Technische Universität München

Fakultät für Sport- und Gesundheitswissenschaften

Leistungsanalyse im Elitefußball

Guilherme de Sousa Pinheiro

Vollständiger Abdruck der von der Fakultät für Sport- und Gesundheitswissenschaften der Technischen Universität München zur Erlangung eines

Doktors der Philospohie (Dr. phil.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Joachim Hermsdörfer

Prüfende der Dissertation:

1. Prof. Dr. Martin Lames
2. Prof. Dr. Tiago Russomano

Die Dissertation wurde am 31.01.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Sport- und Gesundheitswissenschaften am 08.03.2022 angenommen.

I. ABSTRACT

The present dissertation is a combination of publications that aim to expand the scientific and practical knowledge in performance analysis in elite football. Chapter one focused on the analysis of penalty kicks. The process of designing and validating a new instrument for analyzing penalty kicks in football (OSPAF) is described. Utilizing the expert-validated tool (i.e., OSPAF), an empirical study was carried out to distinguish the strategies adopted by the penalty taker and goalkeeper. Subsequently, a novel approach was adopted to automatically detect body angle orientation from video data (i.e., OpenPose), combined with the notational analysis proposed earlier. The developed system (OSPAF) evidenced content validity, inter-and intra-reliability for analyzing penalty kicks in football, using a gold standard methodology for instrument validation. Body orientation analysis using Openpose has shown sufficient reliability and provides practical applications for analyzing the strategies adopted by goalkeepers in penalty kicks in elite football. Chapter two addresses relevant aspects of performance analysis related to investigating the physical and physiological demands of different training tasks. The aim was to examine the differences in external and internal load during pre-season training sessions with different small-sided games (SSGs) and a friendly match in top-class professional football players. The present findings indicated that external and internal loads differ across different SSGs and a friendly match (FM) during the pre-season. Performance analysis is a discipline of sports science that presents a broad analytical approach. The results in chapters one and two provided practical applications for coaches and other football professionals and offered support for future research.

Keywords*: penalty kicks, observational analysis, OpenPose, acceleration, training load, soccer*

CONTENTS

II. SUMMARY

CHAPTER ONE: PERFORMANCE ANALYSIS IN PENALTY KICKS

CHAPTER TWO: FURTHER RESEARCH IN PERFORMANCE ANALYIS

III. LIST OF FIGURES

IV. LIST OF ABBREVIATIONS

OSPAF: Observational System for Penalty Kick Analysis in Football
SSGs: small-sided games
FM: friendly match
U19: under 19 years old category
IFAB: International Football Association Board
FIFA: International Federation of Association Football
PA: performance analysis
TV: television
PhD: Doctor of Philosophy
K: Cohen's kappa
V: Aiken´s V
DLT: direct linear transformation
LR: left-right
LRSh, LRHi, LRSF: penalty kicker's shoulder
LRHi: penalty kicker's hips
LRSF: penalty kicker's non-kicking foot orientation
αSh: penalty kicker's shoulder
αHi: penalty kicker's hip
αSF: penalty kicker's non-kicking foot
CSh: confidence score of the orientation of penalty kicker's shoulder
CHi: confidence score of the hips
CSF: confidence score of the non-kicking foot
αGK: Coordinates belonging to the neck and hip of the goalkeeper at the two moments are mapped to a vector and the angle is calculated
DGKR: the movement distance of the right foot
DGKL: the movement distance of the left foot

CGK: represent confidence score

$\chi 2$: chi square test

ES: effect size

M: median

L: lower 95% confidence interval limit

Vp: Aiken´s V value of the pilot study

VAG: Aiken´s V value of agreement dimension

VAD: Aiken´s V value of adequacy dimension

VU: Aiken´s V value of univocity dimension

rpb: point biserial correlation

ETPS: electronic performance-tracking systems

GPS: Global Positioning System

LPS: Local Positioning System

CBF: Brazilian Football Confederation

CONMEBOL: South American Football Confederation

GK: Goalkeeper

HR: heart rate

Min: minute

Bpm: batiments per minute

m: meter

m/min: meter per minute

h: hour

°C: celsius

V. FULL PUBLICATION'S LIST

I Pinheiro, G.S., Nascimento, V.B., Dicks, M., Costa, V.T., Lames, M. (2021). Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF). *Frontiers in Psychology*, 12, 661179. https://doi.org/10.3389/fpsyg.2021.661179

