A COMPARISON OF RUNNING INTENSITY DURING PROFESSIONAL FOOTBALL TRAINING AND MATCHES AND THE IMPACT OF RUNNING INTENSITY ON THE EXPECTED GOALS PERFORMANCE OUTCOME

Joshua C. Jardine

A thesis submitted in fulfilment of the requirements for the degree of MSc (Hons) Sport and Exercise Science

School of Sport and Exercise Sciences

University of Kent

September 2023

Pages: 64

Abstract

The two main aims of this investigation were to identify whether running intensity in training compares to the running intensity in matches for a professional football team and whether running intensities impacted the expected goals (xG) outcome of the team during a professional football match. The participants, 27 professional football players at Gillingham Football Club, were recorded during training and matches using GPS during the 2022/23 season. Training distance per minute and high-speed distance (> 18km/h) from medium- and large-sided games were collected, these metrics were also collected during competitive matches alongside xG. Following data collection in-depth statistical analysis was concluded, distance per minute covered presented no significant difference between training or matches when considering positions (P = 0.57). Match distance per minute was significantly higher than training when all playing positions were aggregated (P = 0.03). Defenders covered significantly less distance per minute than midfielders (-8.43m.min^{-1}, P = <0.01), and forwards (-7.33m.min^{-1}, P = 0.02). High-speed distance was significantly lower in training than matches by position (P = <0.01), and significantly lower when positions were aggregated (P = <0.01), and when positions were compared against each other midfielders cover significantly more than defenders (P = 0.04) and forwards (0.01). Mean distance per minute aggregated across position showed no significant correlation with xG scored (P = 0.97), this was the same with high-speed distance (P = 0.48). Forwards, midfielders, and defenders showed no significant correlation for distance per minute and xG

(P = 0.59, P = 0.51, P = 0.88, respectively). Forwards and defenders high-speed distance had no significant correlation with xG (P = 0.35, P = 0.41, respectively). However, for midfielders high-speed distance there was a significant correlation with xG (P =<0.01), although an R^2 value of 0.34 suggests a weak correlation. In summary, training distance per minute produces similar values to match data when considering position, implying it is an effective training tool to prepare for the demands of a match.

There was no impact of running intensity on xG performance of the team, although there was a weak positive correlation identified between high-speed distance of midfielders and xG performance.

Acknowledgements

I would like to express my appreciation to my supervisors, Professor John Dickinson, and Professor James Hopker, who provided me support throughout my research project. I would also like to extend my gratitude to both the backroom and playing staff at Gillingham Football Club, who allowed me to collect the data used through the investigation. In particular, James Russell, Thomas Murphy and Huxley Barnacle-Best who aided in the collection of data.

Table of contents

i. Title page

ii. Abstract

iv. Acknowledgements

v. Table of contents

vi. List of Tables and List of Figures

vii. Definitions

1. Chapter I - Introduction

4. Chapter II – Review of Literature

19. Chapter III – Methods

27. Chapter IV – Results

36. Chapter V - Discussion

48. Chapter VI - Conclusion

List of Tables and List of Figures

Figure 1. Distance Per Minute of each positions' relationship between Match and Training – Page 28

Figure 2. High-Speed Distance of each positions' relationship between Match and Training – Page 29

Figure 3. Distance Per Minute of All Positions relationship with xG – Page 30

Figure 4. Distance Per Minute of Forwards relationship with xG – Page 31

Figure 5. Distance Per Minute of Midfielders relationship with xG – Page 31

Figure 6. Distance Per Minute of Defenders relationship with xG – Page 32

Figure 7. High-Speed Distance of All Positions relationship with xG – Page 33

Figure 8. High-Speed Distance of Defenders relationship with xG – Page 33

Figure 9. High-Speed Distance of Forwards relationship with xG – Page 34

Figure 10. High-Speed Distance of Midfielders relationship with xG – Page 35

Definitions

GPS- Global Positioning System

LSG – Large-Sided Games

MSG – Medium-Sided Games

SSG – Small-Sided Games

xG – Expected Goals

Chapter I:

Introduction

Football over the years has been labelled ‘the worlds game’ due to its ever-growing popularity all over the world, from grassroots level up to the professional game (Pifer et al, 2018). Football has incorporated science since the 1970’s (Drust and Green, 2013) but science has substantially advanced since then with professional teams accepting the application of sports science more as years progress (Reilly et al, 2009), whether that be day-to-day in training sessions or for feedback from matches. Sport science is able to provide knowledge on the practice and performance of athletes (Brink et al, 2018) with the aim to improve the quality of both coaching and performance (Martindale and Nash, 2013). Therefore, placing importance on additional research into football using sport science to identify whether its use can benefit the professional football game further. In addition to providing more insight into factors that may affect performance of football athletes.

It is generally accepted in society that training can improve an individual or a team’s performance whether that be as an employee working in a business setting (Elnaga and Imran, 2013), or whether that be in sporting performance, developing physical attributes and technique (Elliott and Mester, 1988). There are many different types of training that can be undertaken by participants in sport, for example, individuals who compete in cross country would aim to train their aerobic fitness, whereas shot putters

may focus more on strength training. Sirotic and Coutts (2007), characterise intermittent team sport as prolonged, high-intensity, intermittent running, as football is classified as an intermittent team sport, this description would imply that different training types, such as aerobic, strength and skills training, are required to improve the performance of an intermittent team sport participant, which in turn allows for more successful outcomes. In the context of this study, comparing the physical training outputs and the physical outputs from a match to identify if current training methods provide the best preparation for competing, will provide further understanding on the current information surrounding training in football, which in turn, allows coaches to provide suitable training programmes to best replicate the demands of football matches during training.

In addition to that, individuals and teams are constantly trying to identify new ways to win, as sport at the elite level is highly concerned with achieving a winning outcome (Bate, 1988) Bate (1988) goes onto state that many strategies are undertaken in the aim to produce as many positive results as possible, specifically highlighting the use of possession-based football. However, Bate's (1988) work dissects different ways elite level football teams can compete to achieve winning results. This study intends to identify whether changes to physical metrics impacts the chance of scoring more goals in professional football, which could improve the possibility of a positive result in football, the findings of this study could therefore allow players and coaches to identify key areas that could improve the possibility of a positive

impact on footballing performance and highlight areas that do not seem to provide benefit to performance.

To summarise, throughout this study there is an aim to answer two research questions, the initial research question focusses on whether training allows for the reproduction of physical outputs that competing requires you to perform successfully. The second research question is aimed to highlight whether these physical outputs can impact the outcome of competition, specifically football in this present study.

Chapter II:

Review of Literature

This literature review will discuss numerous variables that influence football performance at the entry level, through to the elite professional level, exploring what contributes to success in both football training and matches. An in-depth evaluation of training in football, in particular in professional football, is conducted to investigate why training is necessary, and what previous research has identified as typical practices and values for training at the elite level. Links between the previous research findings of such training and subsequent performance in professional level competitive matches are then outlined and explored. Moreover, the use of technology is explored throughout, considering the validity and reliability of its application in sports such as football. Examples of such uses are collecting physical metrics such as total distance covered, distance per minute covered, and high-speed distance. This review also considers the accuracy of the data produced by the Global Positioning System's (GPS) use in sport. GPS's work by using signals for satellites orbiting the earth, the receiver can then calculate the exact position of the GPS unit and how fast the device is moving, which allows sports scientist and coaches to quantify the external load completed by sports athletes (Scott et al, 2016). This literature review is concluded by highlighting existing research comparable to the present research questions and presenting research questions aimed to answer gaps left within previous research.

Use of technology in sport:

Metrics such as total distance or sprint distance have been used to measure performance in elite football for many years leading to one dimensional analysis (Bradley and Ade, 2018). Bradley and Ade (2018) outline a different way of measuring such metrics in football, such as using a multi-dimensional integrated model to quantify football movements, which contextualises match demands. Furthermore, the use of more advanced ways of measuring metrics, such as this integrated model, can be used to aid performance in intermittent activity by providing data with more strategic purpose.

