all clusters are compared with a view to specifying function-based conceptualizations. Later, players with the highest FM 2018 wages within clusters are presented along with the median wages of these groups. The introduction of the highest earners pinpoints the established stars, whereas a comparison of the median wages across clusters helps to demonstrate how different functions were valued by club decision-makers in the season 2017/18.

4.3. Results and Discussion

4.3.1. Identification of Skill Sets

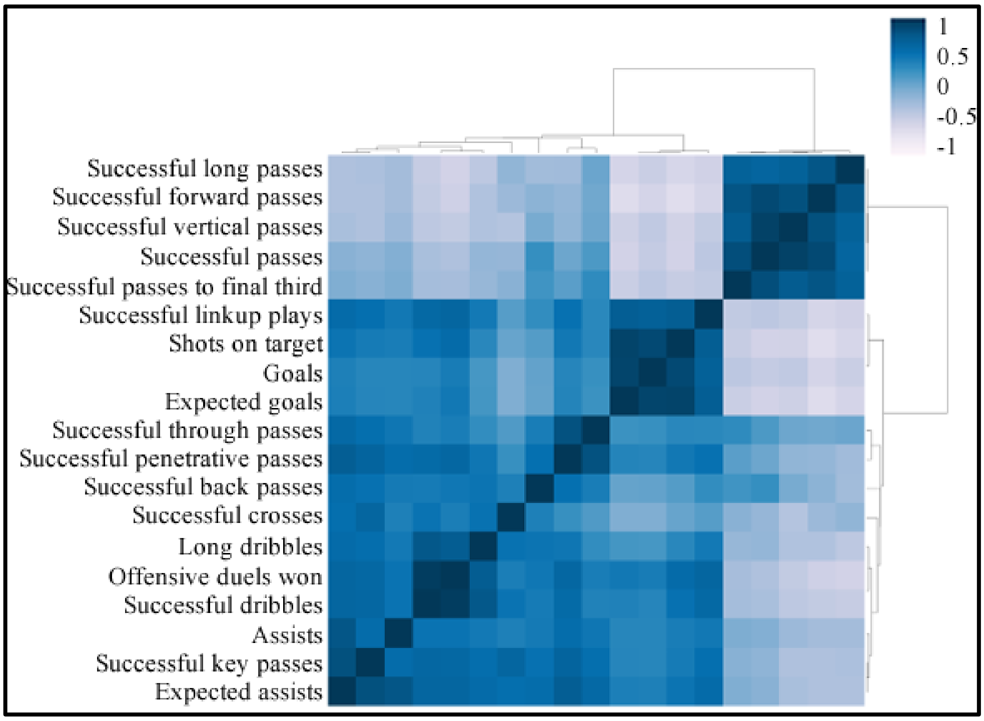

Figure 4.2 Correlation heat map for on-possession variables. The shading indicates the correlation between two variables: a darker colour indicates a stronger positive correlation.

In Figure 4.2, I observe five variables related to the construction of the play with high correlations between them on the top of the heat map. The remaining variables seem to have two separate sub-groups as the colouring indicates that the variable group, including goals and shots on target, represent a distinct skill set to the subset of 10 variables related to creative skills such as assists, successful dribbles and successful penetrative passes. Therefore, I choose to have three PCA phases for on-possession variables based on these three distinct groups: playmaking, shot involvement and creative skill. I apply z-score standardisation to the variables before applying PCA to data subsets to avoid overstating the importance of variables with relatively larger scales, such as successful passes (Everitt et al., 2011). Afterwards, I evaluate the suitability of the variable groups for this method. The values of all KMO tests are over 0.50, with off-possession, playmaking, creative skill and shot involvement samples corresponding to values of 0.64, 0.52, 0.78 and 0.85, respectively. Thus, the samples used can be considered adequate for PCA. Bartlett's tests of sphericity are significant ($p < .001$), indicating that PCA is an appropriate method. As proposed by (Osborne et al., 2008), I utilize the variables with component loading values over 0.30 when assessing which variables are influential in the formation of derived components.

In the off-possession PCA stage, the first three principal components generated explain 76.8% of the variance (see Appendix A for the PC loading tables available). Here, PC1 is positively correlated with variables related to ball-winning, such as ball recoveries, interceptions and defensive duels won. Off-possession PC2 brings together variables of loose ball duels won, aerial duels won and head shots on target, which are expected to be carried out by players with strong athletic abilities. Lastly, off-possession PC3 has strong positive loadings on opponent half recoveries and dangerous opponent half recoveries. Evidently, players who score highly on this component participate in collective pressing and manage to recover the ball in the opponent's half.

PCA performed on five playmaking variables yields one component explaining 81.6% of the total variance. Playmaking PC1 possesses similar positive loadings on all five variables, such as successful forward passes, successful long passes and successful passes to final third. Similarly, the shot involvement PCA generates only one component containing 84.8% of the variance. In the creative skill PCA, the first component accounts for 64.9% of the total variance and has positive loadings on five variables such as expected assists, assists and offensive duels won. Additionally, creative skill PC2 explains 12.4% of the variance and is positively correlated with successful penetrative and through passes. On the other hand, it has strong negative loadings on long dribbles and successful crosses. This contrast makes intuitive sense as it portrays the antagonism between creative passers and dribblers of the ball.

4.3.2. Clustering Solution

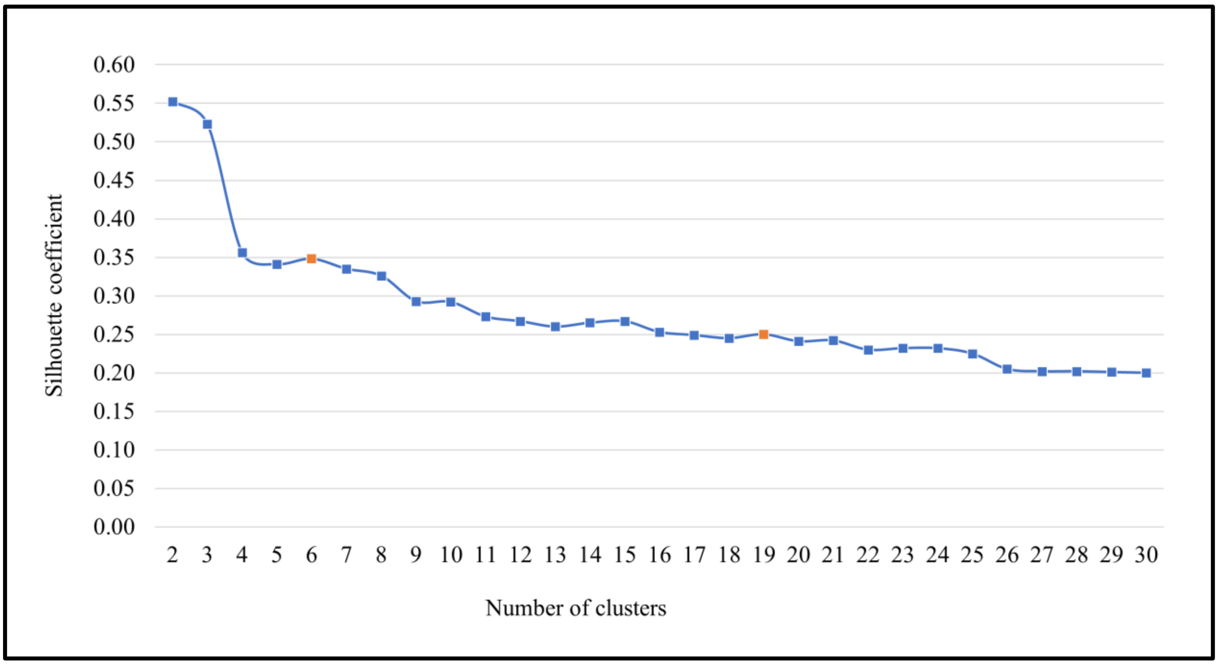

Figure 4.3 Agglomeration schedule of clustering stages. In each stage, the Silhouette coefficient of the obtained clustering solutions is given. The selected solutions are indicated in orange.

Figure 4.3 demonstrates the agglomeration schedule that is utilized to select the cut-off point in the clustering procedure. Here, the aim is to pick two stages of the schedule where a local maximum of the Silhouette coefficient is obtained. This would allow having a smaller number of main clusters and a larger number of sub-clusters that form them. The local maxima in terms of Silhouette coefficients would help to identify the best solution in terms of compactness and separation. Consequently, the first cut-off is fixed at the six-cluster solution. The Silhouette coefficient for this solution is 0.35, which indicates that a moderate clustering

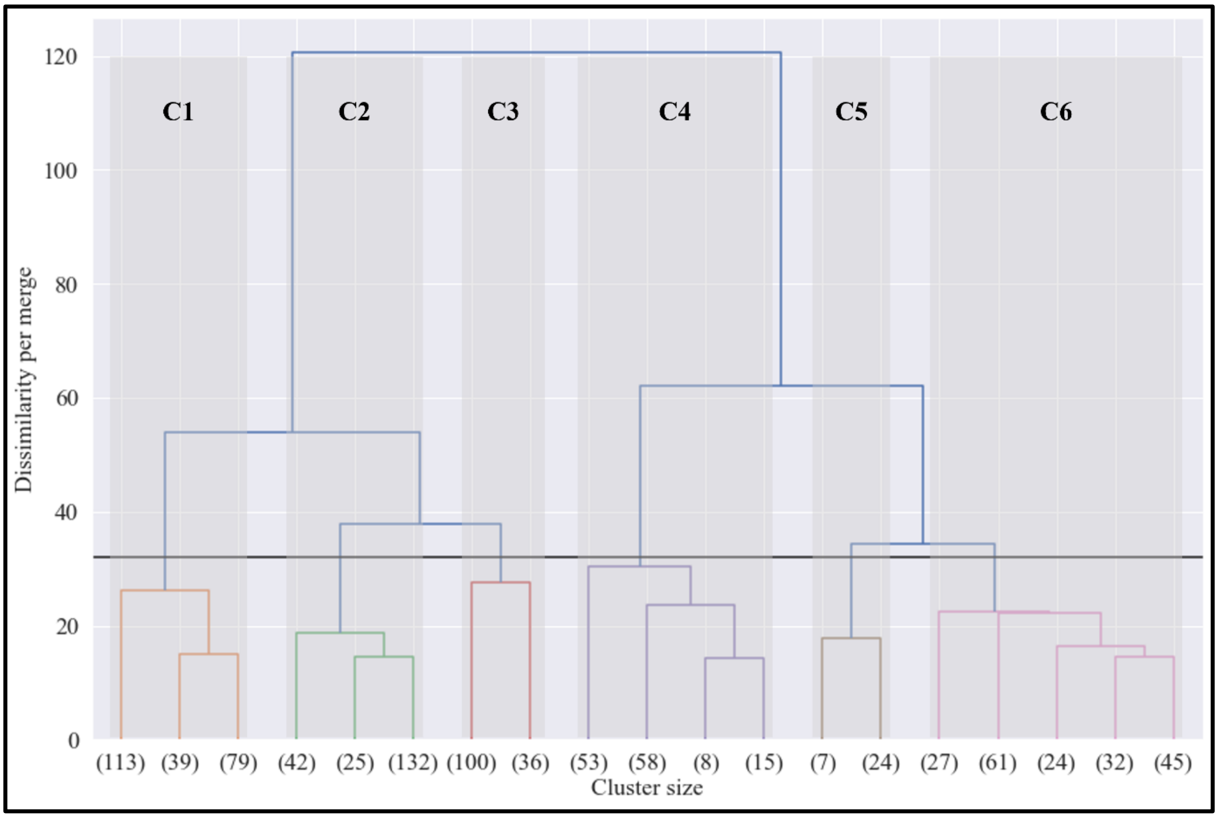

Figure 4.4 Clustering dendrogram: dissimilarity of the clusters merged (in Euclidean distance) and the sizes of the cluster sub-groups. The dissimilarities are indicated by the length of horizontal lines of the dendrogram.

quality is achieved. For the second cut-off point, the 19-cluster solution is chosen at a point where the silhouette coefficient is still moderate with a value of 0.25. From the agglomeration schedule, we see that the Silhouette coefficient keeps getting smaller after this point. Based on these two solutions, we can first analyse the meaning of six main clusters and profile the sub-clusters under these groups. Figure 4.4 shows the clustering dendrogram obtained on the basis of seven PC scores and the mentioned cut-off points. In the next section, I analyse the conceptual meaning of six clusters and label them based on the average PC scores of cluster members.

4.3.3. Cluster Labelling

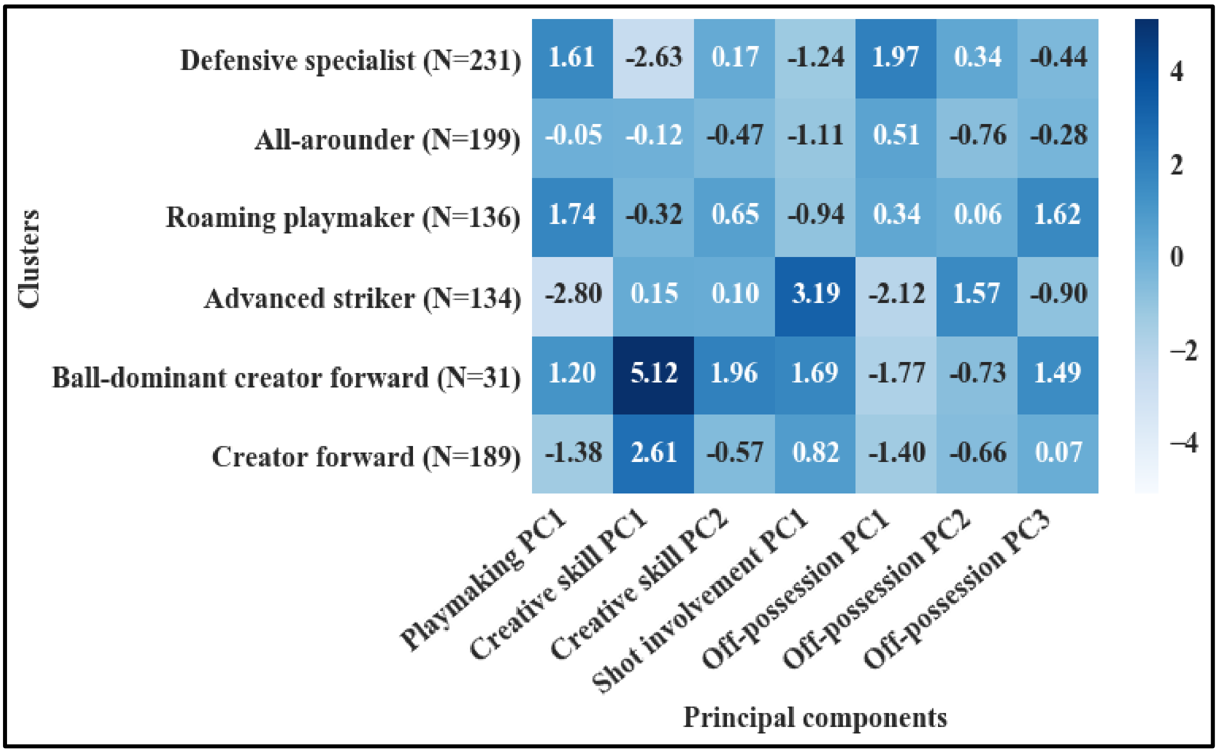

Figure 4.5 Average principal component scores of six clusters. The coloring indicates the magnitude of average principal component scores for each cluster.

Figure 4.5 provides a summary of the six-cluster solution on the basis of average PC scores and cluster sizes. Here, Cluster 1 identifies players with the highest average off-possession PC1 score across all subgroups, which correlates positively with variables related to ball-winning. Moreover, they score very low in creative skill PC1 and shot involvement PC1. Based on the defensive specialist term proposed by Green et al. (2016) in basketball, I label this cluster defensive specialist. Cluster 2 combines players that do not score highly in one specific skill set, so I label this group all-arounder.

Cluster 3 players have the highest playmaking PC1 and the second-highest creative skill PC2 score among clusters. Moreover, they have the highest off-possession PC3 score, which is positively correlated with opponent half recoveries. On the basis of playmaking, creative passing and ball-winning skills portrayed by its members, I call this cluster a roaming playmaker.

Cluster 4 is characterized by players that score the highest on shot involvement PC1, which is correlated highly with variables such as shots on target, expected goals and goals. They also have the highest score in off-possession PC2 that has high loadings on head shots on target, field aerial duels won, and loose ball duels won. These players do not contribute much to other major tasks such as playmaking or ball-winning. As a result, this cluster is called advanced striker.

Cluster 5 has by far the highest creative skill PC1 score and the second-highest shot involvement PC1 score. Additionally, they have the highest creative skill PC2 score across all categories, so they are very proficient in playing penetrative and through passes that outplay opposing defenders. These players seem to bring together the qualities of shot creators and ball-dominant guards proposed by Green et al. (2016). Thus, ball-dominant creator forward is a suitable label for this cluster as these players dominate the offensive performance of their teams through chance creation, playmaking, creative passing and shot involvement. Cluster 6 members have the second-highest score in creative skill PC1 and the third-highest score in shot involvement, which shows their competence in creating goal chances for others, winning offensive duels and performing shots. Yet, they have very low scores in playmaking PC1 and off-possession PC1, as is the case with advanced strikers. Consequently, this cluster is called creator forward.

Figure 4.6 and Figure 4.7 show the comparison of the representative and the most expensive players of the six main clusters. Here, the most expensive players are the highest wage earners, as per the FM database (Football Manager 2018, 2017). The cluster representatives are the players that have the shortest Euclidean distance to the cluster averages in PC scores.