II Pinheiro, G.S., Costa, V.T., Lames, M. (2021). Penalty kicks in elite football: identifying factors related to the player strategy. *International Journal of Sport and Exercise Psychology - 15th World Congress Proceeding of the International Society of Sport Psychology (ISSP)*, 19:sup1, S228-229, doi: 10.1080/1612197X.2021.1982479

III Pinheiro, G.S., Xing, J., Costa, V.T., Lames, M. (2022). Body pose estimation integrated with notational analysis: a new approach to analyze penalty kicks strategy in elite football? *Frontiers in Sports and Active Living, section Sports Science, Technology and Engineering.* doi: 10.3389/fspor.2022.818556

IV Pinheiro, G.S., Quintão, R.C., Nascimento, V.B., Claudino, J.G., Alves, A. L., Costa, I.T., Costa, V.T. (2022). Small-sided games do not replicate all external and internal loads of a football match-play during pre-season: A case study. *International Journal of Sports Science & Coaching*. https://doi.org/10.1177/17479541211069935

1 INTRODUCTION

1.1 Theoretical framework

The penalty kick in football is arguably one of the world's most highly visible and high-pressured sporting situations (Ellis and Ward, 2021). In a football match, both teams' average number of goals is typically low during regulation time (i.e., 2.7; Gürkan et al., 2017). Consequently, the scoring opportunity provided by a penalty kick can strongly influence the outcome of a match (Horn et al., 2020). In addition, during a decisive penalty shoot-out, the importance to the match's result is even more apparent (Paterson et al., 2020), as it may ultimately affect the final standings of a league or competition (Almeida et al., 2016; Göral, 2016).

Most studies have reported success rates of penalty kicks in professional football between 70 and 80.5% (Almeida et al., 2016; Fariña et al., 2013; Göral, 2016; Bar-Eli et al., 2007). However, penalties were often missed at vital moments in professional football matches (Slutter et al., 2021). Many factors influence the overall chances of scoring on a penalty kick (Memmert et al., 2013). Several scientific studies have identified the motivational, strategic, anticipatory, attention, and perception-based factors that can mean a successful or failed penalty kick (Memmert & Noël, 2020). Recent research focusing on the technical dynamics of penalty kicks has also identified multiple key variables that can differentiate the player's strategy (Noël et al., 2015) and enhance the overall chances of scoring a penalty kick (Jamil et al., 2020).

Given the interactive nature of the penalty kicks (Furley et al., 2020), the outcome relies on the performance emerging from the ‘penalty taker - goalkeeper’ dyadic interactions (Lopes et al., 2012; Almeida et al., 2016). Penalty takers and goalkeepers make spatial and temporal decisions about the goal side (Noël et al., 2021). Goalkeepers must (strategically) decide to dive early or wait longer to commit to one side or remain in the middle of the goal (Savelsbergh et al., 2010). At the same time, penalty-takers must decide before the run-up where to shoot (i.e., keeper-independent strategy) or wait for the goalkeeper to commit to a side to subsequently kick to the other side of the goal (i.e., keeper-dependent strategy; van der Kamp, 2006; Noël and van Der Kamp, 2012; Kuhn, 1988). Although penalty kicks are widely studied in football performance analysis, the interaction between the two players has received much less attention (Noël et al., 2021). Research on performance analysis has pointed out that the future of game analysis in football requires building observational instruments that integrate the study of criteria related to the interaction with the opponent (Lames, 2006; Mackenzie and Cushion, 2013; Sarmento et al., 2014).