In team sport, Global Positioning System technology is often used to track and display individuals' movement during training and matches (Coutts and Duffield, 2010). Coutts and Duffield compared the GPS tracking results of two individuals completing a 128.5m circuit replicating team sport demands and found a significance level of < 0.05. They interpreted this as demonstrating acceptable validity of GPS tracking in total distance, intermittent exercise as data was only 5 metres different from actual distance, which can be explained by the route that the individuals ran rather than a GPS error. However, this Coutts and Duffield (2010) study uses team-sport like movements which can be very different between different sports, for example the differences in movement of basketball players to footballers is vast due to basketball being across a smaller playing area, for example. Barr et al. (2019) further supported their results when measuring electronically timed 36.6m sprints among American football players in

training. In this study, the GPS units' measurements of high-velocity sprinting were compared against the number of tackles and blocks measured using video footage. This comparison found that linear high-velocity sprinting values were only slightly different, for instance, only 0.09m.s^{-1} difference in maximal sprint speed, which supports the Global Positioning System validity in assessing the activity profiles of American football players. Although the findings of this study suggest that the GPS units were valid, this study focussed on American footballers which is a different sport with differing demands to the football within this study, for example the stop and start nature of American football compared to the more continuous demands of football. Also, Scott et al. (2016) found that 10Hz GPS units provide reliable comparison between high intensity running actions. These studies support the reliability, validity, and effectiveness of GPS devices to provide accurate values of movement in footballers, when comparing directly with the actual value of the measure parameters such as timing gates. Nonetheless, there are studies offering contradictory findings, such as Gray et al. (2010) who found that the reliability of the GPS measurements significantly decreased with movement intensity, $P < 0.05$, though this study did not focus on football or team sport in general but linear movements, and Coutts and Duffield (2010) who found that 1Hz GPS units were unreliable for higher intensity activities nonetheless higher Hz GPS provided a higher validity as mentioned earlier. Johnston et al (2014), conducted a study into the validity and interunit reliability of Catapult 10Hz GPS units, which were used in this present study, in comparison to a 15Hz unit of another brands (GPSports), their study used team sport athlete movement using metrics such as total distance covered,

high-speed running, average peak speed, along with other metrics. Johnston et al (2014) study identified that the Catapult 10Hz units provided a moderate to good inter unit reliability, with metrics showing less error in every movement demand metric when compared to the 15Hz units, as well as identifying the measurements of movement metrics to be valid. This suggests that the use of Catapult GPS in team sports is recommended for the collection of external load metrics.

Requirements of Football and Professional Football

Football is an intermittent team sport, which requires the use of both aerobic and anaerobic energy systems for the activity to be successfully completed, together with sport-specific skills performed across a one-to-two-hour duration, with both scheduled and unscheduled pauses in play (Baker et al, 2015). Specifically, football requires the aerobic energy system, the process of combustion of carbohydrates and fats in the presence of oxygen (Gastin, 2001) and anaerobic energy system, the process of splitting stored phosphagens, adenosine triphosphate and phosphocreatine and the breakdown of carbohydrates to lactic acid using glycolysis (Gastin, 2001). It requires these to be used over a 90-minute period with a scheduled 15-minute half time stoppage, along with unscheduled stoppages due to substitutions, injuries and the ball going out of play. The aerobic energy system is required for more than 90% of total energy consumption in football matches (Bangsbo, 1994), suggesting that the majority of elite competitive

football is using the aerobic system, with less than 10% using the anaerobic energy system.

Monitoring workload in sport provides many benefits, primarily it is used to reduce chance of injury in athletes when participating in competition, especially over a longer period of time, for example a full footballing season creating an acute: chronic load system (Gabbett, 2017). Another study highlights a benefit of monitoring workload to best prepare for the competition performance as well as monitoring physical adaptations (Pyne and Martin, 2011; cited in Skarbalius et al, 2019). Additionally, workload monitoring can be positively used to identify any changes to fitness and performance outputs, Lima-Alves et al (2022) used both internal and external loads to help monitor changes in team sport athletes, in this case using total distance and heart rate. Load in sport can be measured both internally and externally, internal loads are measured using physiological and psychophysiological markers such as heart rate, heart rate variability and perceived exertion ratings. External loads, however, are measured through physical movements that an individual completes, for example: high-speed running distance, distance per minute and total distance. (Soboleski, 2020).