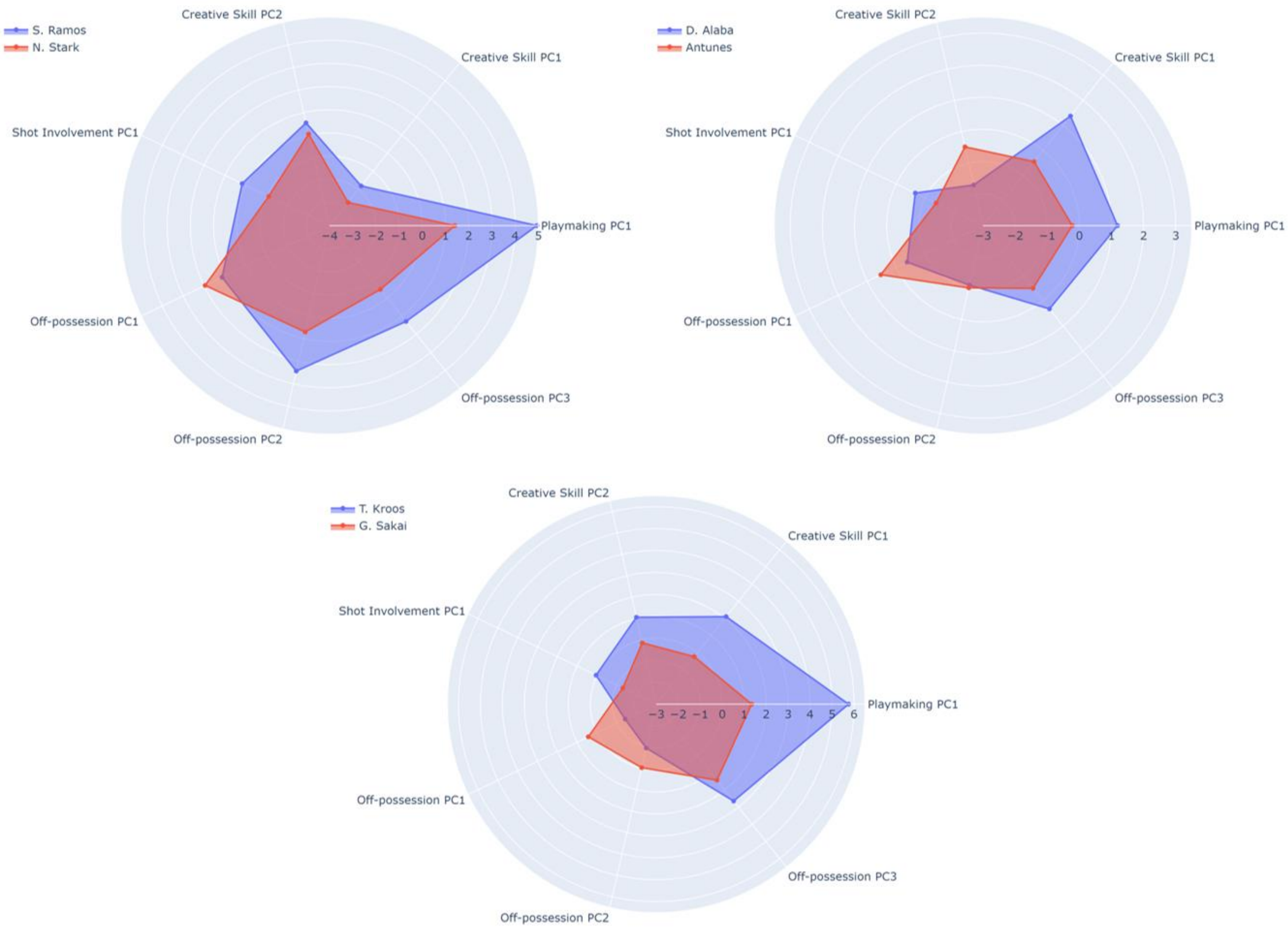

Figure 4.6 Comparison of the representative (in red) and the most expensive players (in blue) of the first three clusters.

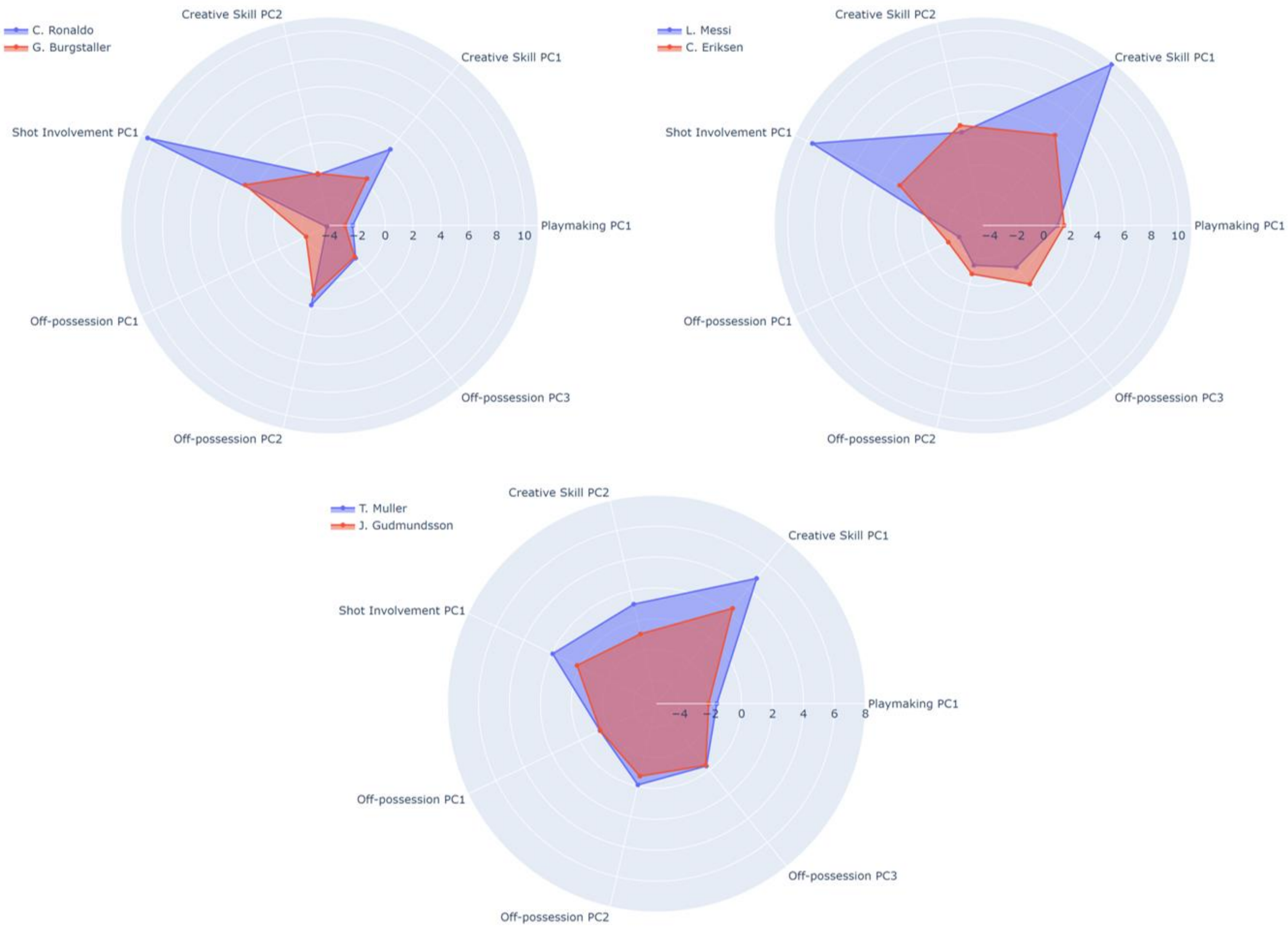

Figure 4.7 Comparison of the representative (in red) and the most expensive players (in blue) of the next three clusters.

4.3.4. Analysis of Sub-Clusters

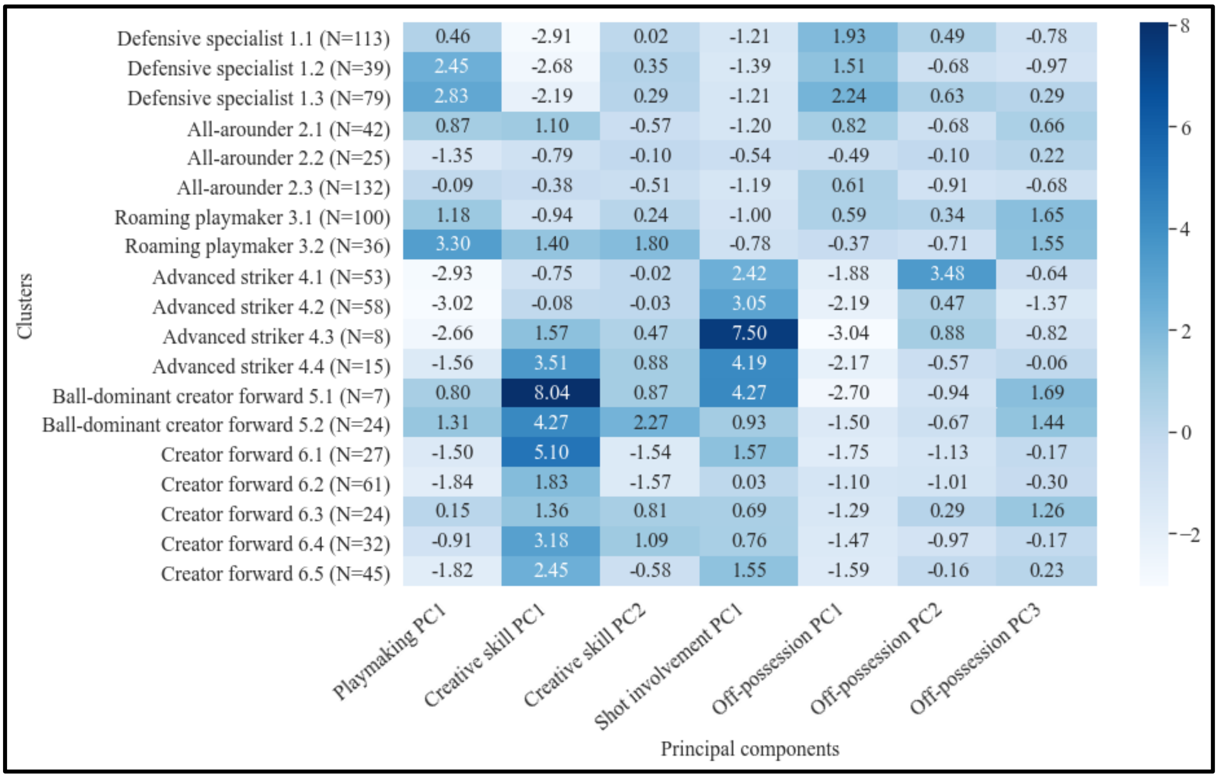

Figure 4.8 Average principal component scores of nineteen sub-clusters. The coloring indicates the magnitude of average principal component scores for each cluster.

This section provides an analysis of the 19 sub-clusters that help to define player typologies (see Appendix B for the radar comparisons of the representative and the most expensive players in each sub-cluster). Figure 4.8 presents an overview of the average PC scores for each sub-cluster, which forms the basis for the sub-cluster analysis.

Defensive Specialists

Cluster 1.1 players have a relatively much lower average score of playmaking PC1 than the other two sub-clusters. In terms of off-possession scores, they are quite close to the average scores of the whole cluster. Therefore, it can be argued that this sub-cluster brings together many players that are suitable for clubs that do not demand a very high level of output in terms of building play. Eric Bailly of Manchester United and Ashley Williams of Everton are two high-profile examples in this sub-cluster.

Cluster 1.2 players have a very high playmaking PC1 score, and this cluster is composed of players who are comfortable in possession that help considerably with ball progression. Yet, they score much lower than the other two clusters in off-possession PC2, which is highly correlated with aerial ability. Matija Nastasic of Schalke 04 and Jordi Amat of Real Betis are two well-known examples of this sub-cluster.

Cluster 1.3 players have high scores in both playmaking PC1 and off-possession PC2. Therefore this sub-cluster includes players that contribute to their teams in playmaking and aerial duels. They also have the highest average score among the defensive specialist cluster in off-possession PC3, which shows competence in opponent half recoveries. Hence, they would be suitable for a playing style that requires pressing the opponent. Virgil van Dijk of Liverpool and Matt Hummels of Borussia Dortmund are two well-known players in this sub-cluster.

All-arounders

Cluster 2.1 players have the highest playmaking PC1 and creative skill PC1 scores in this cluster. These scores highly correlate with playmaking and creating chances for others. Hence, it can be argued that this sub-cluster is composed of technically superior players compared to other sub-clusters. David Alaba of Bayern Munich and Kyle Walker of Manchester City are two high-profile examples of this sub-cluster.

In contrast with Cluster 2.1, Cluster 2.2 players have the lowest playmaking PC1 and creative skill PC1 score. This shows that these players did not contribute to ball progression and chance creation as much as the other two sub-clusters. They also have the lowest off-possession PC1 score within the cluster, so they did also have a lower level of output in ball-winning. On the other hand, they have a much higher average score in off-possession PC2, which correlate highly with aerial ability. Cheikhou Kouyate of West Ham United and Maximilian Eggestein of Werder Bremen are two well-known players in this sub-cluster.

Cluster 2.3 players have average skill set scores that are very similar to the cluster average. Marcos Alonso of Chelsea and Juanfran of Atletico Madrid are two high-profile examples of this sub-cluster.

Roaming Playmakers

Cluster 3.1 players score more highly in off-possession PC1 and PC2, therefore they contribute more to ball winning and aerial activity. Yet, Cluster 3.2 players have higher average scores in key on-possession PC scores such as playmaking PC1, creative skill PC1 and creative skill PC2. This shows that they are more competent in build-up play, chance creation and playing through balls in the final third. Hence, it would be fair to say that Cluster 3.1 represent roaming playmakers that contribute more out of possession, whereas Cluster 3.2 is composed of players that excel more in possession of the ball. Saul of Atletico Madrid and Javi Martinez of Bayern Munich are two examples of Cluster 3.1. On the other hand, Andres Iniesta of Barcelona and Luka Modric of Real Madrid are two high-profile examples of Cluster 3.2.

Advanced Strikers

Cluster 4.1 players have a much higher off-possession PC2 score than the other two sub-clusters within this group and therefore include advanced strikers that are competent in winning aerial duels. On the other hand, they have a much lower creative skill PC1 score, which is highly correlated with creating chances for others. Olivier Giroud of Chelsea and Christian Benteke of Crystal Palace are two well-known examples in this sub-cluster.

Cluster 4.2 players have average skill set scores that are very similar to the cluster average apart from the off-possession PC3 score, which correlates highly with winning the ball in the opposition half. In this aspect, they score much lower than the other sub-clusters. Pierre-Emerick Aubameyang of Arsenal and Jamie Vardy of Leicester City are two high-profile examples in this sub-cluster.

Cluster 4.3 is made of only eight players, and this sub-cluster has an extremely high average score in shot involvement PC1. Therefore, these players have a very high output in finding goal scoring chances. Cristiano Ronaldo of Real Madrid and Harry Kane of Tottenham Hotspur are two high-profile players in this sub-cluster.

In contrast with Cluster 4.1, Cluster 4.4 players have the highest creative skill and the lowest off-possession PC2 score within this cluster. Hence, they are very good at creating chances for others yet are not very effective in the air. They also have the second-highest average score in shot involvement PC1. Roberto Firmino of Liverpool and Raheem Sterling of Manchester City are two well-known examples in this sub-cluster.

Ball-dominant Creator Forwards

Cluster 5.1 is made of only seven players. This sub-cluster has the highest average score in creative skill PC1 and the second-highest average score in shot involvement PC1 among all sub-clusters. These numbers indicate that these players are the best chance creators in European football and second-best goal threats after Cluster 4.3. Lionel Messi of Barcelona and Eden Hazard of Chelsea are two high-profile examples in this sub-cluster.

Cluster 5.2 players have the third-highest average score in creative skill PC1 and the highest average score in creative skill PC2 score among all sub-clusters. Hence, these players are very competent in both creating chances and playing creative passes that penetrate opponent defences. Kevin de Bruyne of Manchester City and Mesut Ozil of Arsenal are two well-known examples in this sub-cluster.

Creator Forwards

Cluster 6.1 players have the second-highest average score in creative skill PC1 and the second-lowest average score in creative skill PC2 score among all sub-clusters. Therefore, they are very good at chance creation and performing long dribbles against opponent defenders. They also have the highest shot involvement PC1 score within this cluster. Arjen Robben of Bayern Munich and Willian of Chelsea are two high-profile examples in this sub-cluster.

Cluster 6.2 players have the lowest average score in creative skill PC2 score among all sub-clusters, so they include very good dribblers of the ball as in Cluster 6.1. Yet, their creative skill PC1 score is not as high, so they do not contribute that highly in terms of chances created. Theo Walcott of Everton and Victor Moses of Chelsea are two well-known examples in this sub-cluster.

Cluster 6.3 players have the highest playmaking PC1 and off-possession PC3 score within the cluster, so they help with the build-up of the game and with pressing the opponent in their own half. Ilkay Gundogan of Manchester City and Corentin Tolisso of Bayern Munich are two high-profile examples in this sub-cluster.

Cluster 6.4 players have the highest average score in creative skill PC1 and PC2 scores within the cluster. As is the case with Cluster 5.2, these players combine chance creation and playing creative passes in the final third of the pitch. Dele Alli of Tottenham Hotspur and Henrikh Mkhitaryan of Arsenal are two well-known examples in this sub-cluster.

Cluster 6.5 players have the highest shot involvement PC1 score within the cluster. In all the other aspects, their scores are close to the average scores of the cluster. Bernardo Silva of Manchester City and Erik Lamela of Tottenham Hotspur are two high-profile examples in this sub-cluster.

4.3.5. *Discussion*

In the decision-making process related to player recruitment and development, decision-makers need to have an optimal set of key performance indicators (Lawrence, 2018). Moreover, the presence of too many variables would negatively affect the cluster analysis as a high number of dimensions would lead to distances being relatively uniform (Beyer et al., 1999). In this study, the combination of AHC and PCA on technical performance variables enables us to identify a small number of crucial and interpretable skill sets in top-level football. These skill sets made it possible to create a data-driven taxonomy of technical player functions.

The structures of clusters based on the first component of four PCA stages are presented in Figures 4.9 and 4.10. These components were chosen to demonstrate the clusters because they account for the highest amount of variance in the four main areas of specialization in football. The score plots suggest that a division of tasks exists between clusters generated. Specifically, some creator forwards and ball-dominant creator forwards seem to overlap, as is the case with numerous all-arounders and roaming playmakers. Hence, it could be possible to detect players with playmaking abilities among sub-groups of all-arounders as well.

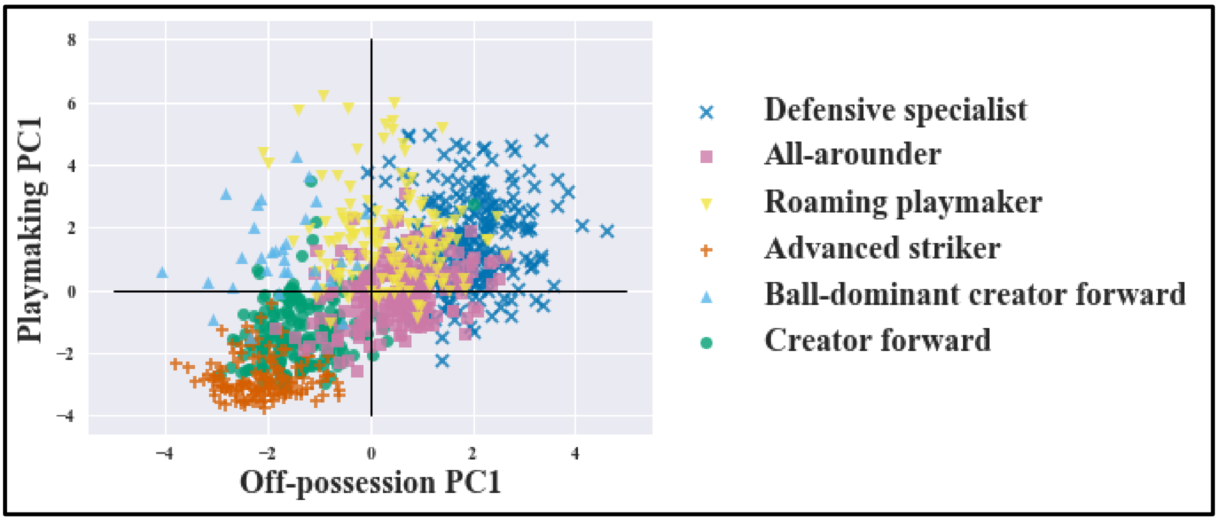

Figure 4.9 Plotted values of the first principal components generated in off-possession and playmaking stages.