The tactical evaluation applying observational systems (also known as "notational systems") has increased in terms of recognition and development over the past years (Nevill et al., 2008). Observational methods are scientific procedures that reveal the occurrence of perceivable behaviors, allowing them to be formally recorded and quantified (Anguera et al., 2001; Lames & Hansen, 2001). To support top-level teams, two purposes of game observation are predominant: preparation against a future opponent and the optimization of training (Lames & Hansen, 2001; Lames & McGarry, 2007). They also enable an analysis of the relationships between these behaviors, such as sequentially, association, and covariation. However, one potential limitation is that one method does not supply all the necessary information entirely. Therefore, there is a need to use a multi-method approach to solve sports analytics problems, analyzing variables using different methods (Aranda et al., 2019). Furthermore, methodology designs that combine different research approaches (e.g., observational and method that produce body angles), also known as mixed methods (Preciado et al., 2019), tend to provide a deeper understanding and reliability of the studied phenomenon (i.e., penalty kicks).

Nevertheless, Brewer and Jones (2002) had previously noticed that information related to validity and reliability concerning deductive processes of systematic observation was insufficient and poorly obtained. Those concepts are essential to improve accuracy in research measurement. Validity generally refers to the ability of a measurement tool to reflect what it is designed to measure, and usually, for performance analysis instruments, it can be determined through expert coaches' opinions in each sports category (O'Donoghue, 2009). Moreover, reliability is the consistency of a measure and is a part of the validity evidence (Sullivan, 2011; Heale and Twycross, 2015). It refers to the reproducibility of values of a test, assay, or other measurements in repeated trials on the same individuals (intra-observer reliability) and repeatability over different observers (inter-observer reliability) (O’Donoghue, 2009). An instrument must be reliable and valid, evidenced by various methods (Heale and Twycross, 2015).

Specifically, it is highly recommended to establish content validity and reliability of notational systems and observational instruments because it can reduce the error caused by human subjectivity (O’Donoghue, 2009; Cobb et al., 2018). Previous research has used a development process to evidence the content validity of an observational instrument (Fernandes et al., 2019). This process includes the following sub-stages: (1) literature review; (2) instrument development; (3) observation training; (4) development of the new observation instrument; (5) pilot study; (6) establishment of content validity with experts; (7) inter-observer reliability and intra-observer reliability assessment. The current dissertation designed a new observational system of penalty kicks analysis in football following the development process. An innovative approach was adopted to analyze players' body angles from video data based on this system. A novelty of this study is the adoption of notational analysis (i.e., OSPAF) combined with body orientation measurements (i.e., OpenPose) to analyze penalty kicks. Multiple practical applications can be provided, from identifying, improving, and refining player strategy in penalty kicks to accurately assessing player orientation in high-level competitive scenarios.

1.2 Scientific problem

Several studies analyze penalty kicks in field settings (Brinkschulte et al., 2020; Higueras-Herbada et al., 2020; Wunderlich et al., 2020; Wood et al., 2015; Dalton et al., 2015; Bowtell et al., 2009). Most of them do not aim at giving a detailed description of the actions of the shooter and goalkeeper but focus more on statistics of results (e.g., quotes for scoring and saving penalties). Few instruments were proposed to evaluate penalty kick strategies in football (e.g., Comas et al., 2018; Noël et al., 2015). Comas et al. (2018) created a system based on the observational methodology to analyze the ball's direction at a penalty kick in football. The authors indicated a relationship between the spatial position of the penalty taker support foot and the opposite arm to the shooting foot with the ball's direction on the penalty kick. However, the non-kicking foot orientation assessment, which is a central variable for the study results, may be difficult to detect through observation, given the small size of the object of interest compared to the large volume containing the necessary information elements for recording a penalty kick. Also, the lack of standardization of video angles used in the study could make it difficult to detect through observation. Noël et al. (2015) developed a method for identifying penalty kick strategies in a non-competitive controlled simulated situation. These researchers included 12 variables in this observational system. A logistic regression model identified three variables (attention to the goalkeeper, run-up fluency, and kicking technique) that, in combination, predicted kick strategy in 92% of the penalties. However, the high prediction accuracy could be inflated because penalty takers were instructed to use either penalty strategy. In addition, the lack of a familiarization and training phase with each type of strategy added to the fact that they are inexperienced players (category U19) could influence the study results. That proposed model and procedures may not be sensitive enough to identify players' interactions. In real competitions, penalty kicks are an interaction process. The observable performance is rather the emergent result of this interaction process than the display of skills and abilities of the two parties (Lames, 2006).