When participating in football it requires a changeability between both high and low intensity movements constantly throughout the activity. Bangsbo (2014) states that the typical distance covered within a top-level football match was between 10km and 13km, combining bouts of low intensity activity such as jogging and walking along with high intensity bouts of

running and sprinting, supported by the research of Stølen et al. (2005) who similarly measured an average of 10km during elite level 90-minute football matches. Bangsbo's (2014) research further identified that in the Danish top division, teams toward the top half of the league table performed a higher percentage of high intensity, high-speed runs compared to teams finishing toward the middle and bottom of the table, around 30-40% more. (Ingebrigtsen et al, 2012; cited in Bangsbo 2014:1) Bangsbo (2014) also highlights that these results can differ due to the playstyle of different leagues, for instance the high incidences of high intensity running (>19km/h) in Championship and League One football compared to Premier League football in England, with the Championship being 803m, League One 881m and the Premier League 681m. High-speed distance default value is typically >19.8km/h, however Abt and Lovell (2009) found that the median high speed in professional football players was >15km/h which is less than the default value ($P < 0.01$). However, this study was conducted 15 years ago so athlete performance could have changed, along with the opinion on high-speed distance of sport scientists.

Research has also identified differences in the physiological demands placed on elite footballers playing in different positions, among players in the English Premier League, Ju et al. (2022), for example, found that central offensive players covered more high-intensity distances compared to other positions. Ju et al (2022) went onto identify the intensities required in each position using these specific positional intensities, which can be used as a comparison level in training, for example, to identify if specific players who

play specific positions are meeting the intensity they require during a competitive match. Additionally, Abbott et al. (2018) conducted research in football academies and observed that central midfielders produced the highest total distances (11,570 metres ± 479), significantly higher than both central defenders (9,830 ± 428) and forwards 1(0,320 ± 420). Overall, in this paper, central defenders produced the lowest values in all metrics, $P < 0.01$, whilst wide defenders and attackers produced the highest high-intensity values. For instance, central defenders had significantly lower very high-speed running distances in comparison to all other positions, $P < 0.001$, and wide defenders and attackers producing significantly higher high intensity acceleration distance, $P < 0.001$. However, Abbott et al. (2018) did also highlight that there were no significant differences during match outcomes in the measured metrics between different positions. Whilst Chmura et al. (2018) found that when a match outcome was a win, the total distance covered by midfield and forward positions were higher than when losing, and observed with winning outcomes a greater high intensity distance was covered, speed of over 21km/h. Additionally, Rhodes et al. (2021) research on the effect of high intensity accelerations and decelerations on match outcome in professional football, found that there was an impact created by the increase of intensity actions, where outputs were highest when the team had won a match, indicating a correlation between physical outputs and performance outcomes. Moreover, Clemente et al. (2013) conducted a study on the activity profiles of football players during the 2010 World Cup, finding that distance per minute was significantly different between positions with midfielders covering the greatest distance per minute compared to defenders

($P < 0.001$) and forwards ($P < 0.001$), and stated that teams who progressed the furthest covered an increased distance, again showing a positive correlation between match outcome and physical metrics. Successful football match outcomes can be determined in many ways, including efficiency, measured using number of goals scored divided by number of shots taken, shots on goal, ball possession, and pass accuracy (Lepschy et al, 2018).