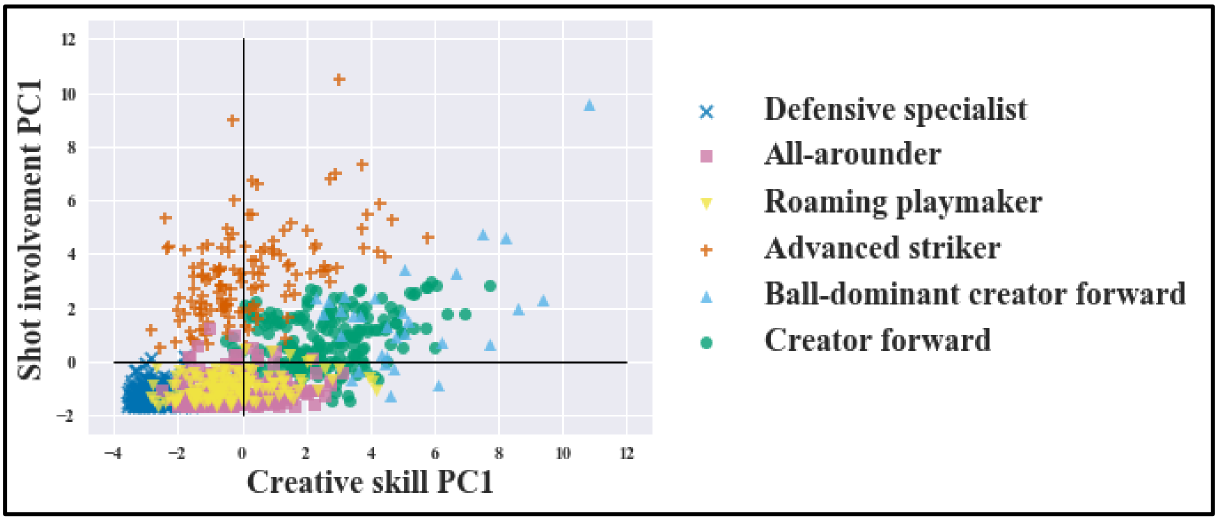

Figure 4.10 Plotted values of the first principal components generated in creative skill and shot involvement stages.

In Figure 4.9, there are some defensive specialists who get highly involved in the build-up of the game. These findings are in line with the trend of central defenders becoming comfortable actors in ball possession and playmaking. As discussed in the previous chapter, the fit between the football identity of a team and its player characteristics is crucial in building a squad (Ancelotti et al., 2016; Keller, 2014). Among defensive specialists, the sub-cluster 1.2 and 1.3 players seem to be more suitable for an elaborate attacking style thanks to their high scores in playmaking. Similarly, the roaming playmaker sub-clusters and the sub-cluster of all-arounder 2.1 have players with high playmaking scores. These groups can be relatively useful in discovering players that have high football identity fit for elaborate attacking styles, such as Bayern Munich and Barcelona.

On the other hand, teams that have direct attacking styles could benefit from sub-clusters with low average creative skill PC2 scores. As these players tend to have high output in long dribbles, they can be effective in fast and direct attacking situations. For example, the sub-clusters of creator forward 6.1 and 6.2 have the highest average scores in this area. While the players in these groups can operate higher up the pitch, all-arounder 2.1 and 2.3 may potentially offer external defenders that are able to move the ball forward from deeper areas.

In terms of defensive playing styles, a crucial variable is off-possession PC3 score, which is highly correlated with ball recoveries in the opposition half. The sub-clusters that have high scores in this area may help to source players that have a high fit with high pressure defending styles, such as Liverpool and Manchester City. The sub-clusters of roaming playmakers and ball-dominant creator forwards have the highest average score in off-possession PC3. Among creator forwards, creator forward 6.3 scores very highly in this aspect. Since the development of a specific playing style in a football club can be a vital source of competitive advantage, the assessment of potential recruits in terms of football identity fit can be viewed as a key benefit of the developed taxonomy created.

Primary Positions and Cluster Membership

Table 8 Distribution of cluster members within five major positional groups. Each player is represented in the position that he is most frequently deployed in.

Clusters / Primary positions	N	CD	ED	CM	EM	CF
Defensive specialist	231	96%	5%	10%	0%	0%
All-arounder	199	1%	82%	16%	5%	2%
Roaming playmaker	136	3%	5%	46%	2%	0%
Advanced striker	134	0%	0%	3%	11%	83%
Ball-dominant creator forward	31	0%	1%	6%	6%	3%
Creator forward	189	0%	6%	18%	76%	13%

N: cluster size, CD: central defender, ED: external defender, CM: central midfielder, EM: external midfielder, CF: central forward

Table 8 shows the distribution of main cluster members within the positional groups of players according to the framework most frequently utilized in sports science literature (Bradley et al., 2009; Dellal et al., 2011; Di Salvo et al., 2007). Here, the most homogeneous positional group in terms of cluster membership is defensive specialists with 96%. After defensive specialists, the advanced striker cluster follows with 83% of all central forwards. Further, 82% of external defenders are in the all-arounder group, whereas 76% of all external midfielders are identified as creator forwards. Central midfielders have the most diverse positional group, with roaming playmaker as the most frequently observed cluster with 46%. This finding is in line with the proposition that it is a challenging task to form a balanced central midfield combination (Gullit, 2016). On the distribution of central midfielders into clusters, it is interesting to see that 10% of central midfielders belong to the defensive specialist cluster dominated by central defenders. These midfielders include Sergio Busquets and Ander Herrera, who are highly competent in both ball-winning and playmaking. Hence, it can be inferred that a part of deep-lying playmakers can be conceptualized as defensive specialists deployed in the central midfield.

In accordance with the trend of external defenders being more active in playmaking and chance creation (Wright, 2017), 5% of external defenders are in the roaming playmaker cluster, whereas 6% fall within the creator forward group. Moreover, Joshua Kimmich and Marcelo are two external defenders who feature in the ball-dominant creator forward group, owing to their high creative passing, chance creation and dribbling output. When managing the talent cycle, the discovery of such players can be influential in building a squad with a high offensive performance capacity. Also, they could be fielded together with teammates that can cover for them when they join attacks in advanced positions.

Additionally, 13% and 3% of central forwards feature in the creator forward and ball-dominant creator forward clusters, respectively. For instance, Diego Costa and Ante Rebić are both tall and physically strong players whose primary position is centre forward. As a result, they are likely to be viewed as advanced strikers or target men. Yet, they are categorised as creator forwards because of their output in creating chances for their teammates and outplaying opponents with dribbles. This finding shows that heuristics based on the physical attributes of players can be misleading in assigning functions to players. In addition, Eden Hazard and Lionel Messi are two examples of ball-dominant creator forwards whose primary position is central forward. Based on the differentiation between advanced and versatile strikers suggested by FIFA (2018), it can be argued that these two clusters could help to detect the mentioned versatile strikers.

Market Valuation of Clusters

In Table 9, the median wages of the main cluster members and the wages of their highest earners are presented. Ball-dominant creator forwards have the highest median wages with a median value of €6.60 million. This cluster has several top earners of numerous wealthiest clubs in Europe. For instance, Kevin de Bruyne of Manchester City FC, Lionel Messi of Barcelona FC and Mesut Özil of Arsenal FC are members of this cluster. The following clusters with the highest median wages are advanced striker, roaming playmaker and creator forward, respectively. An interesting finding here is that roaming playmakers have a slightly higher valuation than creator forwards. This underlines how the major football leagues are catching up in crediting the contributions of a specific breed of playmakers. High-profile examples from this cluster are 2018 Ballon d'Or winner Luka Modric and Barcelona FC legend Andrés Iniesta. On the other hand, the lowest median wages are recorded by all-arounders. The relatively low valuation of this cluster demonstrates that players who do not excel in one area are penalized by the market. As 82% of external defenders are identified as members of this group, club decision-makers can highly benefit from developing and recruiting external defenders that specialize in at least one area.

Table 9 Median wages and top earner wages of clusters (in €million).

Cluster	Median wages	Top earner	Top earner wages
Defensive specialist	1.75	Sergio Ramos	18.65
All-arounder	1.50	David Alaba	9.95
Roaming playmaker	2.05	Toni Kroos	20.00
Advanced striker	2.50	Cristiano Ronaldo	35.20
Ball-dominant creator forward	6.60	Lionel Messi	40.00
Creator forward	2.00	Thomas Müller	14.95

4.4. A Case Study in Player Recruitment: Manchester United

4.4.1. Manchester United: Profile

Manchester United are the highest-revenue generating club and the biggest spender of wages in English football (Deloitte Sports Business Group, 2021). The club has the highest number of first-tier titles in England and won the top UEFA competition in the world, the Champions League, three times (Bird, 2022). Despite this financial and historical success, the club won the Premier League the last time in the season 2012/13, in the final season of its legendary manager, Sir Alex Ferguson (Bird, 2022). Figure 4.11 demonstrates the financial gap between the five highest-revenue generating clubs in EPL and the rest. Here, we see that Manchester United fare below the regression line, which means that they are underperforming with regards to their wage expenditures. Here, we see that teams like Manchester City and Tottenham Hotspur are performing much better while operating on smaller wage budgets. In the next section, the performances of Manchester United are analysed with a view to isolating the results from the performances thanks to the expected goal framework.

Figure 4.11 Wage costs and points won by EPL clubs in the season 2017/18. Clubs known as "Top 6" are represented with abbreviations.

4.4.2. Manchester United: 2017/18 Performance

This section focuses on the analysis of Manchester United's 2017/18 performance in EPL. Table 10 presents the points won as well as actual and expected goals created and conceded by EPL teams during the season 2017/18. The purpose of this table is to analyse if there was a considerable gap between the performances and the results of Manchester United. As explained in the literature review, the expected goal framework focuses on the quality of the goal chances created and conceded in a football game and allows for isolating the performances from the match results (Brechot & Flepp, 2020). As football is a low-scoring sport, the randomness factor plays a higher role than in high-scoring sports such as basketball and rugby.

At the end of the analysed season, Manchester United finished 2nd in the league. To see how the team fared in attack and defence, we look at goals scored and conceded. According to the table, the team scored 1.65 goals per game and conceded 0.69 goals per game. The goal analysis shows that in terms of results, they had the 5th highest value in goals scored and the 2nd lowest value in goals conceded. As a result, the goal difference ranking of the team is 3rd

Table 10 Analysis of the 2017/18 season performances in the EPL.

Club	Points	GF	GC	GD	xGF	xGC	xGD
Manchester City	100	2.51	0.66	1.85	2.31	0.67	1.64
Manchester United	81	1.65	0.69	0.96	1.72	1.17	0.55
Tottenham Hotspur	77	1.73	0.89	0.84	1.80	1.11	0.69
Liverpool	75	2.03	0.94	1.09	1.95	0.93	1.02
Chelsea	70	1.48	0.94	0.54	1.61	1.08	0.53
Arsenal	63	1.80	1.26	0.54	1.90	1.34	0.57
Burnley	54	0.86	0.96	-0.10	1.08	1.50	-0.42
Everton	49	1.04	1.43	-0.39	1.19	1.50	-0.31
Leicester City	47	1.30	1.47	-0.17	1.36	1.41	-0.05
Newcastle United	44	0.94	1.14	-0.20	1.21	1.56	-0.36
Crystal Palace	44	1.08	1.35	-0.27	1.58	1.30	0.28
Bournemouth	44	1.10	1.49	-0.39	1.19	1.75	-0.56
West Ham United	42	1.17	1.66	-0.49	1.19	1.66	-0.47
Watford	41	1.02	1.58	-0.56	1.31	1.48	-0.17
Brighton	40	0.82	1.34	-0.52	1.08	1.58	-0.51
Huddersfield	37	0.64	1.43	-0.79	0.87	1.39	-0.52
Southampton	36	0.88	1.38	-0.50	1.21	1.38	-0.18
Swansea City	33	0.67	1.38	-0.71	0.90	1.67	-0.76
Stoke City	33	0.83	1.67	-0.84	1.05	1.74	-0.69
West Bromwich Albion	31	0.76	1.38	-0.62	1.01	1.30	-0.30

GF = goals for, GC = goals conceded, GD = goal difference, xGF = expected goals for, xGC = expected goals conceded, xGD = expected goal difference

with 0.96 goals per game behind Liverpool, which has a goal difference of 1.09 and the champions Manchester City, which ended up with a goal difference of 1.85 goals.

Next, we examine the expected goals that teams recorded to be able to have a deeper understanding of their attacking and defensive performances. In terms of expected goals scored, Manchester United rank 5th with 1.72 goals. This number is very close to the goals scored value of 1.65 per game, and the actual and expected goal rankings are the same. Therefore, we conclude that there is not a big gap between the results and performances in this aspect. As a result, the attacking side of the game should be considered a major area for improvement for the team.

To assess the defensive performance, we look at goals conceded and expected goals conceded. During the season 2017/18, the team conceded 0.69 goals per game. This value is very close to that of Manchester City, which is 0.66 goals per game. However, Manchester City conceded 0.67 expected goals per game, which is very close to the goals conceded value. Therefore, it can be concluded that their defence is by far the best in terms of the quality of chances conceded. On the other hand, Manchester United conceded 1.17 expected goals per game, which is almost twice the value of goals conceded. This gap shows that the opposing teams have found good goal scoring chances that they were not able to convert to goals. This can be attributed to a combination of the quality of goalkeeping and the lack of quality of the opposing teams' attackers in front of the goal. Despite the fact that the team's starting goalkeeper David de Gea is known to be one of the best goalkeepers in the world, the level of performance that he showed in 2017/18 might not be sustainable. Consequently, the team had the 5th best defensive performance in the league while having the 2nd best results, and the defensive side of the game also needs major improvements to challenge for the first spot in the league the following season.

To be able to analyse the overall performance of teams, xG difference per game can be considered a good metric. This metric looks at the difference between the quality of chances found and conceded. In this metric, Manchester City is by far the best team with 1.64 xG per game, whereas Liverpool comes second with 1.02 xG per game. The third team is Tottenham Hotspur, with 0.69 xG per game. The rest of the top six revenue-generating teams are quite close to each other, with Arsenal, Manchester United and Chelsea recording 0.57, 0.55 and 0.53 xG per game, respectively. These figures show the level of the competition within the Top 6 and underline how difficult it is to win this league.

Figure 4.12 provides an overview of the playing styles of teams competing in the EPL in the analysed season. As per the playing style framework presented in Chapter 3, attacking styles can be described as elaborate and direct (Castellano & Pic, 2019). On the other hand, teams can prefer defensive styles that are deep defending or high pressure defending (Castellano & Pic, 2019). Here, two variables are chosen to reflect the playing styles defined in the framework: possession percentage and PPDA. Possession percentage indicates the share of the possession of a team during the season. Teams that play direct tend to have a relatively lower possession percentage, whereas teams that have elaborate attacks have a higher share of possession. PPDA is the number of opponents' passes allowed per defensive action, where a relatively lower PPDA indicates a high pressing intensity. Therefore, we can expect teams with lower PPDA values to have high pressure defending styles, whereas teams with higher PPDA figures tend to prefer deep defending approaches.

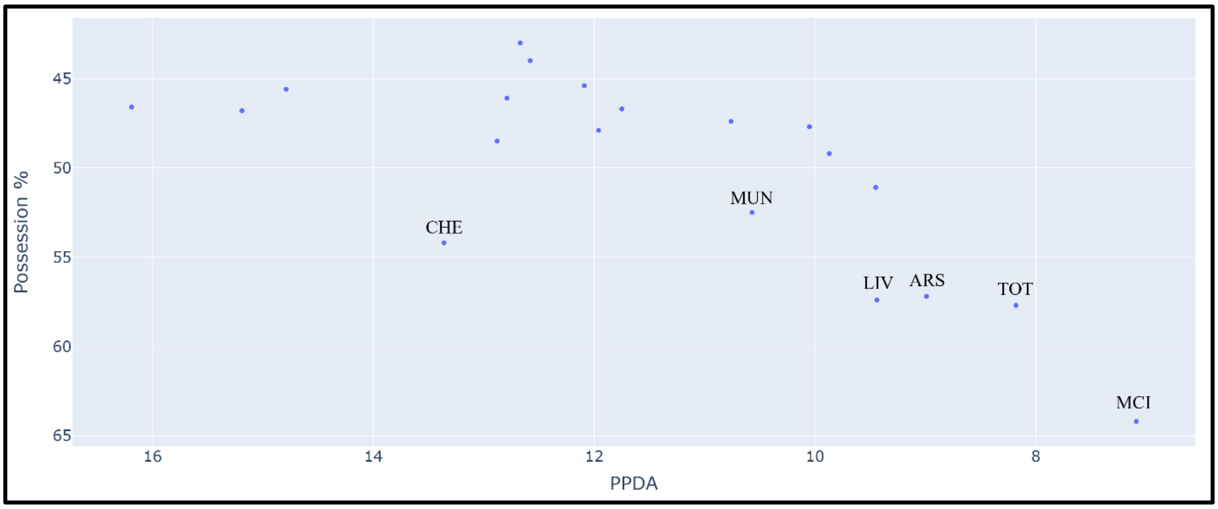

Figure 4.12 Football identity map of EPL clubs in the season 2017/18. Clubs known as "Top 6" are represented with abbreviations.

This figure shows that there is a considerable stylistic gap between Manchester United and four of the highest revenue-generating EPL clubs, Arsenal, Liverpool, Tottenham Hotspur and Manchester City. With the addition of Chelsea, these clubs are known as the “Top 6” as a result of the gap in the recent performance and success levels between the six clubs and the rest. At the bottom right of the figure, Manchester City can be seen as the team that has the highest possession share and the highest pressing intensity. Three other teams, Arsenal, Liverpool and Tottenham Hotspur, make up a group with varying PPDA levels and a very high possession percentage. Therefore, these teams seem to capitalise on their ability to press the opponent and win the ball, in addition to keeping more of the ball than their opponents. In these two aspects, we see that Manchester United are quite far apart from the other four teams. Considering that Liverpool and Tottenham Hotspur have a bigger xG difference than Manchester United and more similar playing styles to Manchester City, it can be said that these two teams seem to be better candidates to challenge the 2017/18 champions the following season.