Although the distinction between strategies in penalty kicks is common in the scientific literature and football, there is a lack of a valid and reliable method for distinguishing the penalty strategies. Using a reliable method may have implications for scientific researchers, performance analysts, and coaches who seek to identify penalty takers' likely strategy. It would allow for replications of studies to track, for example, long-time trends and for comparisons between different settings (e.g., countries, leagues, age groups, gender). Researchers would be able to identify determinants of successful kicks (e.g., kick coordination, patterns of gaze, anticipation), especially under high pressure. In addition, practitioners in professional football could distinguish penalty kick strategies and so inform coaching, training, and scouting (Noël et al*.,* 2015).

Body orientation has been indicated as a critical factor under covering the success in penalty kicks (Li et al., 2015). However, it is a yet-little-explored area in penalty kick analytics. There is a need within human movement sciences for a markerless motion capture system (e.g., OpenPose), which is easy to use and sufficiently accurate to evaluate motor performance (Nakano et al., 2020). The innovative approach proposed in this dissertation, using the OSPAF and applying technological methods to analyze its variables, such as computer techniques for body pose estimation, may provide a deeper understanding of the factors influencing the players' strategy in penalty kicks.

2 LITERATURE REVIEW

2.1 The penalty kick

A penalty kick is awarded if a player commits a direct free kick offense inside their penalty area or off the field as part of play, as outlined in Laws 12 and 13. A goal may be scored directly from a penalty kick. The ball must be stationary on the penalty mark, and the goalposts, crossbar, and goal net must not be moving. The player taking the penalty kick must be clearly identified. The defending goalkeeper must remain on the goal line, facing the kicker, between the goalposts, without touching the goalposts, crossbar, or goal net, until the ball has been kicked. The players other than the kicker and goalkeeper must be at least 9.15 m from the penalty mark, behind the penalty mark, inside the field of play, outside the penalty area. As soon the players have taken positions following this Law, the referee signals for the penalty kick to be taken. The player taking the penalty kick must kick the ball forward; back heeling is permitted provided the ball moves forward. When the ball is kicked, the defending goalkeeper must have at least part of one-foot touching, or in line with, the goal line. The ball is in play when it is kicked and moves. The kicker must not play the ball again until it has touched another player. The penalty kick is completed when the ball stops moving, goes out of play, or the referee stops play for any offense (IFAB, 2021-2022).

UEFA introduced penalty shootouts to major tournaments in 1976 (FIFA followed in 1978) to decide matches in the knockout phase of major tournaments when the score is a draw at the end of the match. The penalty shootout has become the standard tie-breaking procedure in knockout tournaments (Avugos et al., 2020). Since 1978 about every fifth knockout game at the FIFA World Cup has been decided by penalty shootouts. This fact results in an estimated chance of 67% that a future world champion must complete at least one penalty shootout on his way to the title (Memmert and Noël 2020). Since 2009 almost 100,000 penalty shots have been taken on football pitches around the globe. Penalty takers scored 75.49% of the penalties; goalkeepers saved 17.57%, 4.07% went wide, and 2.87% hit posts or crossbars (Instat, 2020).

2.2 Performance analysis and the case of the penalty kicks

Performance analysis (PA) of sport is a relatively recent discipline of sports science. Its history is composed of biomechanics and the history of notational analysis, with these two disciplines coming together within performance analysis in 2001 (O'Donoghue, 2010). PA investigates actual sports performance or performance in training (O'Donoghue, 2010). The main reason for doing PA is to understand sports that can inform decision-making by those seeking to enhance sports performance. The complexities and dynamic nature of many sports mean that observation and measurement are needed to improve our understanding of performance (O'Donoghue, 2010). The rationale for using PA is to overcome the limitations of subjective observation alone and to provide objective information to achieve a greater understanding of the performance. This information, in turn, assists decision-making by coaches and may, therefore, play a vital role in performance enhancement (O'Donoghue, 2010).