Another form of measuring successful performance in a football match is using expected goals (xG), a calculation of a team's chances of either scoring or conceding a goal (Rathke, 2017). Rathke (2017) outlines that xG is a calculation of many variables, such as shot location, distance, speed, and angle, that together provide a value for the likelihood for that shot to result in a goal. These factors leading to the value of xG are known as technical performance indicators, this framed against physiological factors such as high-speed running has shown to provide varying results within professional football, for example, a professional Defender and Midfielder from the Spanish League have a stronger impact on their teams expected goals value when performing more high intensity running distance. (Llana et al, 2022). Furthermore, Llana et al (2022) research shows that depending on positional and tactical roles, such as specific directions required of players, various impacts in expected goals can be identified. This could therefore imply that research into this area could be team or tactic specific. Furthermore, Adams et al (2023) research surrounding strategy, a tactical indicator, and player performance, a technical indicator, identify impact on xG, therefore highlighting how adapting tactical and technical instructions

can impact xG, however this study researched handball which tactically is different to football. Overall, the research around the relationship between xG and physical metrics in football, such as distance per minute and high-speed distance, is limited, so identifying if xG is affected by these physical outputs could provide further insight to any importance it could provide for match performance.

In sum, these studies demonstrate the differences in physical metrics, such as high-speed distance and distance per minute, and differing requirements between positions during competition level football, yet present contradictory results regarding the impact of this on performance outcomes, such as winning or losing matches. Moreover, these studies did not compare distance per minute between positions with performance outcome, which could further differentiate running intensities of each playing position and the effect of running intensities on the outcome of performance.

Professional football and elite sport, in general, can be affected by the smallest of margins due to the high level of competition and many variables can contribute to the success or failure of a football teams' performance, including the physical condition of the players, technical and tactical factors (Rosch et al, 2000), such as ball possession, player performance and player role (Liu et al, 2021), that can all be measured when assessing performance. This statement by Rosch et al. (2000) highlights the necessity to further understand the effect of physical measures, such as running distance per minute and high-speed distance running, on football training

and performance. Furthermore, research such as that of Reilly and Gilbourne (2003), has provided evidence of a higher work rate in the modern game compared to 2003, when this research was published, and decades prior to 2003. This would suggest that a team's distance per minute, positional distance per minute and fluctuations in distances per minute, could affect the scoring chances created by either team from match to match in contemporary football. Reilly and Gilbourne (2003) go on to suggest that game-specific training regimes are required more throughout a training week to enhance ability to cope with match demands, and Gimenez et al. (2020) suggests an increase of pitch size and team size to replicate competitive match physical outputs. The research Rosch et al. (2000) and Reilly and Gilbourne (2003), together with Gimenez et al. (2020) referenced earlier, emphasises that physical measures can directly impact performance, such as the ability to cope with match demands, whether that be in matches or the prior training.

Physical Metrics of Football Training

Undertaking training at the elite level of sport is crucial to elite performance, as it aids the ability to develop the automation of motor skills, enhance structural and metabolic functions, as well as facilitating increased self-confidence, and building tolerance for higher training quality and improved competition (Smith, 2003). This in turn, enhances skills an overall improvement of elite level performance. These statements could imply that increasing the frequency and duration of training would create the most

substantial improvement in performance, however this is not the case. Smith (2003) goes on to explain that it is critical that overtraining is avoided, as this can lead to excessive fatigue and/or injury, reverting any progress that has been made by the individual in their training, as for an athlete to achieve optimal performance their optimum fitness is essential. This is echoed by Nikolaidis (2014), who compared physical fitness measures, including the mean power in the Wingate anaerobic test and maximal force in a force-velocity test, to overtraining questionnaires and found that physical fitness measured worsened with a higher overtraining questionnaire score, which demonstrated the negative effect of overtraining on the physical fitness, reducing the overall performance in semi-professional footballers. Further support for avoiding overtraining comes from Urhausen and Kindermann (2002), who explain that when looking at overtrained athletes, both anaerobic performance and time taken to reach exhaustion during high intensity endurance reduced, which was also would suggest a reduced overall footballing performance. A decrease in anaerobic measures of performance, such as high-speed distance, could be used to identify whether overtraining has occurred, as if this were to occur during matches this could potentially impair match outcomes.