4.4.3. The Selection of Recruitment Candidates

In this section, recruitment alternatives for Manchester United in five positional groups will be outlined. As in the previous chapter, the unit of analysis covers the season 2017/18 in German Bundesliga, English Premier League and Spanish La Liga. To be able to focus on the most attractive clusters, three squads of the league champions are considered as the benchmark: Barcelona, Bayern Munich and Manchester City. Based on the number of minutes played within these squads, clusters of the most-utilised ten outfield players (i.e. starters) are pinpointed (see Appendix C for the list of starting players of Barcelona, Bayern Munich and Manchester City in 2017/18). This is because every starting eleven is composed of a goalkeeper and ten outfield players. Later, the identified clusters are explored to discover similar players to the benchmarked players in each positional group.

As discussed in Chapter 3, clubs need to effectively use different age groups to be able to build competitive squads. While pre-peak players have the potential to grow into star players within the club and to bring future transfer profits, peak and post-peak players are seen as less risky investments thanks to their performance records over longer durations. Based on the literature produced on age and market valuation (Kalén et al., 2019), three players under the age of 26 (i.e. pre-peak) and three players aged 26 and over (i.e. peak and post-peak) are pinpointed for each position. Therefore, the selection combines a solid base for recruitment in terms of age profiles.

For external defenders and midfielders, left-sided and right-sided options are presented separately since these players tend to specialise in operating on one side of the pitch. In the central midfield, two types of central midfielders are used in line with Dellal et al. (2011). However, instead of calling them central defensive midfielders and central attacking midfielders, I use the term No.6/8 and No.8/10. As explained in the literature review, there is a consensus in the football world on three types of central midfield positions: No.6, 8 and 10. As many central midfielders are flexibly used in defensive and central, or central and attacking roles, I utilise the relevant shirt numbers attributed to these positions. Since top-level players are multifunctional players nowadays, terms such as defensive and attacking can be misleading in defining players. The proposed positions reflect this versatility. For instance, the position of N'Golo Kanté is characterised as No.6/8, whereas Kevin de Bruyne's position is defined as a No.8/10 in the selection.

As pointed out at the beginning of this chapter, there is a huge financial gap between the top 5 revenue-generating clubs and the rest of the clubs in EPL. Since the other four teams in EPL are direct competitors, it would be very difficult for Manchester United to recruit from these clubs. Therefore, the players of Arsenal, Chelsea, Liverpool and Manchester City are not taken into consideration in the case study. Another limitation utilised in the study is

Table 11 Principal components, their abbreviations and the variables with the strongest correlations.

Principal Component	Abbreviation	Key Variable
Playmaking PC1	PLAY1	Successful forward pass
Creative skill PC1	CREA1	Expected assist
Creative skill PC2	CREA2	Successful through pass (Long dribble*)
Shot involvement PC1	SHOT1	Expected goal
Off-possession PC1	OFF1	Interception
Off-possession PC2	OFF2	Aerial duel won
Off-possession PC3	OFF3	Opponent half recovery

* Long dribble has a high negative correlation with CREA2 scores.

defined based on the individual player wages. A maximum of 10 million euros per annum is defined as a wage threshold as it can be argued that players earning more than this figure, such as Lionel Messi, Cristiano Ronaldo and Luka Modric have already proven to be world-class players. As a result, 38 players are discarded from the selection.

In this case study, PC scores will be represented with abbreviations for ease of reading. Table 11 provides a list of these abbreviations as well as the key variables that have the strongest correlation with the utilised PC scores.

The six players that were chosen for each position are presented in Table 12. Among these players, one option from the pre-peak age group and one from the peak and post-peak age group per playing position are compared in Figures 4.13 and 4.14. The rationale behind the selection of these pairs is explained in the position-based sub-sections below.

Table 12 The selected recruitment candidates for each playing position.

Player	Club	Cluster	Age	Wages	PLAY1	CREA1	CREA2	SHOT1	OFF1	OFF2	OFF3
Central defenders											
Mats Hummels	Bayern Munich	1.3	29	10.0	4.77	-2.20	1.05	-1.06	3.33	1.79	1.91
Marc Bartra	Real Betis	1.3	27	3.6	3.14	-1.00	0.84	-0.79	3.84	1.64	1.99
Niklas Süle	Bayern Munich	1.3	22	3.0	3.65	-2.58	0.17	-0.81	1.23	1.45	0.99
Dayot Upamecano	RB Leipzig	1.3	19	2.4	1.97	-2.65	0.01	-0.84	3.11	1.38	-0.43
Willi Orban	RB Leipzig	1.3	25	2.0	2.02	-2.60	0.26	-1.01	3.03	1.64	0.09
Zouhair Feddal	Real Betis	1.3	28	0.7	2.20	-2.72	0.39	-0.65	2.66	1.52	-0.03
External defenders (Left)											
Jordi Alba	Barcelona	2.1	29	8.6	1.45	2.25	1.24	-1.06	1.13	-1.30	-0.43
Marcelo	Real Madrid	5.2	30	8.3	2.88	6.10	0.80	-0.84	-1.06	-0.98	2.92
Filipe Luis	Atletico Madrid	2.1	32	4.7	1.91	0.37	-0.52	-1.30	1.96	-1.38	-0.59
Jose Gaya	Valencia	2.1	23	3.3	0.50	1.01	-0.67	-1.49	0.39	-0.75	0.57
Douglas Santos	Hamburg	2.1	24	2.4	0.97	1.02	-1.48	-1.13	1.98	-0.07	1.80
Jonny	Celta	2.1	24	1.7	0.89	0.24	-0.13	-1.21	2.33	-0.49	0.17
External defenders (Right)											
Daniel Carvajal	Real Madrid	2.1	26	8.3	1.72	2.01	-0.89	-1.16	0.86	0.05	2.02
Rafinha	Bayern Munich	2.1	32	5.0	3.08	1.32	-0.21	-1.29	0.66	-1.08	1.13
Kieran Trippier	Tottenham Hotspur	2.1	27	2.3	0.77	2.22	1.57	-1.62	0.81	-0.30	0.83
Joshua Kimmich	Bayern Munich	5.2	23	1.5	2.49	4.63	1.04	-1.27	-0.10	-0.66	0.78
Benjamin Henrichs	B. Leverkusen	2.1	21	1.5	0.45	1.41	-0.52	-1.33	1.72	-0.64	-0.45
Pablo Maffeo	Girona	2.1	20	0.3	-1.41	1.18	-1.66	-1.21	0.61	0.00	0.87
Central midfielders (No.6/8)											
Ever Banega	Sevilla	3.2	30	6.5	6.22	3.01	1.27	-0.27	-0.93	-0.52	2.51
Mateo Kovacic	Real Madrid	3.2	24	6.2	3.33	1.56	-0.18	-1.36	-0.41	-0.44	2.92
Dani Parejo	Valencia	3.2	29	5.3	5.30	0.92	1.68	0.40	0.28	-0.47	0.99
Eric Dier	Tottenham Hotspur	1.3	24	3.5	3.53	-1.51	1.12	-1.18	0.79	0.80	0.65
Nuri Sahin	B. Dortmund	3.2	29	2.5	4.82	2.00	4.15	0.01	0.24	0.66	3.42
Mikel Merino	Newcastle United	3.1	22	2.1	2.41	0.25	0.16	-0.75	1.18	2.16	4.96

Central midfielders (No.8/10)											
Mahmoud Dahoud	B. Dortmund	5.2	22	5.0	0.63	3.42	2.77	-0.66	-1.11	-0.57	2.25
Naby Keita	RB Leipzig	5.2	23	4.0	0.80	4.53	0.40	0.18	-0.27	-0.33	2.85
Sergio Canales	R. Sociedad	5.2	27	1.6	0.63	4.49	-0.18	1.34	-1.67	-1.23	0.64
Pablo Fornals	Villareal	5.2	22	1.6	0.22	4.74	4.13	-0.27	-1.07	-0.74	-0.35
Jonathan Viera	Las Palmas	5.2	28	1.5	2.01	6.21	2.89	0.69	-2.27	-0.76	2.75
Emre Colak	Deportivo	5.2	27	0.8	1.33	4.44	3.96	-0.07	-1.72	-1.74	0.61
External midfielders (Left)											
Nolito	Sevilla	5.2	31	6.6	-1.52	3.69	1.99	1.72	-2.35	0.91	3.16
Karim Bellarabi	B. Leverkusen	6.1	28	5.0	-1.45	5.75	-1.39	1.23	-0.72	-0.80	-0.68
Heung-Min Son	Tottenham Hotspur	4.4	26	5.0	-1.81	2.95	-0.18	3.47	-2.52	-0.95	0.42
Yannick Carrasco	Atletico Madrid	6.1	24	3.0	-2.10	4.21	-2.16	1.18	-2.04	-2.01	1.17
Goncalo Guedes	Valencia	6.1	21	2.5	-1.12	6.00	-0.54	0.55	-1.93	-1.42	0.63
Leon Bailey	B. Leverkusen	6.1	20	2.0	-1.34	5.23	-1.40	2.59	-1.75	-0.88	-0.44
External midfielders (Right)											
Marco Asensio	Real Madrid	6.1	22	8.4	0.67	5.73	-0.17	2.79	-2.22	-1.21	-0.70
Kingsley Coman	Bayern Munich	6.1	22	8.0	-1.63	7.71	-5.07	2.87	-2.17	-1.00	-0.77
Arjen Robben	Bayern Munich	6.1	34	7.0	-0.66	6.04	-0.82	2.86	-2.07	-1.69	-1.11
Riyad Mahrez	Leicester City	6.1	27	5.9	-0.02	5.39	0.95	2.60	-1.57	-0.71	0.18
Lucas Vazquez	Real Madrid	6.1	27	3.7	-0.51	6.41	0.13	1.79	-1.03	-0.47	0.62
Angel Correa	Atletico Madrid	6.5	23	1.9	-1.47	3.05	-0.56	2.38	-1.25	-0.33	0.96
Central forwards											
Harry Kane	Tottenham Hotspur	4.3	24	5.9	-2.83	0.30	0.28	6.70	-3.25	1.75	-0.44
Serge Gnabry	Hoffenheim	4.4	22	4.5	-2.11	3.77	-0.44	4.26	-1.24	-0.95	-1.12
Max Kruse	Werder Bremen	5.2	30	3.0	0.29	3.20	1.29	2.42	-3.17	-0.03	2.91
Lars Stindl	B. M’gladbach	5.2	29	3.0	0.32	2.29	2.64	2.36	-0.81	0.47	2.43
Andrej Kramaric	Hoffenheim	4.4	27	2.5	-1.24	4.43	0.57	3.93	-2.90	-1.48	-0.19
Luka Jovic	E. Frankfurt	4.3	20	0.4	-2.87	-0.25	0.80	6.01	-3.02	1.02	-0.87

Figure 4.13 Comparison of the representative (in red) and the most expensive players (in blue) of the first four positions analysed.

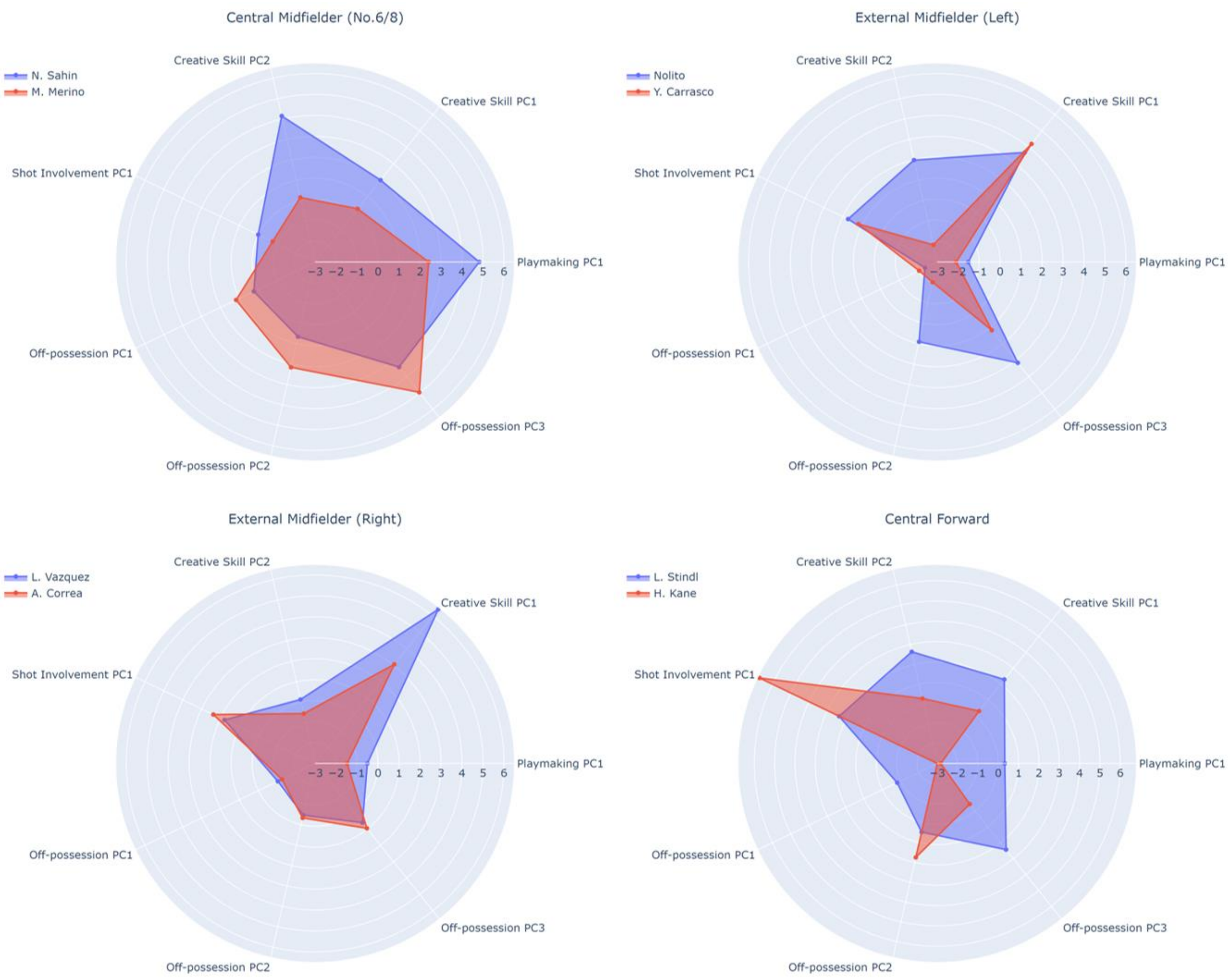

Figure 4.14 Comparison of the representative (in red) and the most expensive players (in blue) of the next four positions analysed.

Central Defenders

According to the benchmark analysis, five of the six starters from the three analysed teams belong to the cluster defensive specialist 1.3. These players combine playmaking, ball-winning and aerial ability skills. Therefore, cluster 1.3 is defined as the focal point and players that score highly in PLAY1, OFF1 and OFF2 are chosen. On the other hand, the starters of Manchester United, Chris Smalling and Phil Jones, are in the cluster defensive specialist 1.1. This cluster has a relatively lower playmaking score than cluster 1.3. As a result, improvements in this position can be very beneficial for the team, as technically excellent central defenders can help to adopt an elaborate attacking approach, as is the case with Manchester City and Liverpool.

In the peak and post-peak age group, the selected players are Mats Hummels, Marc Bartra and Zouhair Feddal. Here, Mats Hummels would be the most expensive and attractive choice based on his playing experience in Bayern Munich and the German national team. In the pre-peak age group, Niklas Süle, Dayot Upamecano and Willi Orban are chosen. Süle and Upamecano are 22 and 19 years old with annual wages of 3 and 2 million euros respectively. Recruiting these players could be a very profitable investment in the long run.

Since all these players have high scores in PLAY1, they would be suitable for an elaborate style play in attack. Also, Hummels and Bartra have high scores in CREA2, which correlates highly with penetrative passes that outplay several opponents at once. Hence, two players seem to have a higher technical ability than the rest of the players on the list.

Moreover, Hummels, Bartra and Süle have very high scores in OFF3, which is highly correlated with opponent half recoveries. Therefore, these players would have a high fit with a high pressure defending style. Consequently, three players seem to be ideal central defender options for an elaborate and high pressure defending playing style. As he is on 3 million euros per annum, Süle represents a more affordable option than Hummels in terms of wage costs.

External Defenders

The benchmark analysis shows that four of the six starting external defenders belong to the cluster all-arounder 2.1. In addition, Fabian Delph of Manchester City is a roaming playmaker 3.1, which brings together players with higher defensive output among the group of roaming playmakers. On the other hand, Joshua Kimmich is one of the few external defenders in the ball-dominant creator forward 5.2 cluster. Based on this mix, these three clusters are explored with a focus on OFF1, PLAY1 and CREA1 scores. The Manchester United starters in this position, Antonio Valencia and Ashley Young, are from this group as well. However, they are past their peak ages and are 32 and 33 years old, respectively. Therefore, the team would benefit highly from recruiting peak and pre-peak age players in this position.

(1) External Defenders: Left

The selected players from the peak and post-peak age groups are Jordi Alba, Marcelo and Filipe Luis. Among this group of players, Alba scores highly on all the three scores that the benchmark analysis provides and has annual wages amounting to 8.6 million euros. Marcelo is another rare external defender in the ball-dominant creator forward cluster with very high PLAY1 and CREA1 scores, yet his OFF1 score is relatively low. This shows that he excels on the attacking side of the game, whereas his defensive side can be seen as a weakness. In contrast, Luis has a very high OFF1 score, but his low CREA1 score points to a relatively lower potential to create goalscoring chances for his teammates.

In the pre-peak age group, the selected players are Jose Gaya, Douglas Santos and Jonny. Among this group of players, Santos has high scores in all three identified key aspects.

Moreover, his OFF3 score is very high, so he shows the ability to press the opponent effectively. Jonny scores highly on OFF1 yet has relatively lower scores on the two other on-possession aspects of the game. On the other hand, Gaya has a high CREA1 score, but his PLAY1 and OFF1 output are relatively lower. Therefore, Santos can be seen as the low-cost alternative to Alba, with annual wages of 2.4 million euros.

(2) External Defenders: Right

In the peak and post-peak age group, the selected players are Daniel Carvajal, Rafinha and Kieran Trippier. Daniel Carvajal is the player that stands out in the group with high scores in all areas, in addition to having the highest OFF3 score. He is also the most expensive option with 8.3 million euros per annum. Rafinha has very high scores in the two on-possession aspects yet has a relatively lower OFF1 score. Although Tripper has a very high CREA1 score, his PLAY1 and CREA1 output are relatively lower.

The selected players in the pre-peak age group are Joshua Kimmich, Benjamin Henrichs and Pablo Maffeo. In this group, there is no player that excels in all the three identified skill sets. However, although Kimmich's OFF1 score is moderate, it can be argued that his excellent on-possession abilities can compensate for his low defensive output. His annual wages are quite low compared to those of Carvajal, as he is on 1.5 million euros. Henrichs has very high scores in CREA1 and OFF1, but his PLAY1 output is relatively lower. On the other hand, Maffeo has a very high CREA1 and high OFF1 output yet records a very low PLAY1 score.