Penalty kicks have been mainly analyzed in two contexts: First, in a laboratory or other non-game controlled settings (video-simulation and in-situ experimental conditions), aiming at the analysis of perceptual-motor and cognitive aspects of performance (e.g., Dicks et al., 2010; Lopes et al., 2012; Weigelt and Memmert, 2012; Navarro et al., 2013); and second, in real match situations, enabling the identification of prominent factors that affect both players' performances and the penalty kick outcome using mainly observational methods (e.g., Chiappori et al., 2002; Jordet et al., 2007; White and O'Donoghue, 2013; Horn et al., 2020). While in the first context, a common theoretically motivated focus has been developed to enhance the representative design of methods used to examine the expertise of penalty takers and goalkeepers (Dicks et al., 2009), in the second, researchers have attempted to improve data collection procedures based on game video analysis (Almeida et al., 2016). In the present thesis, the data collection procedures will be investigated.

2.3 Penalty kick strategies

A prerequisite for increasing the probability of success in a penalty kick is implementing the best penalty kick strategy. Previous research has identified two main strategies for taking a penalty (Kuhn, 1988; van der Kamp, 2006). Firstly, the keeper-independent strategy, where the kicker selects the target location to shoot toward before the run-up and does not attend to the actions made by the goalkeeper during the run-up. The decision of where to aim depends on the penalty taker's kicking preference (Noël et al., 2015). On the contrary, the kicker tries to obtain information from the goalkeeper's reactions during the run-up in the keeper-dependent strategy. Nevertheless, the outcome of a penalty is determined by an interaction between the shooter's strategy (e.g., technique, speed) and the goalkeeper's strategy (Hunter, et al., 2018). Therefore, the optimal strategy depends on the keeper's behavior and the relative benefits of speed, accuracy, and unpredictability within each situation. There are two approaches regarding the goalkeeper strategy: the dependent and independent penalty taker. The goalkeeper who behaves according to the first group defines his movement based on the actions of the penalty taker. The second type of goalkeeper is the one who risks jumping to a corner independently of the kicker's movement (Kuhn, 1988).

Kuhn (1988) suggested that around three-quarters of penalty takers use the keeper-dependent strategy, but he did not report whether the strategy is more successful than the keeper-independent strategy. (Kuhn referred to these strategies as ‘‘closed’’ and ‘‘open loop’’, respectively.) By anticipating the side that the goalkeeper will dive, the penalty kicker intends to decrease the probability that the goalkeeper can reach the ball and save the kick. It can be supported by the use of early advanced information concealed in the goalkeeper’s postures and movements. Still, it may also be facilitated by knowledge about a particular goalkeeper’s preferred side. The keeper-dependent strategy appears particularly advantageous when the goalkeeper commits himself early. It is perhaps for this reason that many players prefer the keeper-dependent strategy over the keeper-independent strategy (van der Kamp, 2006).

Morya et al. (2003) suggested that taking into account the goalkeeper’s actions might seriously impede the successful conversion of the penalty kick. The employment of a keeper-dependent strategy appears no guarantee that the ball is placed in the direction opposite to the goalkeeper's movement (Morya et al., 2003). This fact is particularly evident when information about the goalkeeper's dive is detected shortly before ball contact, as there may remain insufficient time to alter the direction of the kick. A late decision may not only result in placing the ball to the same side as the goalkeeper but might also result in a relatively inaccurate placement. Therefore, the time available to alter the direction of the penalty kick appears a critical factor for the success or failure of the keeper-dependent strategy.

Previous research has pointed that the keeper-dependent strategy may also be problematic as a penalty-taking method because of the constraints placed on perception (or visual attention) during the run-up (van der Kamp, 2006). A gaze fixation on the target would ensure accurate control of aiming movements (Land & Furneaux, 1997; Norman & Shallice, 1986). A penalty taker searching for predictive information about the goalkeeper's intention will fixate the goalkeeper instead of the target location (van der Kamp, 2006). This may reduce the accuracy of the penalty kick, even if the direction of the penalty kick and thus the kicking action is not altered (i.e., independent of the constraints on the action). However, the keeper-independent strategy neither curbs the spatial location of gaze fixations nor their timing. On the contrary, it permits an optimal pattern of gaze fixations for accurate aiming of the kick (van der Kamp, 2006).