Training in football requires structure and differing physical stimulus' such as warm-ups, technical drills, and tactical drills, to best physically prepare players to compete in matches across a season (Briggs, 2013). These training sessions often include aerobic training, anaerobic training, and skill-specific training to allow for all components of match play to be practiced and

improved. Dellal et al. (2012) explains how previous research has indicated the need for training to mimic competitive demands, in order to optimise a players match performance. In professional football, the use of different sized pitches during different drills allows for different physiological and tactical parameters to be met. Owen et al. (2014) found a significantly faster playing tempo during small-sided games, in comparison to medium- and large-sided games (150.5, 108.3 and 120.4m.min^{-1}, respectively. $P < 0.01$), but significantly less high intensity efforts, such as accelerations and decelerations, high intensity running (SSG 7m vs LSG 39m, $P < 0.01$) and high-speed sprint distance (SSG 0m vs LSG 11m, $P < 0.01$) when comparing small-sided games and large-sided games. Dellal et al. (2012) identified two styles of training games; small-sided games (SSG) and large-sided games (LSG). In their study, the SSG included 4-minute 5 versus 5 matches, and the LSG were 11 versus 11 matches. Contradictory to other research, Dellal et al (2012) also detailed that typically, SSG resulted in players having higher measurements of metrics such as increased distance per minute and high-speed distance compared to LSG, which conflicts previous research comparing SSG and LSG. These studies together highlight the differences that previous research has found regarding training preparation for matches, with one study finding medium- and large-sided games produce significantly higher high intensity outputs, such as high intensity running, and another finding that SSG produce higher high intensity outputs compared to LSG. From what we know from the research addressed earlier in this review of literature, it suggests that SSG, MSG and LSG can all produce relevant match-like physical metrics. Nonetheless, there is not a vast amount of

research on whether match data and training data for running intensities, such as distance per minute and high-speed distance directly compare, to be able to confidently conclude whether the output of intensity in training correlates with players performance in matches.

Ammann and Altman (2023) studied player and positional training, and match loads, in professional football, and found that the loads that occurred in matches were unique in comparison to those in training. This stance was supported by research from Anderson et al. (2016), who found that with high-speed running set at >19.8km per hour, an approximately of 100 metres of high-speed distance was covered during training compared to 900 metres during matches. Moreover, Oliva-Lozano et al (2022) research offers further support, detecting that matchday workloads were significantly more demanding for all intensity variables, including high-speed distance running, than training workloads ($P < 0.01$). Gualtieri et al (2023) similarly compared match demands to training strategies and identified that training drills greater than 225 metres squared were adequate in achieving high-speed running distance, this study was conducted by reviewing multiple research papers discussing training strategies and match demands, which resulted in the conclusion that training drills over the 225 squared metre size is advisable to elicit relevant high speed running measures. Critically, distance per minute as a running intensity metric does not seem to have been widely researched when identifying the differences between training and matches in football.

Overall, the research explored here identifies a difference between high-speed distance in matches compared to in training. Moreover, it suggests that actions of high intensity in matches are unique compared to the values produced in training.

Summary of Review of Literature

To conclude, the previous research provides evidence to suggest that training is important in the preparation for competitive football matches. However, research has also provided contradictory results when identifying the most appropriate method to produce match-like physical metrics. Furthermore, there is limited research on the impact of running intensity on performance outcomes, specifically the effect of distance per minute and high-speed distance. Therefore, highlighting a research gap surrounding whether specific training drills, MSG and LSG produce physical metrics required in matches and the impact that running intensity has on match performance. Addressing these gaps could be useful for football teams in the professional setting by allowing teams to identify if their current training methods produce adequate running intensity metrics and if they can improve performance through intensity metrics such as distance per minute and high-speed distance. The present study aims to identify additional evidence to understand further, the impact of intensity in professional football, both in training and in competitive matches.

Research questions and hypotheses'

Subsequent to the literature review, the first research question for this study was: Does training running intensities compare to match running intensities in professional football?

At the beginning of this investigation the null hypothesis was: There will be no significant difference between the intensity of medium- and large-sided games and competitive matches.