Central Midfielders

(1) No.6/8

According to the benchmark analysis, two of the four starters in the No.6/8 position belong to the cluster roaming playmaker 4.2, which brings together players that are skilled in playmaking, performing penetrative passes and winning the ball in the opposition half. In addition, Javi Martinez of Bayern Munich is from the cluster roaming playmaker 3.1. Lastly, Sergio Busquets of Barcelona is a rare central midfielder from the cluster defensive specialist 1.3 that is dominated by central defenders. Therefore, these three clusters are explored with a view to identifying players that score highly in PLAY1, CREA2 and OFF3 scores. The only Manchester United starter in this position is Nemanja Matic. Matic is another rare defensive specialist 1.3 that is deployed in central midfield. It can be argued that a relatively more multifunctional No.6/8 can be a useful addition to the team.

The selected players from the peak and post-peak age groups are Ever Banega, Dani Parejo and Nuri Sahin. All of these players score highly in all three identified key areas. As Sahin has higher OFF1 and OFF2 scores, he can be viewed as the most multifunctional option that can balance the off-possession and on-possession sides of the game. Hence, he seems to be the most attractive option in this position and is relatively affordable, with annual wages of 2.5 million euros.

In the pre-peak age group, the selected players are Mateo Kovacic, Eric Dier and Mikel Merino. In this group, there is no player that scores highly in all the three identified skill sets. Dier has a relatively lower output in OFF3, whereas Merino and Kovacic have a moderate score in CREA2. As Merino has a much higher defensive output thanks to his OFF1 and OFF2 scores, he seems to be the relatively multifunctional option for this position. He represents another low-cost recruitment option with 2.1 million euros per annum.

(2) No.8/10

The benchmark analysis shows that all three players deployed in this position are ball-dominant creator forwards. This group of players scores highly on the creative side of the game and has the ability to win the ball in the opposition half. Hence, this cluster is explored with a focus on CREA1, CREA2 and OFF3 scores. Compared with other positional groups, Manchester United are relatively strong in this area with the two starting players of Paul Pogba and Alexis Sanchez.

In the peak and post-peak age group, the selected players are Sergio Canales, Jonathan Viera and Emre Colak. Among this group, Viera scores highly on all the three scores identified. He also represents a low-cost option with 1.5 million euros per annum. Canales has a relatively lower output in CREA2, whereas Colak's OFF3 score is lower than the other two options.

The selected players in the pre-peak age group are Mahmoud Dahoud, Naby Keita and Pablo Fornals. In this group, Dahoud is the most attractive option as he excels in all three key aspects that the benchmark analysis provides. Despite being only 22 years old, his annual wages are 5 million euros per annum. Keita has a relatively lower CREA2 score, whereas Fornals has a low output in OFF3.

External Midfielders

According to the benchmark analysis, this is the most diverse positional group in terms of the player clusters. Here, there are two ball-dominant creator forwards, one roaming playmaker 3.2, one creator forward 6.1, one creator forward 6.5 and one advanced striker 4.4. The commonly observed high scores are CREA1, SHOT1 and OFF3, which correlate with the quality of chances created, the quality of shots taken and opponent half recoveries. Therefore, we explore these clusters with a focus on the mentioned three scores. In this position, the two starting players of Manchester United are Marcus Rashford and Juan Mata. Rashford is a creator forward 6.5, whereas Mata is a ball-dominant creator forward 5.2. This is another area where the team has high-quality starters. However, Mata is already 30 years old, and external midfield is a position where high performing teams have the necessity to possess depth in numbers. This is because teams with lower resources are inclined to adopt a deep defending defensive style against teams like Manchester United, and these players might be able to pose different kinds of attacking solutions to unlock the opposition. Therefore, the team can benefit from quality additions in this positional group as well.

(1) External Midfielders: Left

The selected players from the peak and post-peak age groups are Nolito, Karim Bellarabi and Heung-Min Son. Among this group, Nolito scores highly on all the key areas identified. He is on 6.6 million euros per annum. Bellarabi and Son have very high scores in CREA1 and SHOT1 yet have a relatively lower output in OFF3.

In the pre-peak age group, the selected players are Yannick Carrasco, Goncalo Guedes and Leon Bailey. Here, Carrasco is the most attractive option as he excels in all three scores that the benchmark analysis provides. His annual wages are 3 million euros, so he represents a relatively affordable alternative to Nolito.

(2) External Midfielders: Right

In the peak and post-peak age group, the selected players are Arjen Robben, Riyad Mahrez and Lucas Vazquez. Among this group, Vazquez scores highly in all the three areas identified. He is on 3.7 million euros per annum. Robben and Mahrez have very high scores in both on-possession scores yet have lower scores in OFF3.

The selected players in the pre-peak age group are Marco Asensio, Kingsley Coman and Angel Correa. Here, Correa is the most attractive option, with very high scores in all the key areas. He is another low-cost recruitment candidate with an annual wage of 1.9 million. As is the case with Mahrez and Robben, Asensio and Coman have relatively low OFF3 scores.

Central Forwards

The benchmark analysis shows that the five players deployed in this position come from three different clusters: advanced striker 4.3, ball-dominant creator forward 5.1 and creator forward

6.4. These three clusters are explored with a focus on SHOT1, CREA1 and OFF2. The combination of these characteristics would lead to identifying a central forward that excels in finding and creating good goal scoring opportunities as well as winning aerial duels. Manchester United's starting central forward, Romelu Lukaku, is an advanced striker 4.2 that has a very high SHOT1 score and high CREA1 and OFF2 scores. As is the case with external midfielders, this is a playing position that can benefit from squad depth in creating different attacking solutions when playing against deep defending teams.

The selected players from the peak and post-peak age groups are Max Kruse, Lars Stindl and Andrej Kramaric. All these players score highly in the on-possession aspects, yet Kruse and Kramaric score relatively low in OFF2. Hence, Stindl is the most attractive option and is on 3 million euros per annum.

In the pre-peak age group, the selected players are Harry Kane, Serge Gnabry and Luka Jovic. In this group, none of the players satisfied the three identified criteria. Gnabry has a very low OFF2 score, whereas Kane and Jovic score relatively low on CREA1. Here, Kane appears to be the most desirable option as he scores relatively high in all three areas than Jovic.

4.5. Managerial Implications and Conclusions

In this chapter, a data analytical model was proposed based on player skill sets and functions identified in the top three European leagues. This research contributes to the sports management field both theoretically and empirically. Examining player skill sets and functions offers a deeper theoretical perspective beyond the conventional player positions for further research in sports management and sciences. The concepts introduced can be used to examine football players in relation to talent management, market valuation and physical performance. Moreover, the methodology can be adapted to other invasion sports when investigating skill sets and functions. Empirically, it was demonstrated that PCA and hierarchical cluster analysis could be helpful in identifying patterns and creating actionable knowledge to enhance organizational performance in sports clubs. Lastly, the case study offered a practical example of how the generated skill set scores and clusters could be beneficial in a football club setting.

The research findings are relevant for decision-makers who assume responsibilities in the talent management cycle, including the transfer and contract negotiations. For chief executive officers, sporting directors and first-team managers, who are accountable for sporting and financial results, a common language that can enable effective resource management on the basis of empirical evidence was provided. Currently, a former domestic league winner with a stadium capacity of over 35,000, Hamburger SV are competing in the second tier due to misspending on player wages and transfer fees. On the other hand, clubs with competent talent evaluation practices in place, such as Sevilla and Borussia Dortmund, have managed to boost their revenue-making potential thanks to net transfer fee income from the sales of the players discovered at early ages and developed within a well-defined sporting strategy.

Via the application of the model to other domestic and international competitions, decision-makers can find playing talent in affordable leagues all over Europe and the world. As demonstrated in the case study, the composition of high performing teams could be used as benchmarks for decision-makers in player recruitment and squad planning.

Evidently, the utilization of five conventionally utilized positions is useful in characterizing players. Yet, the diversity of player functions identified demonstrates that decision-makers in clubs might make costly mistakes based on this simplified framework. Moreover, heuristics based on physical attributes could lead to systematic biases in decisions made regarding the talent cycle (Mills et al., 2018). The proposed framework helps to overcome such cognitive biases on the basis of a data-driven approach as utilizing player skill sets, functions and positions at the same time can lead to a fair comparison between players. As for youth development, the lack of communication between A and youth teams tends to hinder the progression of young players into first teams (Relvas et al., 2010). The framework proposed can help to bridge the gap between the two levels because it enables decision-makers to assign benchmarks and personalized development plans to young players based on skill sets and functions. As well as for football clubs, the findings are applicable to national federations and leagues that oversee the development of footballers at youth levels.

5. Conclusions

5.1. Overview of the Thesis

5.1.1. Research Questions and Findings: Chapter 3

The research question tackled in this chapter was as follows:

- How can a decision-making framework be built to guide player recruitment activities and processes in European football clubs?

The complex nature of this decision-making context made it a challenging task to cover all the different perspectives involved in a player recruitment decision. For instance, the financial implications of such a decision could prove to be more important for a chief executive officer than they are for a first-team manager/head coach. Categorisation of the player recruitment objectives based on merely sporting or financial considerations would be theoretically viable. However, this type of categorisation might have hindered the synergies that could potentially be created on the basis of a multidisciplinary approach. Therefore, the choice of VFT was highly beneficial in providing a holistic conceptual model built on a set of fundamental and means objectives, which emphasised the significance of the potential objectives in their fields. As this method is known to be utilised in complex problem contexts that involve multiple stakeholders (Barclay, 2014; Keeney, 1999a), its application enabled a multi-perspective approach to the identification of objectives. Thanks to the available literature on football management and performance, four fundamental objectives were selected: to maximise sporting results, to effectively use football identity, to minimise wage costs and to maximise net transfer fee income.

Following that, the investigation of means objectives in the existing literature led to the discovery of two types of objectives. The first type was relatively straightforward as it related to five aspects of player characteristics regarding technical performance, physical performance, age, contract type and playing status as per domestic regulations. However, limiting the means objectives to player-related criteria would have lacked the contextual characteristics regarding a club's financial situation and sporting approach. These factors are known to be critical in guiding the player recruitment efforts of a football club (Ancelotti et al., 2016; Bridgewater, 2016; Keller, 2014; Lawrence, 2018). Therefore, another set of means objectives called contextual objectives was included to consider the organisational context while striving to achieve the fundamental objectives. This group of objectives is composed of the following: to construct a football identity, to recruit players based on football identity fit, to recruit players based on financial fit, and to recruit players based on national teams.

5.1.2. Research Questions and Findings: Chapter 4

The second and third research questions of this thesis were as follows:

- How can the technical skill sets of players competing in European football clubs be conceptualised?
- How can the technical functions of players competing in European football clubs be identified?

A major challenge here was that the data set included 70 performance variables. The utilisation of correlation analysis and the elimination of the redundant variables led to the selection of an optimal subset of 27 variables. In the PCA stage, three off-possession and four on-possession skill sets were obtained from these performance variables. Since the presence of too many variables is not desirable in cluster analysis (Beyer et al., 1999), having seven features was suitable for the next stage. Next, the skill set scores were utilised in AHC and helped derive six major clusters and nineteen sub-clusters that categorise technical player functions.

The obtained clustering solution was visualised with a dendrogram, which portrays the hierarchical formation of the clusters, the similarities of the clusters and the number of players in each sub-cluster. The six major clusters and nineteen sub-clusters were introduced on the basis of PC scores of their members. Following that, the distribution of the players belonging to five positional groups in the six clusters created was examined. An interesting finding was that only 46% of all central midfielders belonged to the same cluster, which is the roaming playmaker cluster. On the other end of the spectrum, 96% of all central defenders ended up in the defensive specialist cluster. To investigate the market valuation of players belonging to the six clusters, the median and maximum wages of the players in the sample were examined. According to this analysis, ball-dominant creator forwards such as Lionel Messi of Barcelona and Kevin de Bruyne of Manchester City had the highest median salaries. Finally, the case study of Manchester United was provided to illustrate how the model developed could support the recruitment decision-making of players and building a football squad via benchmarking against best practices in England, Germany and Spain.

When deciding on the number of main clusters and sub-clusters, I utilised the Silhouette coefficient to assess cluster validity. Silhouette coefficient is a robust metric used to measure the performance of a clustering solution by evaluating the compactness and separation of clusters (Bolshakova & Azuaje, 2003; Handl et al., 2005; Rousseeuw & Kaufman, 1990; Vendramin et al., 2010). Here, the result varies from +1 to -1, and a measure between 0.2 and 0.5 is considered a reasonable solution (Sarstedt & Mooi, 2014). The Silhouette coefficient obtained based on six and nineteen groups of technical functions was 0.35 and 0.25, respectively. Hence, a moderate clustering quality was achieved, and both research questions were addressed and positively answered.

5.2. Contributions

5.2.1. Theoretical Contributions

In Chapter 3, 'Building a Decision-Making Framework for Player Recruitment', both the identification of pursued player recruitment objectives and the resulting framework help to advance the theoretical understanding of this critical decision-making context in football. The proposed framework provides a holistic picture of the sporting, organisational and financial aspects on the basis of the available literature and therefore enhances the understanding of the complexity behind player recruitment decisions in European club football.

Existing literature on talent identification in football mainly focuses on the sporting performance of youth players (Johnston et al., 2018; Larkin & O'Connor, 2017; Larkin & Reeves, 2018; Río, 2014; Roberts et al., 2019). On the other hand, there has been a growing sports analytics literature examining the technical performance indicators used to analyse players and teams (Barnes et al., 2014; Barron et al., 2018; Dellal et al., 2011; Fernandez-Navarro et al., 2016; Lago-Peñas et al., 2010; Liu et al., 2016; Rampinini et al., 2009; Yi et al., 2018). However, the financial and organisational context that clubs operate in is not covered in these lines of research. This framework brings together the literature on performance analysis of players and the research conducted on the organisational goals of football clubs in both sporting and financial aspects. For instance, the playing identity of a team is a crucial determinant affecting the criteria required for different playing positions on the pitch. Additionally, many clubs rely on transfer fee profits from the relatively young players that they sign and later sell to wealthier clubs. Therefore, another vital point is whether there is an expectation of net transfer fee income based on the future sales of a player. The consideration of these and other objectives provides a multi-dimensional theoretical grounding for the decision frame regarding first-team player recruitment.

Furthermore, the methodology utilised in this framework can be applied to similar decision-making contexts in sports and football management. In other domains, where recruiting talent has vital financial and strategic implications for organisations, it can help to define a comprehensive decision frame based on the relevant objectives.

This chapter also contributes to the field of operations research, as it demonstrates that VFT can help to structure a decision-context in sports, football and human resources management. Current research conducted with the application of VFT includes the problem domains regarding defence, environment & energy, and other government and corporate settings (Parnell et al., 2013). Within the corporate sectors, the majority of the research focused on e-commerce and project management (Barclay, 2014). Therefore, we contribute to the corpus of VFT literature by showing how it can help to define and analyse additional decision-making contexts.

The findings of Chapter 4, ‘Identifying Means Objectives for Player Recruitment’, contribute to the literature on sports management, performance and analytics by providing a taxonomy of technical player functions based on the skill sets exhibited at the top level of European football. The most common categorisation used for football players is based on their playing positions on the pitch (Bradley et al., 2009; Dellal et al., 2011; Di Salvo et al., 2007). In the analytical model presented, the skill sets of players determine their categories. The derived technical functions help to present a deeper theoretical understanding of player typologies that go beyond the playing positions. Further, it is shown that central midfielders are the most heterogeneous positional group by analysing the overlaps between player groups based on playing positions and technical functions. 46% of the players whose primary position is central midfielder ended up in the roaming playmaker cluster, whereas 54% of these players were spread in the other five clusters. Existing research on physical performance, market valuation and peak ages of football players has utilised playing positions to understand better the differences between players fielded in different positions. This chapter contributes to these research domains by offering a technical function-based categorisation that can be utilised to gain a different perspective on player performance and valuation. The methodology employed in the development of this model can be adapted to other sports when investigating player skills and typologies.

This chapter also demonstrates that the utilisation of PCA and AHC can be beneficial in the context of football and sports. Up to now, PCA has been utilised in a variety of sports-related settings (Bruce, 2016; Fernando et al., 2015; Gløersen et al., 2018; Parmar et al., 2018). However, to my knowledge, AHC has not been used to discover player types in academic research in sports. Since this method provides a hierarchical structure of derived groups and sub-groups as well as their distances, it can allow for a clear conceptual picture regarding a complex problem domain.

5.2.2. Practical Contributions

The proposed decision-making framework in Chapter 3 offers a conceptual basis for key decision-makers responsible for managing and organising the player recruitment activities in European club football. The identification of four fundamental and nine means objectives allows for a structured approach to problem-solving in the analysed decision-making context. The proposed framework can be used for the effective and sustainable management of player costs. As pointed out by Barros (2006), the strategic management of player expenditures sustains the existence of many football clubs. Additionally, Baroncelli and Lago (2016) emphasise the contributions of the net transfer fee income to club finances, which is realised by discovering young talent and selling their contractual rights to wealthier clubs.

Further, a significant source of competitive advantage for clubs is the playing identity (Honigstein, 2015; Michels, 2001; Perarnau, 2016; Wilson, 2018). Despite that some major

European clubs have very distinct football styles, this might not be the case for many clubs all over Europe. Furthermore, even the top clubs might also go through phases where they evolve their approaches to playing football based on the new manager/head coach hires. For instance, former Barcelona head coach Luis Enrique adopted a more direct attacking style, which was different from the well-known possession-based identity of the club (Haugstad, 2017). The section introduced on the construction of a football identity can allow for a shared understanding of the on-the-pitch requirements that an ideal playing squad should fulfil.

Since the framework covers a set of contextual characteristics that may vary from club to club, it can be adapted to different football leagues and club sizes. Further, it can promote a strategic discussion regarding the balance in the squad in terms of values such as age, technical performance, physical performance and the potential to bring net transfer income. For instance, if there are five identified playing position-based needs in a summer transfer window, the relevant objectives for each case could be decided with the rest of the squad in mind. As a result, decision-makers in clubs could benefit from the adaptability of the objectives in squad planning and building.

Another significant practical contribution of this framework is the determination of the time dimension that VFT makes possible. Decision-makers in clubs are under intense pressure from the fan bases to build squads that will immediately be successful. However, organisational goal setting based on immediate results would generally not be viewed as an acceptable practice in other sectors. Hence, the framework proposed can help to set customised return on investment targets for different players. For instance, a pre-peak aged player might be expected to perform very well in the space of six months, whereas a peak aged player could be required to deliver on the pitch immediately. On the other hand, the former can potentially bring a net transfer profit while the latter may not be expected to do so.

Based on the clarification of case-based expectations, decision-makers can participate in succession planning for several years.