The differentiation between the two strategies has been investigated in relation to numerous factors, such as spatiotemporal (e.g., ball speed, run-up; Kuhn, 1988), perceptual (e.g., visual search behaviors; van der Kamp, 2006), individual (e.g., footedness, age, positional role; Almeida et al., 2016; Avugos et al., 2020), or psychological (Memmert and Noël, 2020). According to Kuhn (1988), ball speed could distinguish strategies (even though a keeper-dependent strategy does not necessarily preclude a forceful kick), it is not the only distinctive feature. Ratings for all items for the run-up (i.e., its fluency and length and the length of the last step) differed between strategies. For penalty kicks with the keeper-independent strategy, the run-up seems to be more fluent, and the full run-up and last step distance is longer than for kicks with the keeper-dependent strategy (Noël et al., 2015). The difference in fluency is probably a consequence of penalty takers who use a keeper-dependent strategy to increase time at the end of the run-up by waiting for the goalkeeper to commit to one side of the goal (van der Kamp, 2006). Noël and Van der Kamp (2012) found that those penalty takers who use a keeper-dependent strategy spend more time looking at the goalkeeper throughout the run-up and kick execution than penalty takers who use a keeper-independent strategy. The later attend longer to the target area (in the preparatory phase before the run-up to the ball) and the ball (during the run-up and kick execution). These differences in gaze are correlated with the accuracy of the kick; the longer penalty takers look at the goalkeeper – rather than the ball – the closer the kick is to the goalkeeper (van der Kamp, 2006).

However, Memmert et al. (2013) argued that it is primarily not possible to determine which shooting strategy will be the most successful. Instead, it depends on the shooter's technical ability (e.g., it is typically easier for a player to score when aiming to the opposite side of his strong foot) and the goalkeeper's ability. Given this interaction process between the goalkeeper and the penalty kick taker in football, it is necessary to adopt instruments capable of assessing the interaction in this set piece. Therefore, it is essential to understand the factors that limit performance in kicking and defending penalties to determine the best strategies to prepare for this often-critical game moment (Reilly et al., 2005).

3 OBJECTIVES

3.1 General objective

This thesis seeks to expand the sport-scientific knowledge of performance analysis in penalty kicks in elite football by: (i) developing and applying a method to analyze the penalty taker and goalkeeper strategy, and (ii) by applying an innovative approach method to extract body angles in the penalty kick, integrated with notational analysis.

3.2 Specific objectives

3.2.1 Specific objective: paper 1

Paper 1, entitled "Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF)," carried out a methodological design containing three studies (pilot study, main study, and video requirements study). The pilot study aimed to get formal feedback on variables for penalty kick analysis suggested by professionals in the area; the main study aimed at designing and validating an observational system applied to in-match penalty kick analysis; and the video study served to evaluate the influence of the video footage (i.e., viewing angles, number of angles and video quality) on penalty kick analysis through an observational system.

3.2.2 Specific objective: Paper 2

Paper 2, entitled "Penalty kicks in elite football: identifying factors related to the player strategy," aimed to identify the relationship between a set of observable variables and the players' strategy using an expert-validated observational system for penalty kick analysis in football (OSPAF).

3.2.3 Specific objective: Paper 3

Paper 3 title is "Body pose estimation integrated with notational analysis: a new approach to analyze penalty kicks strategy in elite football?". It analyzed whether OpenPose can detect relevant body orientation angles from video data of penalty kicks in elite football; and (ii) investigated the relationship between these body angles and observable behaviors analyzed via OSPAF with the penalty taker and goalkeeper strategy.