The alternative hypothesis was: There will be a significant difference between the intensity of the medium- and large-sided games and competitive matches.

The second research question considered by this study following the review of literature was: Will running intensity during a competitive football match impact the performance outcome of expected goals?

At the beginning of this investigation the null hypothesis was: There will be no significant difference between the running intensity during a competitive football match and the performance outcome of expected goals.

The alternative hypothesis was: There will be a significant positive difference between the running intensity during a competitive football match and the performance outcome of expected goals.

Chapter III

Methods:

Participant information

The data of 27 professional footballers from Gillingham Football Club was collected from all training sessions and competitive matches, focussing on the matches played in the English Football League, League Two 2022/23 season and the 2022/23 F.A. Cup competition between the months of January 2023 and May 2023. The 27 football players chosen were the number of professional contracted footballers at the football club participating in the medium- and large-sided training drills and competitive matches between the 1st of January and the 8th of May. The 27 players selected for this study include 7 Forwards, 11 Midfielders and 9 Defenders ranging between the ages of 18 and 37 years.

These participants had a mean age of 25.9 ± 4.17 years, mean height of 182cm ± 7.35cm and mean body mass of 79.8kg ± 7.2kg. Each participants height and weight were measured at the beginning of the study or immediately when joining the study using a stadiometer and a scale. The criterion to be included in this study required the participants to be contracted to Gillingham Football Club during the 2022/23 season, to have participated in a competitive match or medium- or large-sided game during training sessions between the 1st of January 2023 and the 8th of May 2023 and to be

over the age of 18 years old. Players were excluded if they had not participated in any medium- or large -sided games during training, not competed in a competitive match between the previously stated dates or were under the age of 18.

Consent to use the data has been willingly released by Gillingham Football Club to be used in this study. Gillingham Football Club were informed that all data used throughout the study would be anonymised.

Collection of Physical and Performance metrics

Each player was individually assigned a Global Positioning System that they wore during all training sessions and matches. Using the GPS Catapult Vector Core Plus (CATAPULT, Melbourne, Australia) physical metrics completed by each individual player were collected, these were collected during training sessions and competitive matches and analysed post training and matches throughout the second half of the 2022/23 EFL League two season and the F.A Cup, from the 1st of January 2023 and the 8th of May 2023.

Along with this, the use of Wyscout (https://platform.wyscout.com/app/) allowed for the collection of match performance statistic, expected goals.

Following both training and matches, the GPS data was extracted onto the Catapult software for analysis. For this investigation the ethics were approved by SSES local ethics committee.

Data Collection

Inclusion and exclusion criteria for data

The inclusion criteria for the data included: match data from competitive matches and data collected during medium- and large-sided games. Players were included if they had completed one or more medium- or large sided games during training.

The exclusion criteria for data included: match data recorded in non-competitive matches, data with missing values (such as signal dropouts) and training data not collected during medium- or large-sided games.

Training

The data collected in the training aspect of the study included: Distance per minute, high speed distance (> 18km/h) collected during training drills with match-like aspects, such as medium- and large-sided games. Research suggests that the median high-speed distance of professional football players is 15km/h, however the default value is 19.8km/h (Abt and Lovell, 2009), therefore, 18km/h was chosen due to it being Gillingham FC's default value, who themselves are a professional football establishment, as well as being between the two values stated from research. This data was collected from 17 training sessions which lasted between 45 minutes to 90 minutes. The medium- and large sided games within these sessions typically lasted

between 8 to 12 minutes, with constant ball in-play during the duration of the games.

Match

The data collected in the match aspect of the study included the physical metrics, including distance per minute and high-speed distance. Alongside this, the metric collected on WYScout used during this study was expected goals. This data was collected from 25 competitive matches throughout the 2022/23 season, with each match lasting 90 minutes in addition to injury time but excluding any extra time.