A recent trend in EPL is the increased adoption of the sporting director model, which is the traditionally accepted model in German and Spanish football. In this model, the sporting director is also responsible for identifying the best possible experts in live, video and data scouting and managing the expectations of the owner and the executive board. On this power shift, Kelly and Harris (2010) point to a lack of trust between the first team managers/head coaches and the executives and owners. Since this model covers both sporting and financial aspects of a recruitment decision, it can promote improved coordination among multiple stakeholders involved in complex problem-solving. This would also mitigate against the risks of over-dependence of football agents. As stated by Rossi et al. (2016), the relationship between agents and club decision-makers can be so strong that a club might only recruit through one agent acting as a gatekeeper. Limiting the talent scope to the player lists created by a small number of agents can lead a club to miss out on relatively more cost-effective solutions.

The analytical model introduced in Chapter 4 provides a data-driven taxonomy of technical player functions based on the on-possession and off-possession skill sets of football players, whereas the following case study illustrates how the model can be used in decision support at a team level. The club decision-makers involved in the recruitment activities face the challenge of bringing together players that possess complementary skill sets within the limitations of their budgets. Also, the talent pool to choose from is much larger than before as match video and performance data of the players have increased considerably in the last two decades. This model offers an evidence-based conceptualisation that can be utilised for the effective management of player costs and recruitment activities. Many clubs in European football rely on the transfer fee income generated from player sales, so they might have to sell some of their best players. As pointed out by Szymanski (2015), it is not an easy task to keep selling good players and staying competitive in a given league. The recruitment case study provided shows that this model can allow for the discovery of high-performing and cost-effective players.

Since the model is derived from the performance data of the three top revenue-generating European leagues, it reflects the top level of European club football. As pointed out by Soriano (2011), decision-makers in football clubs should be open to the idea of borrowing from the successful practices of other clubs. The scope of available data and video regarding other clubs provides the clubs with an opportunity that is non-existent in many other business sectors. For instance, after Borussia Dortmund won two consecutive championships with a much smaller budget thanks to a high-pressing style, Bayern Munich started adopting a similar sporting approach (Honigstein, 2015). Similarly, clubs can use this model in identifying benchmarks in terms of establishing a certain football identity and building a competent team on the basis of the skill sets and functions exhibited within a playing squad.

The utilisation of playing positions in categorising and analysing players is very common in the football industry. However, top coaches such as Jürgen Klopp and Josep Guardiola point to the dangers of assigning fixed positions to footballers (Perarnau, 2016; Smith, 2017). It might be the case that some players might be able to play well in a number of positions or that they were not fielded in their most productive position for many years. For instance, Josep Guardiola started deploying the German international external defender, Phillip Lahm, as a central midfielder at the age of 30 (Perarnau, 2016). Hence, the proposed model can help to mitigate position-based biases when assessing the capabilities of players. Existing research on talent identification shows that scouts and coaches might suffer from cognitive heuristics based on the physical attributes of players (Kuper & Szymanski, 2010; Mills et al., 2018).

Similarly, position-based biases can prevent decision-makers from accurately assessing a player’s skill sets and potential contributions.

Another common pitfall when developing talent is an excessive focus on goal scoring and assist making. As illustrated in one of the case studies, an impressive goal-scoring performance by a player can result in an intensified competition between several clubs. Facing higher fee and wage costs, the smaller clubs are expected to be priced out in such instances. Therefore, the framework can be beneficial in helping identify player strengths and weaknesses regardless of their scoring contributions.

In the area of player development, one of the main challenges that club decision-makers face is the integration of the youth players into the first teams. This is due to the lack of communication between the coaches in A and youth teams (Relvas et al., 2010). Based on the proposed player categories, youth players can be developed and mentored based on the benchmarks in a given playing squad, as well as the top players in a given country. Since utilising elite competitors as benchmarks can be a beneficial way to improve performance (Lyttleton, 2017), this practice can help coaches to bridge the gap between actual and desired performance.

5.3. Limitations and Future Research

A limitation of this study is that the utilised performance data does not account for contextual information on the locations of the match events. Further, physical performance indicators such as running distance and sprint distance covered are not included in the data set. As is the case with the physical performance data, crucial data on the personality characteristics of players is very hard to gather on a large scale. Therefore, this is another limitation of this study. Additionally, some players are known to be relatively more competent than others in off-the-ball movements and orientating themselves by anticipating what will happen next.

This kind of skill can be captured by tracking data based on how the players utilise the available spaces on the pitch or through the live scouting of players.

Future work can focus on utilising the presented objectives and criteria in Chapter 3 via the application of quantitative methods such as multi-criteria decision making and goal programming. As pointed out by Parnell et al. (2013), management science and operations research techniques have great potential to leverage the benefits of VFT.

Similarly, these methods can be applied to a recruitment decision-making problem as in the case study presented. In this case study, players are selected on the basis of satisfying three skill set scores that were identified through benchmark analysis. Via the application of multi-criteria decision-making methods, players can be assessed and compared with a larger set of variables in mind, including the age, career appearances and contract expiry data of players.

Regarding the analysis of skill sets and functions, combining technical performance with physical performance and accounting for the match context can give a more detailed perspective on player skill sets and functions. Further, the utilisation of match event locations can lead to valuable insights based on the spaces that players utilise on the pitch.

Moreover, players from the derived clusters may be analysed and compared through the lens of effectiveness and production functions, as was done on a club performance basis (Carmichael et al., 2011; Carmichael et al., 2016; Gerrard, 2007). Player wages, transfer fees and Transfermarkt values can be utilised with a view to examining the productivity of different groups of players. Finally, the study is limited to one season and the three top-level European leagues. Therefore, studies applying the research design to other leagues would allow for assessing the generalisability of the findings.

6. Appendix

APPENDIX A

Table 13 Results of the principal component extraction of off-possession variables: component loadings of variables and variance explained by each component. The bold values relate to the variables allocated to principal components.

Variables	PC1 (35%)	PC2 (23%)	PC3 (19%)
Interceptions	**0.55**	0.00	0.11
Ball recoveries	**0.52**	0.08	0.21
Defensive duels won	**0.48**	-0.03	0.29
Field aerial duels won	0.19	**0.60**	-0.16
Loose ball duels won	0.03	**0.56**	0.13
Head shots on target	-0.17	**0.54**	-0.23
Opponent half recoveries	-0.25	0.09	**0.66**
Dangerous opp. half recoveries	-0.28	0.14	**0.57**

Table 14 Results of the principal component extraction of playmaking variables: component loadings of variables and variance explained by each component. The bold values relate to the variables allocated to principal components.

Variables	PC1 (82%)
Successful passes	**0.47**
Successful forward passes	**0.47**
Successful vertical passes	**0.45**
Successful passes to final third	**0.44**
Successful long passes	**0.40**

Table 15 Results of the principal component extraction of creative skill variables: component loadings of variables and variance explained by each component. The bold values relate to the variables allocated to principal components.

Variables	PC1 (58%)	PC2 (12%)
Expected assists	**0.37**	0.41
Successful key passes	**0.35**	0.41
Successful dribbles	**0.35**	0.40
Offensive duels won	**0.34**	0.39
Assists	**0.30**	0.34
Successful through passes	0.27	**0.58**
Successful penetrative passes	0.33	**0.40**
Long dribbles	0.31	**-0.38**
Successful crosses	0.25	**-0.37**

Table 16 Results of the principal component extraction of shot involvement variables: component loadings of variables and variance explained by each component. The bold values relate to the variables allocated to principal components.

Variables	PC1 (85%)
Expected goals	**0.52**
Shots on target	**0.52**
Goals	**0.51**
Successful linkup plays	**0.45**

APPENDIX B

Figure 6.1 Comparison of the representative (in red) and the most expensive players (in blue) of the sub-clusters: No.1 to 4.

Figure 6.2 Comparison of the representative (in red) and the most expensive players (in blue) of the sub-clusters: No.5 to 8.

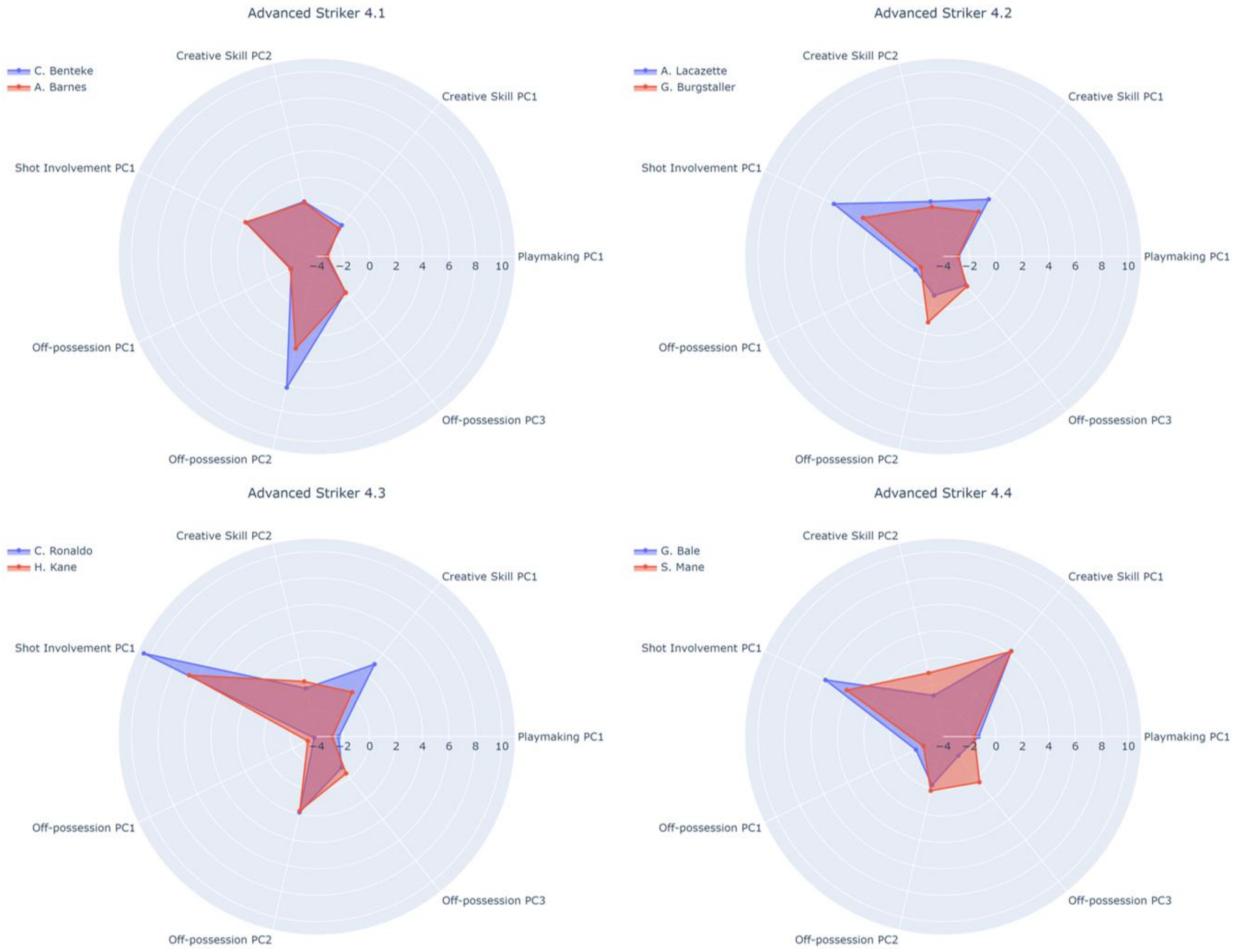

Figure 6.3 Comparison of the representative (in red) and the most expensive players (in blue) of the sub-clusters: No.9 to 12.

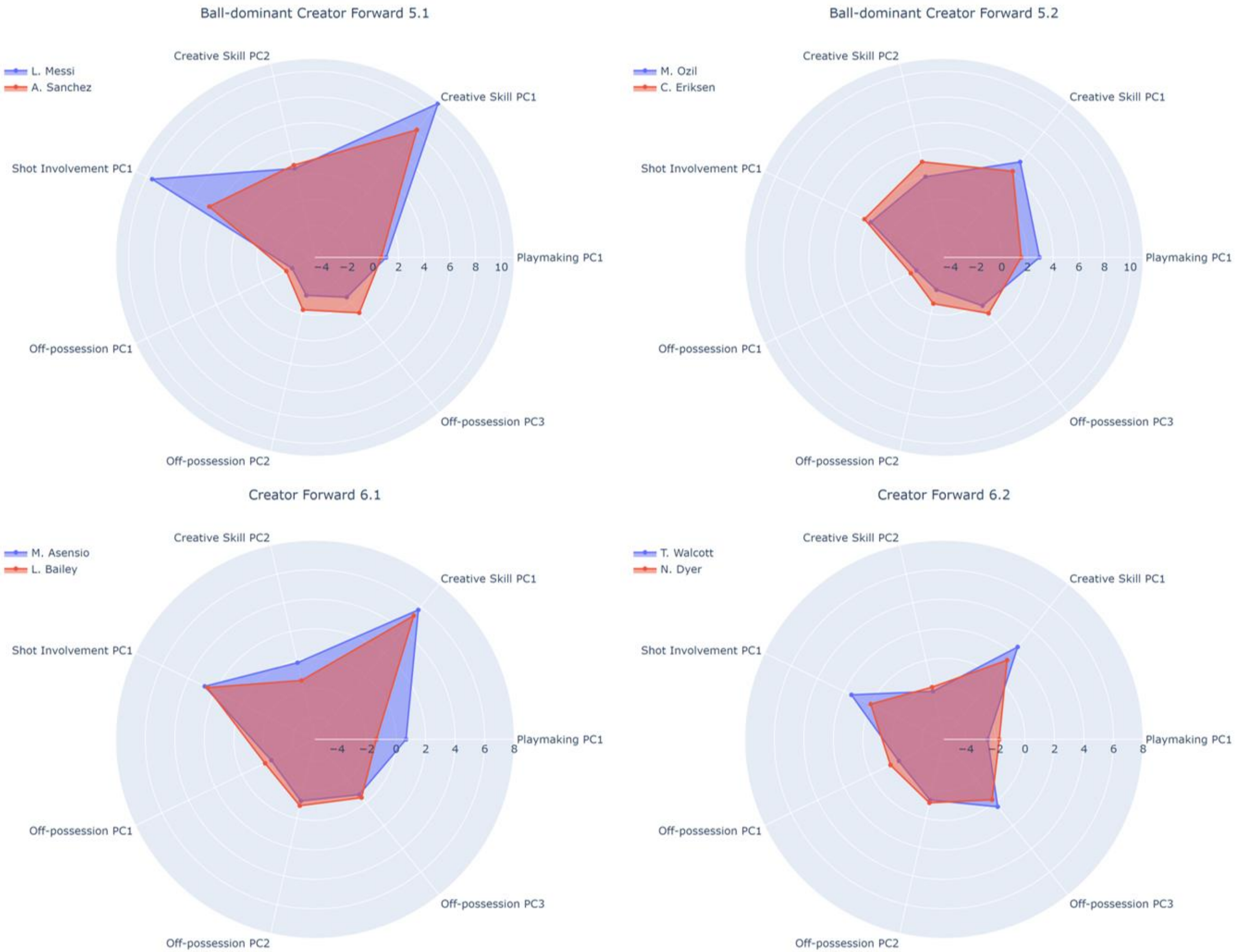

Figure 6.4 Comparison of the representative (in red) and the most expensive players (in blue) of the sub-clusters: No.13 to 16.

Figure 6.5 Comparison of the representative (in red) and the most expensive players (in blue) of the sub-clusters: No.17 to 19.

APPENDIX C

Table 17 The list of starting players of Barcelona, Bayern Munich and Manchester City in 2017/18.

Player	Club	Cluster	Age	Wages	PLAY1	CREA1	CREA2	SHOT1	OFF1	OFF2	OFF3
Central defenders											
G. Pique	Barcelona	1.2	31	12.8	2.61	-2.51	0.25	-0.91	2.27	-0.56	-1.17
N. Otamendi	Man City	1.3	30	10.5	4.59	-2.61	0.2	-0.75	1.81	0.67	0.25
M. Hummels	Bayern Munich	1.3	29	10.0	4.77	-2.2	1.05	-1.06	3.33	1.79	1.91
A. Laporte	Man City	1.3	24	8.8	3.48	-2.18	0.44	-1.54	2.19	0.11	0.76
S. Umtiti	Barcelona	1.3	24	6.5	2.71	-2.82	-0.15	-1.25	2.79	-0.02	-0.14
N. Süle	Bayern Munich	1.3	22	3.0	3.65	-2.58	0.17	-0.81	1.23	1.45	0.99
External defenders											
D. Alaba	Bayern Munich	2.1	26	10.0	1.19	1.37	-1.70	-0.66	-0.38	-1.10	0.32
S. Roberto	Barcelona	2.1	26	10.0	1.60	2.57	-0.11	-1.27	0.69	-1.32	1.40
J. Alba	Barcelona	2.1	29	8.6	1.45	2.25	1.24	-1.06	1.13	-1.30	-0.43
K. Walker	Man City	2.1	28	6.4	2.21	0.98	-0.18	-1.57	0.52	-0.55	0.49
F. Delph	Man City	3.1	28	4.7	2.86	-0.47	-0.53	-1.33	0.72	-0.35	0.22
J. Kimmich	Bayern Munich	5.2	23	1.5	2.49	4.63	1.04	-1.27	-0.10	-0.66	0.78
Central midfielders (No.6/8)											
S. Busquets	Barcelona	1.3	29	15.0	4.08	-1.14	0.85	-1.28	0.37	0.12	1.98
I. Rakitic	Barcelona	3.2	30	13.4	4.51	0.78	1.33	-0.98	-0.27	-0.36	2.09
Fernandinho	Man City	3.2	33	10.5	4.43	0.20	2.05	-0.29	0.68	0.90	2.03
J. Martinez	Bayern Munich	3.1	29	6.0	1.65	-1.04	0.64	-0.96	1.54	1.34	1.02
Central midfielders (No.8/10)											
K. De Bruyne	Man City	5.2	27	13.4	2.31	7.73	3.64	0.66	-1.29	-0.83	1.14
J. Rodriguez	Bayern Munich	5.2	26	11.1	3.68	5.16	2.01	1.49	-1.19	-1.62	0.17
D. Silva	Man City	5.2	32	9.4	2.73	5.06	3.26	1.03	-2.20	-1.09	2.01
External midfielders											
P. Coutinho	Barcelona	5.1	26	23.4	1.00	6.67	1.77	3.30	-1.69	-1.47	1.00
A. Iniesta	Barcelona	3.2	34	16.1	4.07	4.19	2.60	-1.04	-1.98	-1.75	2.42

F. Ribery	Bayern Munich	5.1	35	16.0	0.10	5.06	-1.39	3.43	-2.69	-0.85	1.69
R. Sterling	Man City	4.4	23	8.8	-2.10	4.65	0.01	5.30	-1.88	-0.92	-0.24
A. Robben	Bayern Munich	6.1	34	7.0	-0.66	6.04	-0.82	2.86	-2.07	-1.69	-1.11
L. Sane	Man City	6.5	22	5.6	-2.52	4.74	-0.09	1.78	-1.78	-1.35	-0.18
Central forwards											
L. Messi	Barcelona	5.1	31	40.0	1.04	10.84	2.61	9.56	-2.53	-1.46	-0.53
L. Suarez	Barcelona	4.3	31	23.4	-2.12	2.75	1.66	6.83	-2.17	-0.96	-0.56
R. Lewandowski	Bayern Munich	4.3	29	16.0	-2.90	-0.32	-0.45	9.00	-3.41	1.99	-2.14
T. Müller	Bayern Munich	6.4	28	15.0	-1.60	4.88	1.11	1.94	-1.45	-0.10	-0.33
S. Aguero	Man City	4.3	30	9.4	-2.91	3.73	-0.02	7.35	-3.07	0.61	-1.94

7. References

Aalbers, B., & Van Haaren, J. (2018). *Distinguishing Between Roles of Football Players in Play-by-Play Match Event Data.* Paper presented at the International Workshop on Machine Learning and Data Mining for Sports Analytics.

Alencar, M. H., Priori Jr, L., & Alencar, L. H. (2017). Structuring objectives based on value-focused thinking methodology: Creating alternatives for sustainability in the built environment. *Journal of Cleaner Production, 156*, 62-73. doi:10.1016/j.jclepro.2017.03.221

Ancelotti, C., Brady, C., & Forde, M. (2016). *Quiet Leadership: Winning Hearts, Minds and Matches*: Penguin UK.

Anderson, C., & Sally, D. (2013). *The numbers game: Why everything you know about soccer is wrong*: Penguin.

Association of the Liga Nacional de Futbol Profesional. (2019). Distribution of broadcasting revenues of the 2019/20 season to meet the reporting obligations set out in Royal Decree-Law 5/2015. Retrieved from https://www.laliga.com/en-GB/transparency/economic-management/tv-rights

Atkins, C. (2014). Ranking the Top 10 Players to Have Represented Barcelona and Ajax. Retrieved from https://bleacherreport.com/articles/2237823-ranking-the-top-10-players-to-have-represented-barcelona-and-ajax

Atkinson, C. (2018). Game on: Facebook, Google and Amazon are coming for your TV sports. Retrieved from https://www.nbcnews.com/tech/internet/game-facebook-google-amazon-are-coming-your-tv-sports-n856906

Austin, S. (2017a). Bierhoff: Why Germany treats players like 'independent entrepreneurs'. Retrieved from https://trainingground.guru/articles/bierhoff-why-germany-treats-players-like-independent-entrepreneurs

Austin, S. (2017b). Chelsea loan players to 87 clubs in five seasons. Retrieved from https://trainingground.guru/articles/chelsea-loan-players-to-87-clubs-in-five-seasons

Austin, S. (2019). Tim Waskett: The maths behind Liverpool's title charge. Retrieved from https://trainingground.guru/articles/tim-waskett-the-maths-behind-liverpools-title-charge

Balague, G. (2017). *Brave New World: Inside Pochettino's Spurs*: W&N.

Barbier, J. (2012). Don Leo's Odyssey. Retrieved from https://www.theblizzard.co.uk/article/don-leos-odyssey

Barclay, C. (2014). Overview of the Value-Focused Thinking Methodology. In *Advances in Research Methods for Information Systems Research* (pp. 183-196).

Barnes, C., Archer, D. T., Hogg, B., Bush, M., & Bradley, P. S. (2014). The evolution of physical and technical performance parameters in the English Premier League. *Int J Sports Med, 35*(13), 1095-1100. doi:10.1055/s-0034-1375695

Baroncelli, A., & Lago, U. (2016). Italian Football. *Journal of Sports Economics, 7*(1), 13-28. doi:10.1177/1527002505282863

Barron, D., Ball, G., Robins, M., & Sunderland, C. (2018). Artificial neural networks and player recruitment in professional soccer. *PLoS One, 13*(10). Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/30379858

Barros, C. P. (2006). Portuguese football. *Journal of Sports Economics, 7*(1), 96–104.

Bate, A. (2018). Norwich top of the Championship: and even Jurgen Klopp is a fan. Retrieved from https://www.skysports.com/football/news/11709/11578660/norwich-top-of-the-table-stuart-webber-explains-how-it-happened

Bate, A., & Mackenzie, R. (2016). The Changing Face of Scouting: The challenge of buying from abroad

Retrieved from https://www.skysports.com/football/news/11096/10694789/the-changing-face-of-scouting-the-challenge-of-buying-from-abroad
Bekkers, J., & Dabadghao, S. (2019). Flow Motifs in Soccer: What can passing behavior tell us? *Journal of Sports Analytics, 5*(4), 299-311.
Beyer, K., Goldstein, J., Ramakrishnan, R., & Shaft, U. (1999). *When is "nearest neighbor" meaningful?* Paper presented at the International conference on database theory.
Biermann, C. (2019). Football Hackers. The Science and Art of a Data Revolution. In. London: Blink Publishing
Bird, J. (2022). How many trophies has Manchester United won? Complete list of silverware in the Red Devils' trophy case. Retrieved from https://www.sportingnews.com/uk/soccer/news/manchester-united-trophies-won-complete-list-silverware/uuyb6wc5o7isrfwa1pjssb07
Bleaney, R. (2014). Football Manager computer game to help Premier League clubs buy players. In.
Bolshakova, N., & Azuaje, F. (2003). Cluster validation techniques for genome expression data. *Signal processing, 83*(4), 825–833.
Bradley, P. S., & Ade, J. D. (2018). Are current physical match performance metrics in elite soccer fit for purpose or is the adoption of an integrated approach needed? *International journal of sports physiology and performance, 13*(5), 656-664.
Bradley, P. S., Sheldon, W., Wooster, B., Olsen, P., Boanas, P., & Krustrup, P. (2009). High-intensity running in English FA Premier League soccer matches. *Journal of Sports Sciences, 27*(2), 159–168.
Brechot, M., & Flepp, R. (2020). Dealing with randomness in match outcomes: how to rethink performance evaluation in european club football using expected goals. *Journal of Sports Economics, 21*(4), 335-362.
Brewster, C., & Cerdin, J.-L. (2018). HRM in mission driven organizations. *Cham: Palgrave Macmillan*.
Bridgewater, S. (2016). *Football management*: Springer.
Bruce, S. (2016). A scalable framework for NBA player and team comparisons using player tracking data. *Journal of Sports Analytics, 2*(2), 107-119. doi:10.3233/jsa-160022
Bryson, A., Frick, B., & Simmons, R. (2013). The returns to scarce talent: Footedness and player remuneration in European soccer. *Journal of Sports Economics, 14*(6), 606–628.
Caley, M. (2015). Premier League Projections and New Expected Goals. Retrieved from https://cartilagefreecaptain.sbnation.com/2015/10/19/9295905/premier-league-projections-and-new-expected-goals
Carling, C., Williams, A. M., & Reilly, T. (2007). *Handbook of soccer match analysis: A systematic approach to improving performance*: Routledge.
Carmichael, F., Forrest, D., & Simmons, R. (1999). The labour market in association football: who gets transferred and for how much? *Bulletin of Economic Research, 51*(2), 125–150.
Carmichael, F., McHale, I., & Thomas, D. (2011). Maintaining Market Position: Team Performance, Revenue and Wage Expenditure in the English Premier League. *Bulletin of Economic Research, 63*(4), 464-497. doi:10.1111/j.1467-8586.2009.00340.x
Carmichael, F., Rossi, G., & Thomas, D. (2016). Production, Efficiency, and Corruption in Italian Serie A Football. *Journal of Sports Economics, 18*(1), 34-57. doi:10.1177/1527002514551802
Castellano, J., Casamichana, D., & Lago, C. (2012). The use of match statistics that discriminate between successful and unsuccessful soccer teams. *Journal of human kinetics, 31*, 137–147.

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), 50-90.

Chaudhuri, O. (2019). Our Own Frenkie De Jongs. Retrieved from https://www.21stclub.com/2019/02/07/our-own-frenkie-de-jongs/

Chelsea FC - All Transfers. (2021). Retrieved from https://www.transfermarkt.com/fc-chelsea/alletransfers/verein/631

Cruyff, J. (2016). *My Turn: The Autobiography*: Macmillan.

Decroos, T., & Davis, J. (2020). *Valuing on-the-ball actions in soccer: a critical comparison of XT and VAEP*. Paper presented at the Proceedings of the AAAI-20 Workshop on Artifical Intelligence in Team Sports.

Dellal, A., Chamari, K., Wong, d. P., Ahmaidi, S., Keller, D., Barros, R., . . . 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.

Deloitte Sports Business Group. (2015). Deloitte Football Money League 2015. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Audit/gx-football-money-league-2015.pdf

Deloitte Sports Business Group. (2017). Deloitte Football Money League 2017. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/sports-business-group/deloitte-uk-sport-football-money-league-2017.pdf

Deloitte Sports Business Group. (2018). *Deloitte Annual Review of Football Finance 2018*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/sports-business-group/deloitte-uk-sbg-annual-review-of-football-finance-2018.PDF

Deloitte Sports Business Group. (2020). *Deloitte Annual Review of Football Finance 2020*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/sports-business-group/deloitte-uk-annual-review-of-football-finance-2020.pdf

Deloitte Sports Business Group. (2021). *Deloitte Football Money League 2021*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/sports-business-group/deloitte-uk-deloitte-football-money-league-2021.pdf

Dendir, S. (2016). When do soccer players peak? A note. *Journal of Sports Analytics, 2*(2), 89-105.

Di Salvo, V., Baron, R., Tschan, H., Montero, F. J. C., Bachl, N., & Pigozzi, F. (2007). Performance characteristics according to playing position in elite soccer. *International Journal of Sports Medicine, 28*(03), 222–227.

Dobson, S., & Gerrard, B. (1999). The determination of player transfer fees in English professional soccer. *Journal of Sport Management, 13*(4), 259-279.

Draebye, M. (2018). Strategic management in football organisations. *Routledge Handbook of Football Business and Management*.

Drevin, L., Kruger, H. A., & Steyn, T. (2007). Value-focused assessment of ICT security awareness in an academic environment. *Computers & Security, 26*(1), 36-43.

Elberse, A. (2013). Ferguson’s formula. *Harvard Business Review, 91*(10), 116-125.

Evans, S. (2020). Soccer: Norwich reap reward for astute recruitment in Premier League push. Retrieved from https://jp.reuters.com/article/instant-article/idINKBN1QU1ON

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis: Wiley series in probability and statistics*. Chichester: Wiley.

Feess, E., Frick, B., & Mühlheusser, G. (2004). Legal restrictions on outside trade clauses–theory and evidence from German soccer. *Institut Zukunft der Arbeit/Discussion paper, 1140*.
Felker, J. A. (2012). EU's new member states: an untapped (and underappreciated) source of talent for MNCs? *The International Journal of Human Resource Management, 23*(2), 255–277.
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.
Fernandez-Navarro, J., Fradua, L., Zubillaga, A., & McRobert, A. P. (2018). Influence of contextual variables on styles of play in soccer. *International Journal of Performance Analysis in Sport*, *18*(3), 423-436.
Fernando, T., Wei, X., Fookes, C., Sridharan, S., & Lucey, P. (2015). *Discovering methods of scoring in soccer using tracking data.* Paper presented at the KDD Workshop on Large-Scale Sports Analytics.
Ferri, L., Macchioni, R., Maffei, M., & Zampella, A. (2017). Financial Versus Sports Performance: The Missing Link. *International Journal of Business and Management, 12*(3). doi:10.5539/ijbm.v12n3p36
FIFA. (2006). *The Technical Report and Statistics 2006 World Cup* Retrieved from https://resources.fifa.com/image/upload/germany-2006-part-500834.pdf?cloudid=ddjn8odoi3vcgscxqds3
FIFA. (2010). The TMS explained. *FIFA*.
FIFA. (2018). *The Technical Report 2018 FIFA World Cup Russia*. Retrieved from https://resources.fifa.com/image/upload/2018-fifa-world-cup-russia-technical-study-group-report.pdf?cloudid=evdvpfdkueqrdlbbrrus
FIFA. (2020). Regulations on the Status and Transfer of Players. Retrieved from https://resources.fifa.com/image/upload/regulations-on-the-status-and-transfer-of-players-june-2020.pdf?cloudid=ixztobdwje3tn2bztqcp
FIFA. (2021a). FIFA - MEN'S RANKING. Retrieved from https://www.fifa.com/fifa-world-ranking/ranking-table/men/
FIFA. (2021b). *Global Transfer Market Report 2020*. Retrieved from https://img.fifa.com/image/upload/ijiz9rtpkfnbhxwbqr70.pdf
Flint, S. W., Ratten, Nnamdi Madichie, V., Plumley, D., & Wilson, R. (2016). You're getting sacked in the morning: managerial change in the English Premier League. *Marketing Intelligence & Planning, 34*(2), 223-235. doi:10.1108/mip-09-2014-0189
Football Manager 2018. (2017).
Franck, E. (2014). *Financial Fair Play in European Club Football What is it all about?* Retrieved from https://EconPapers.repec.org/RePEc:zrh:wpaper:328
Franck, E., & Nüesch, S. (2011). The effect of wage dispersion on team outcome and the way team outcome is produced. *Applied Economics, 43*(23), 3037–3049.
Franck, E., & Nüesch, S. (2012). Talent and/or popularity: what does it take to be a superstar? *Economic Inquiry*, *50*(1), 202–216.
Freeburn, L. (2009). European Football's Home-Grown Players Rules and Nationality Discrimination Under the European Community Treaty. *Marq. Sports L. Rev., 20*, 177.
Frick, B. (2007). THE FOOTBALL PLAYERS'LABOR MARKET: EMPIRICAL EVIDENCE FROM THE MAJOR EUROPEAN LEAGUES. *Scottish Journal of Political Economy*, *54*(3), 422–446.
Frick, B. (2011). Performance, salaries, and contract length: Empirical evidence from German soccer. *International Journal of Sport Finance*, *6*(2), 87-118.

Galardini, G. (2019). Mediapro Offered '$1.45 Billion-A-Year' To Make Italian Football Great Again. Retrieved from https://www.forbes.com/sites/giacomogalardini/2019/12/02/mediacom-deal-italian-football-serie-a/?sh=aa085df7c339
Gangoso, E. (2019). From Cruyff to De Jong. Retrieved from https://www.fcbarcelona.com/en/news/981669/from-cruyff-to-de-jong
Garcia-del-Barrio, P., & Pujol, F. (2007). Hidden monopsony rents in winner-take-all markets—sport and economic contribution of Spanish soccer players. *Managerial and Decision Economics*, *28*(1), 57-70. doi:10.1002/mde.1313
Geey, D. (2020). *Done Deal: An Insider's Guide to Football Contracts, Multi-Million Pound Transfers and Premier League Big Business*. London: Bloomsbury Publishing PLC.
Gerrard, B. (2007). Is the Moneyball approach transferable to complex invasion team sports? *International Journal of Sport Finance, 2*(4), 214–230.
Gerrard, B. (2016). Analytics, technology and high performance sport. *Critical issues in global sport management*, 205–218.
Gløersen, Ø., Myklebust, H., Hallén, J., & Federolf, P. (2018). Technique analysis in elite athletes using principal component analysis. *Journal of Sports Sciences, 36*(2), 229–237.
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).
Green, E., Menz, M., Benz, L., Zanuttini-Frank, G., & Bogaty, M. (2016). Clustering NBA Players. Retrieved from https://sports.sites.yale.edu/clustering-nba-players
Gullit, R. (2016). *How To Watch Football*: Viking.
Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*: Elsevier.
Handl, J., Knowles, J., & Kell, D. B. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics, 21*(15), 3201–3212.
Haugstad, T. (2017). How Barcelona lost their identity under Luis Enrique. Retrieved from https://www.fourfourtwo.com/features/how-barcelona-lost-their-identity-under-luis-enrique
Hennig, C. (2015). Clustering Strategy and Method Selection. *Handbook of Cluster Analysis*, 703–731.
Herm, S., Callsen-Bracker, H.-M., & Kreis, H. (2014). When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review, 17*(4), 484–492.
Herzberg, F. (2005). Motivation-hygiene theory. *Organizational behavior one: Essential theories of motivation and leadership, eds JB Miner, ME Sharpe Inc, New York*, 61-74.
Honigstein, R. (2015). *Das Reboot: How German Football Reinvented Itself and Conquered the World*: Random House.
Honigstein, R. (2020). Did Bayern really save Dortmund? Retrieved from https://theathletic.co.uk/1754341/2020/04/22/borussia-bayern-munich-dortmund-debt-frings/
How Borussia went from bust to boom. (2013). Retrieved from https://www.aljazeera.com/sports/2013/5/21/how-borussia-went-from-bust-to-boom

Hughes, M., Hughes, P., Mellahi, K., & Guermat, C. (2010). Short-term versus long-term impact of managers: evidence from the football industry. *British Journal of Management*, *21*(2), 571-589.

Hughes, S., & Pearce, J. (2020). The secrets of how Liverpool scout. Retrieved from https://theathletic.com/2216701/2020/11/30/liverpool-scouting-secrets/

Hytner, D. (2014). Arsenal's 'secret' signing: club buys £2m revolutionary data company. Retrieved from https://www.theguardian.com/football/2014/oct/17/arsenal-place-trust-arsene-wenger-army-statdna-data-analysts

Ilkay Gundogan. (2021). Retrieved from https://www.transfermarkt.com/ilkay-gundogan/profil/spieler/53622

Jackson, J. (2019). Pep Guardiola credits scouting for '80%' of Manchester City's success. Retrieved from https://www.theguardian.com/football/2019/nov/01/pep-guardiola-scouting-manchester-city

Jarosz, O., Kornikov, K., & Söderman, S. (2015). *ECA Club Management Guide*: ECA European Club Association.

Johnston, K., Wattie, N., Schorer, J., & Baker, J. (2018). Talent identification in sport: a systematic review. *Sports Medicine, 48*(1), 97-109.

Jolliffe, I. (2002). *Principal component analysis*: Springer Verlag.

Jones, P. D., James, N., & Mellalieu, S. D. (2004). Possession as a performance indicator in soccer. *International Journal of Performance Analysis in Sport*, *4*(1), 98–102.

Kalén, A., Rey, E., de Rellán-Guerra, A. S., & Lago-Peñas, C. (2019). Are soccer players older now than before? aging trends and market value in the last three decades of the UEFA champions league. *Frontiers in Psychology, 10*, 76.

Kane, A., & Shiri, N. (2016). *Selecting the top-k discriminative features using principal component analysis*. Paper presented at the Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference.

Keeney, R. (1996a). *Value-focused thinking*: Harvard University Press.

Keeney, R. (1996b). Value-focused thinking: Identifying decision opportunities and creating alternatives. *European Journal of Operational Research*, *92*(3), 537-549.

Keeney, R. (1999a). Developing a Foundation for Strategy at Seagate Software. *Interfaces, 29*(6), 4-15. doi:10.1287/inte.29.6.4

Keeney, R. (1999b). The Value of Internet Commerce to the Customer. *Management Science, 45*(4), 533-542.