Protocol

Training protocol

Prior to the beginning of training sessions, the participants were required to collect their GPS units before training commenced, these were provided to each player and turned on by the Sports Science team before the players arrived for training. To ensure the players received their GPS a member of the Sport Science team would hand out the GPS. The players were then instructed to complete the training session planned for the team as normal, which included a 10-to-20-minute warm-up and 5-to-10-minute technical drill before undertaking the focussed drills of this investigation, to allow for

accurate results without biases from players or staff to affect the data. Medium- and large-sided game training drills were where the data collection was focussed for this observational study, game realistic drills ranging from 5 vs 5 to 11 vs 11 players, with pitch areas varying with increasing participants, 5 vs 5, 73.6m^2 per player, 6 vs 6, 110.3m^2 per player, 7 vs 7, 107.1m^2 per player, 8 vs 8, 150m^2 per player, 9 vs 9, 166.67m^2 per player, 10 vs 10, 202m^2 per player, 11v11, 324.9m^2 per player. The areas for each MSG or LSG were decided by Gillingham Football Club coaching staff, familiarisation of these drills had been completed throughout the first half of the season. During the LSG or MSG drills the players were not permitted to view the data to ensure there were no performance effects during training, although verbal encouragement was used by the staff throughout the drills.

Following this, the GPS was removed from the individual's player vests post training, this data was then downloaded to the Catapult software where it could be viewed and analysed. Analysis of the data was then completed by cropping the data into individualised sections of the training sessions, focussing on the sections where an MSG or LSG were completed, the procedure of analysing data was the method used by Gillingham Football Club which involved cropping the data into different drills, within the MSG and LSG drills scheduled breaks were not excluded. This protocol was repeated for every training session from the 1st of January to the final training session on the 7th of May.

Match day protocol:

Prior to competitive matches, the GPS were inserted into the players individual vests and placed at their personal seats in the changing room. Following, the players undertook a 25-to-35-minute pre-match warm-up which included stretching, possession and, defensive and attacking minded drills. The players then participated in the match, typically 324.9m^2 per player, and data was downloaded from the GPS to the Catapult software where the data was viewed and analysed. Analysis of Match Day data was completed by cropping the data into individualised sections; the warm-up, the first half and, second half, excluding the scheduled break between the two halves to allow for accurate distance per minute measurements.
During the matches, the expected goals metric was recorded, this was completed by an analyst, the analyst monitored and entered statistics to the WYScout software during and post matches. This data was then uploaded to WYScout, the expected goals statistic was then obtained from WYScout and used for analysis alongside the high-speed distance and distance per minute data obtained from the GPS software.

Statistical analysis

Following the completion of the data collection from protocols for both training and matches, IBM SPSS Statistics 27 software and Microsoft Excel were used to analyse that data collected.

Statistical Analysis of research question: Does training running intensities compare to match running intensities in professional football?

Statistical tests were executed to identify the statistical significance of each component of this observational field study, including a Univariate Analysis of Variance analysis with an LSD Post-Hoc test to identify the statistical significance of the match values high-speed distance and the distance per minute GPS metrics in comparison to the high-speed distance and distance per minute training session values.

Statistical Analysis of research question: Will running intensity during a competitive football match impact the performance outcome of expected goals?

A bivariate correlation statistical analysis was used to determine the statistical significance of match intensity on expected goals. The variables used included: total mean high-speed distance and distance per minute for all positions together and, high-speed distance and distance per minute of each position separately, Forwards, Midfielders, and Defenders, these variables were used analysed against expected goals.

A Shapiro-Wilks test was also performed for all data sets to indicate if assumptions for parametric statistical tests were satisfied. Verification of the statistical tests performed during this investigation was realised, $P > 0.05$.

Statistical significance:

Using a statistical significance value of $P < 0.05$ for this study means a significance value less than 0.05 would indicate a significant relationship between the training values and match day values of high-speed distance and distance per minute, as well as a significant correlation between match high-speed distance and distance per minute and expected goals. However, if the statistical significance value is more than 0.05, the data would be deemed as being not significant. Lehmann (1958) stated that, fixing the significance level to be accepted as statistically correlating at $P < 0.05$ would allow for a control over the probability of a false rejection, type I error.