Keeney, R., & McDaniels, T. L. (1992). Value-Focused Thinking about Strategic Decisions at BC Hydro. *Interfaces, 22*(6), 94-109. doi:10.1287/inte.22.6.94

Keller, R. H. (2014). Is team based tacit knowledge transferable? Players as strategic resources. *Journal of Management Policy & Practice, 15*(4).

Kelly, A. L., Wilson, M. R., Jackson, D. T., Turnnidge, J., & Williams, C. A. (2021). Speed of thought and speed of feet: examining perceptual-cognitive expertise and physical performance in an English football academy. *Journal of Science in Sport and Exercise, 3*(1), 88-97.

Kelly, S. (2008). Understanding the role of the football manager in Britain and Ireland: A Weberian approach. *European Sport Management Quarterly, 8*(4), 399–419.

Kelly, S., & Harris, J. (2010). Managers, directors and trust in professional football. *Sport in Society, 13*(3), 489-502. doi:10.1080/17430431003588150

Kempe, M., Vogelbein, M., Memmert, D., & Nopp, S. (2014). Possession vs. direct play: evaluating tactical behavior in elite soccer. *International Journal of Sports Science, 4*(6A), 35-41.

Kesenne, S. (2007). The peculiar international economics of professional football in Europe. *Scottish Journal of Political Economy, 54*(3), 388-399.

Kharrat, T., McHale, I. G., & Peña, J. L. (2020). Plus–minus player ratings for soccer. *European Journal of Operational Research, 283*(2), 726-736.
Kidd, R. (2019). The World's Biggest Soccer Clubs Find A 'Smarter' Way To Scout Transfer Targets. Retrieved from https://www.forbes.com/sites/robertkidd/2019/04/14/the-worlds-biggest-soccer-clubs-find-a-smarter-way-to-scout-transfer-targets/?sh=14ec92fb2adb
Kuper, S., & Szymanski, S. (2010). *Why England Lose: & other curious football phenomena explained*: HarperCollins UK.
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.
Lago-Peñas, C., Lago-Ballesteros, J. n., Dellal, A., & Gómez, M. (2010). Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league. *Journal of sports science & medicine, 9*(2), 288–293.
Larkin, P., & O'Connor, D. (2017). Talent identification and recruitment in youth soccer: Recruiter's perceptions of the key attributes for player recruitment. *PLoS One, 12*(4), e0175716.
Larkin, P., & Reeves, M. J. (2018). Junior-elite football: time to re-position talent identification? *Soccer & Society, 19*(8), 1183-1192.
Lawrence, I. (2018). *Football Club Management: Insights from the Field*: Routledge.
Leon, O. G. (1999). Value-focused thinking versus alternative-focused thinking: Effects on generation of objectives. *Organizational Behavior and Human Decision Processes, 80*(3), 213-227.
Lewis, M. (2004). *Moneyball: The Art of Winning an Unfair Game*: WW Norton & Company.
Lewis, M. (2016). *The undoing project: A friendship that changed the world*: Penguin UK.
Linoff, G. S., & Berry, M. J. A. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*: John Wiley & Sons.
Liu, H., Gómez, M.-A., Gonçalves, B., & Sampaio, J. (2016). Technical performance and match-to-match variation in elite football teams. *Journal of Sports Sciences, 34*(6), 509–518.
Lucifora, C., & Simmons, R. (2003). Superstar effects in sport: Evidence from Italian soccer. *Journal of Sports Economics, 4*(1), 35–55.
Lyttleton, B. (2017). *Edge: What Business Can Learn from Football*: HarperCollins.
Madsen, D. Ø., Stenheim, T., Hansen, S. B., Zagheri, A., & Grønseth, B. O. (2018). Wage expenditures and sporting success: An analysis of Norwegian and Swedish football 2010–2013. *Cogent Social Sciences, 4*(1), 1457423.
Maguire, K. (2019). *The price of football*: Agenda Publishing.
Manchester City - All Transfers. (2021). Retrieved from https://www.transfermarkt.com/manchester-city/alletransfers/verein/281
Manchester City. (2021). Retrieved from https://www.transfermarkt.com/manchester-city/alletransfers/verein/281
Manninen, K., & Huiskonen, J. (2019). Sustainability Goal Setting with a Value-Focused Thinking Approach. In *Sustainable Business Models* (pp. 89-118).
Manning, C. D., Raghavan, P., & Schutze, H. (2008). *Introduction to Information Retrieval*: Cambridge University Press.
Markham, C. (2020). Liverpool owner John W Henry says Reds' achievements eclipse Boston Red Sox's wins. Retrieved from https://www.independent.co.uk/sport/football/premier-league/liverpool-news-john-w-henry-premier-league-trophy-boston-red-sox-a9652011.html

McHale, I. G. (2018). Sports business analytics. *Routledge Handbook of Football Business and Management*.
McNicholas, J., & Crafton, A. (2021). Arsenal to/from Barcelona and Real Madrid: Why £250m of talent has been traded. Retrieved from https://theathletic.co.uk/2393904/2021/02/24/arsenal-to-from-barcelona-real-madrid-why-250m-of-talent-has-been-traded/
Michels, R. (2001). *Team Building: The Road to Success*: Reedswain Spring City, PA.
Mills, J. P., Ing, C., Markham, T., & Guppy, F. (2018). It is not Black and White: A comparison of skin tone by playing position in the Premier League and English football. *International Journal of Sports Science & Coaching*, *13*(3), 398–404.
Morris, B. (2014). Billion-Dollar Billy Beane. Retrieved from https://fivethirtyeight.com/features/billion-dollar-billy-beane/
Morrow, S. (1999). *The new business of football: Accountability and finance in football*: Springer.
Muller, O., Simons, A., & Weirunann, M. (2017). Beyond crowd judgments: Data-driven estimation of market value in association football. *European Journal of Operational Research*, *263*(2), 611-624. doi:10.1016/j.ejor.2017.05.005
Neiger, D., Churilov, L., & Muehlen, M. z. (2006). Integrating risks in business process models with value focused process engineering.
Neville, G. (2015). Gary Neville: The era of the gaffer is over. Retrieved from https://www.telegraph.co.uk/sport/football/11366206/Gary-Neville-The-era-of-the-gaffer-is-over.html
Nielson, E. G. (2021a). "Only spend the money you have": Borussia Monchengladbach count on financial prudence to navigate through pandemic. Retrieved from https://offthepitch.com/a/only-spend-money-you-have-borussia-monchengladbach-count-financial-prudence-navigate-through
Nielson, E. G. (2021b). Owner of four clubs: "Boards and investors are always chasing something, be it signing a right back, winning a trophy or wanting to get promoted. The number one goal should be balancing the budget". Retrieved from https://offthepitch.com/a/owner-four-clubs-boards-and-investors-are-always-chasing-something-be-it-signing-right-back
Nwokolo, C. (2021). Top 10 Most Lucrative Football Shirt Sponsorship Deals in Europe 2021. Retrieved from https://topsoccerblog.com/10-football-shirt-sponsorship-deals/
O'donoghue, P. (2009). *Research methods for sports performance analysis*: Routledge.
O'Donoghue, P. (2008). Principal components analysis in the selection of key performance indicators in sport. *International Journal of Performance Analysis in Sport*, *8*(3), 145–155.
Osborne, J. W., Costello, A. B., & Kellow, J. T. (2008). Best practices in exploratory factor analysis. *Best practices in quantitative methods*, 86–99.
Paola, M. D., & Scoppa, V. (2011). The Effects of Managerial Turnover. *Journal of Sports Economics*, *13*(2), 152-168. doi:10.1177/1527002511402155
Parekh, J., Jaffer, A., Bhanushali, U., & Shukla, S. (2020). Disintermediation in medical tourism through blockchain technology: an analysis using value-focused thinking approach. *Information Technology & Tourism*. doi:10.1007/s40558-020-00180-4
Parmar, N., James, N., Hearne, G., & Jones, B. (2018). Using principal component analysis to develop performance indicators in professional rugby league. *International Journal of Performance Analysis in Sport*, *18*(6), 938–949.
Parnell, D., Groom, R., Widdop, P., & Ward, S. (2018). The Sporting Director: Exploring Current Practice and Challenges within Elite Football. *Routledge Handbook of Football Business and Management*, 155–170.

Parnell, D., Widdop, P., Groom, R., & Bond, A. (2018). The emergence of the sporting director role in football and the potential of social network theory in future research. *Managing Sport and Leisure, 23*(4-6), 242-254.

Parnell, G. S., Hughes, D. W., Burk, R. C., Driscoll, P. J., Kucik, P. D., Morales, B. L., & Nunn, L. R. (2013). Invited Review-Survey of Value-Focused Thinking: Applications, Research Developments and Areas for Future Research. *Journal of Multi-Criteria Decision Analysis, 20*(1-2), 49-60. doi:10.1002/mcda.1483

Pawar, R. S., Sonje, S. A., & Shukla, S. (2020). Food subsidy distribution system through Blockchain technology: a value focused thinking approach for prototype development. *Information Technology for Development*, 1-29. doi:10.1080/02681102.2020.1841714

Peeters, T., & Szymanski, S. (2014). Financial fair play in European football. *Economic Policy, 29*(78), 343–390.

Peña, J. L., & Navarro, R. S. (2015). Who can replace Xavi? A passing motif analysis of football players. *arXiv preprint arXiv:1506.07768*.

Perarnau, M. (2016). *Pep Guardiola: The Evolution*: Birlinn Ltd.

Plumley, D., Wilson, R., & Ramchandani, G. (2017). Towards a model for measuring holistic performance of professional Football clubs. *Soccer & Society, 18*(1), 16–29.

Polden, J. (2018). Ballon d'Or winners: List of every victor including Ronaldo and Messi. Retrieved from https://www.mirror.co.uk/sport/football/news/ballon-dor-winners-list-every-13612425

Rampinini, E., Impellizzeri, F. M., Castagna, C., Coutts, A. J., & Wisløff, U. (2009). Technical performance during soccer matches of the Italian Serie A league: Effect of fatigue and competitive level. *Journal of Science and Medicine in Sport, 12*(1), 227–233.

Relvas, H., Littlewood, M., Nesti, M., Gilbourne, D., & Richardson, D. (2010). Organizational Structures and Working Practices in Elite European Professional Football Clubs: Understanding the Relationship between Youth and Professional Domains. *European Sport Management Quarterly, 10*(2), 165-187. doi:10.1080/16184740903559891

Río, J. F. (2014). Talent detection and development in soccer: a review. *Journal of sport and health research, 6*(1), 7-18.

Rivoire, X. (2011). *Arsene Wenger: The Biography*: Aurum Press Limited.

Roberts, S. J., McRobert, A. P., Lewis, C. J., & Reeves, M. J. (2019). Establishing consensus of position-specific predictors for elite youth soccer in England. *Science and Medicine in Football, 3*(3), 205-213.

Rossi, G., Semens, A., & Brocard, J. F. (2016). *Sports agents and labour markets: evidence from world football*: Routledge.

Rousseeuw, P. J., & Kaufman, L. (1990). *Finding groups in data*: Wiley Online Library Hoboken.

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

Sarstedt, M., & Mooi, E. (2014). *A concise guide to market research*: Springer.

Schoenfeld, B. (2019). How Data (and Some Breathtaking Soccer) Brought Liverpool to the Cusp of Glory. Retrieved from https://www.nytimes.com/2019/05/22/magazine/soccer-data-liverpool.html

Selart, M., & Johansen, S. T. (2011). Understanding the Role of Value-Focused Thinking in Idea Management. *Creativity and Innovation Management, 20*(3), 196-206. doi:10.1111/j.1467-8691.2011.00602.x

Simmons, R. (1997). Implications of the Bosman ruling for football transfer markets. *Economic Affairs, 17*(3), 13-18.
Singh, K. (2019). Introducing Expected Threat (xT). Retrieved from https://karun.in/blog/expected-threat.html
Skandalis, A., Byrom, J., & Banister, E. (2015). Brand scouting: co-creation of value in the Football Manager community. *ACR North American Advances*, 409-414.
Skillcorner. (2020). Let's Get Physical: Comparing the Big 5 European Football Leagues. Retrieved from https://medium.com/skillcorner/lets-get-physical-comparing-the-big-5-european-football-leagues-bcb04ebb835c
Skillcorner. (2022). Retrieved from https://www.skillcorner.com/
Slemen, D. (2020). What makes an effective Sporting Director? Retrieved from https://trainingground.guru/articles/what-makes-an-effective-sporting-director
Smith, A. (2020). Football manager job security at all-time low, Sky Sports study finds. Retrieved from https://www.skysports.com/football/news/11096/10811856/football-manager-job-security-at-all-time-low-sky-sports-study-finds
Smith, R. (2017). In Praise of Soccer's Jack-of-All-Trades, Now Making a Comeback. Retrieved from https://www.nytimes.com/2017/10/31/sports/soccer/premier-league-soccer.html
Soriano, F. (2011). *Goal: The Ball Doesn't Go in by Chance: Management Ideas from the World of Football*: Palgrave Macmillan.
Southwell, C. (2019). Stuart Webber: Climb of the Canaries. Retrieved from https://trainingground.guru/articles/stuart-webber-climb-of-the-canaries
Sportlogiq. (2022). Retrieved from https://sportlogiq.com/
Stats Perform. (2022). Retrieved from https://www.statsperform.com/
Statsbomb. (2022). Retrieved from https://statsbomb.com/
Sunderland, T. (2018). Mauricio Pochettino Hails 'Brave' Tottenham After Making No Summer Signings. Retrieved from https://bleacherreport.com/articles/2790382-mauricio-pochettino-hails-brave-tottenham-after-making-no-summer-signings
Szymanski, S. (2015). *Money and Soccer: A Soccernomics Guide: Why Chievo Verona, Unterhaching, and Scunthorpe United Will Never Win the Champions League, Why Manchester City, Roma, and Paris St. Germain Can, and Why Real Madrid, Bayern Munich, and Manchester United Cannot Be Stopped*: Nation Books.
Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Boston, MA: Pearson
Taylor, D. (2020). Revealed: How Manchester City really run their recruitment department – 'the smell', the global network and the Odegaard story. Retrieved from https://theathletic.com/1560112/2020/01/29/manchester-city-odegaard-transfers-youth-recruitment/
Tenga, A., Ronglan, L. T., & Bahr, R. (2010). Measuring the effectiveness of offensive match-play in professional soccer. *European Journal of Sport Science, 10*(4), 269–277.
The Football Association. (2020). Plan Agreed On Entry Requirements for Overseas Players in England Post-Brexit in 2021. Retrieved from https://www.thefa.com/news/2020/dec/01/gbe-points-based-system-launched-011220
The Football Association Premier League Limited. (2020). *Premier League Handbook Season 20/21*. Retrieved from https://resources.premierleague.com/premierleague/document/2020/09/11/dc7e76c1-f78d-45a2-be4a-4c6bc33368fa/2020-21-PL-Handbook-110920.pdf

Thomas, L. (2018). Why does Willian use Wyscout? Digital scouting tool explained. Retrieved from https://www.skysports.com/football/news/11668/11277117/why-does-willian-use-wyscout-digital-scouting-tool-explained

Tippett, J. (2017). *The Football Code: The Science of Predicting the Beautiful Game*: Self-publisher.

Tomlinson, S. (2020). The Import-export Business. Retrieved from https://www.21stclub.com/2020/01/09/the-import-export-business/

Transfermarkt. (2022). Retrieved from http://www.transfermarkt.com/

Tremlett, G. (2017). Manchester City’s plan for global domination. Retrieved from https://www.theguardian.com/news/2017/dec/15/manchester-city-football-group-ferran-soriano?CMP=share_btn_tw

Tufféry, S. (2011). *Data mining and statistics for decision making*: John Wiley & Sons.

UEFA. (2020a). *The European Club Footballing Landscape: Club Licensing Benchmarking Report Financial Year 2018*. Retrieved from http://www.uefa.com/MultimediaFiles/Download/OfficialDocument/uefaorg/Clublicensing/02/63/79/75/2637975_DOWNLOAD.pdf

UEFA. (2020b). UEFA Champions League Statistics Handbook 2020/21. Retrieved from https://www.uefa.com/MultimediaFiles/Download/EuroExperience/competitions/Publications/02/28/56/89/2285689_DOWNLOAD.pdf

Velte, C. J., Scheller, K., & Steinhilper, R. (2018). Circular economy through objectives–development of a proceeding to understand and shape a circular economy using value-focused thinking. *Procedia CIRP*, *69*, 775-780.

Vendramin, L., Campello, R. J. G. B., & Hruschka, E. R. (2010). Relative clustering validity criteria: A comparative overview. *Statistical Analysis and Data Mining: the ASA Data Science Journal*, *3*(4), 209–235.

Vogelbein, M., Nopp, S., & Hökelmann, A. (2014). Defensive transition in soccer–are prompt possession regains a measure of success? A quantitative analysis of German Fußball-Bundesliga 2010/2011. *Journal of Sports Sciences*, *32*(11), 1076-1083.

Weatherspoon, C. (2021). Analysis: “The transfer market could become like when houses fall in value and no one wants to crystallise the loss”. Retrieved from https://offthepitch.com/a/analysis-transfer-market-could-become-when-houses-fall-value-and-no-one-wants-crystallise-loss

Weimar, D., & Wicker, P. (2017). Moneyball revisited: Effort and team performance in professional soccer. *Journal of Sports Economics*, *18*(2), 140–161.

Wilson, J. (2013). The Question: what does the changing role of holding midfielders tell us? Retrieved from https://www.theguardian.com/football/blog/2013/dec/18/question-holding-midfielders-changing-role

Wilson, J. (2016). The Question: can De Bruyne and Silva prosper in their ‘free No8’ roles? Retrieved from https://www.theguardian.com/football/blog/2016/aug/18/the-question-kevin-de-bruyne-silva-free-roles

Wilson, J. (2018). *Inverting the Pyramid: the History of Soccer Tactics*: W&N.

Wright, N. (2017). The evolution of full-backs: How they became integral to success. Retrieved from https://www.skysports.com/football/news/11095/10752408/the-evolution-of-full-backs-how-they-became-integral-to-success

Wyscout. (2022a). Retrieved from https://wyscout.com/

Wyscout. (2022b). Events Manual. Retrieved from https://footballdata.wyscout.com/events-manual/

Yam, D. (2019). *A data driven goalkeeper evaluation framework.* Paper presented at the 13th MIT Sloan Sports Analytics Conference.

Yi, Q., Gómez, M. A., Wang, L., Huang, G., Zhang, H., & Liu, H. (2019). Technical and physical match performance of teams in the 2018 FIFA World Cup: Effects of two different playing styles. *Journal of Sports Sciences, 37*(22), 2569-2577.

Yi, Q., Jia, H., Liu, H., & Gómez, M. Á. (2018). Technical demands of different playing positions in the UEFA Champions League. *International Journal of Performance Analysis in Sport*, *18*(6), 926–937.

Zambom-Ferraresi, F., Rios, V., & Lera-López, F. (2018). Determinants of sport performance in European football: What can we learn from the data? *Decision Support Systems, 114*, 18-28.