4 PUBLICATION LIST

The following publications are presented in support of chapter one of the thesis:

I Pinheiro, G. S., Nascimento, V. B., Dicks, M., Costa, V. T., & Lames, M. (2021). Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF). *Frontiers in Psychology*, 12, 661179. https://doi.org/10.3389/fpsyg.2021.661179

II Pinheiro, G.S., Costa, V.T., Lames, M. (2021). Penalty kicks in elite football: identifying factors related to the player strategy. *International Journal of Sport and Exercise Psychology - International Society of Sport Psychology (ISSP) 15th World Congress Proceeding*, 19:sup1, S228-229, doi: 10.1080/1612197X.2021.1982479

III Pinheiro, G.S., Xing, J., Costa, V.T., Lames, M. (2022). Body pose estimation integrated with notational analysis: a new approach to analyze penalty kicks strategy in elite football? *Frontiers in Sports and Active Living, section Sports Science, Technology and Engineering.*

5 MATERIALS AND METHODS

5.1 Ethical procedures

Participants (study 1) provided informed consent after details of the study were communicated in written form before participation in the study. The following studies 2 and 3 did not involve data collection with human beings, so submission to the ethics committee was unnecessary. All procedures performed in the study were in strict accordance with the Declaration of Helsinki and the ethical standards of the Technical University of Munich.

The videos were recorded from TV broadcasters and were registered and analyzed post-event. As the video recordings were public, confidentiality was not an issue, and authorization was not required from the players observed or their representatives.

5.2 Experimental design: Paper 1

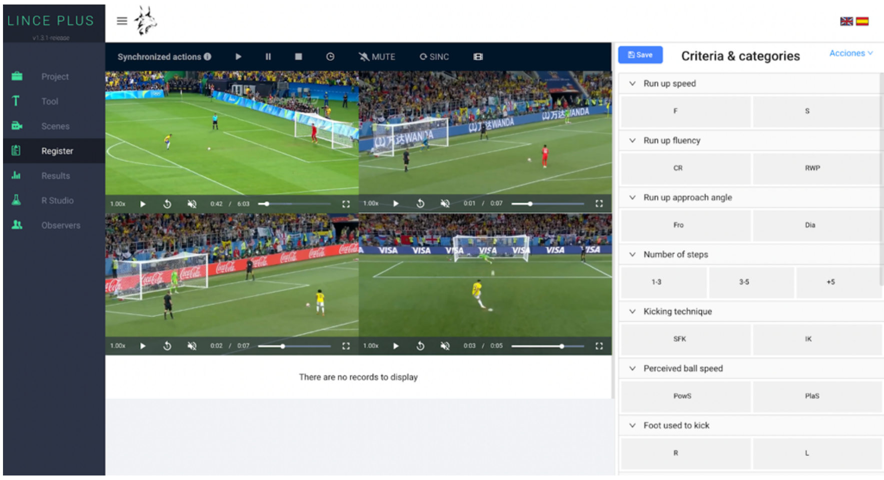
5.2.1 Subjects

Participants in the pilot study were four sports scientists and three high-level football coaches (43.32 ± 15.48 years). A panel of 20 experts (41.85 ± 13.96 years), from Brazil, England, Germany, Israel, Netherlands, Romania, and Spain, who met the following criteria: (1) Ph.D. in sports sciences, and (2) experience of publishing in penalty kick research was contacted and voluntarily agreed to participate. The inclusion criteria established for forming part of the panel of sports scientists were: (1) postgraduate master in sports sciences or Ph.D. in sports sciences, (2) to have had at least three years experience as a university researcher in sports sciences, (3) experience in performance analysis research (final master's thesis, doctoral thesis or scientific publication); and for high-level football coaches were: (1) graduate in physical activity or sport sciences, (2) have an official license as a football coach, (3) more than three years as a football coach in a team of an official competition. In addition, more detailed characteristics about the experts were collected, such as sports biography and open items on the experts' general judgment on each criterion.

5.2.2 Procedures

A methodological design containing three studies (pilot study, main study, and video requirements study) was carried out. In the pilot and main study, a survey with two different versions was developed in Google Forms to assess content validity with the experts. The video study utilized another online survey containing penalty kick videos (i.e., 2016 Olympics, World Cups between 2010 and 2018, and major European leagues from 2015 to 2020). The experts were instructed to answer the questionnaire on a computer or notebook, and there was no time limit to answer the questions. Experts could watch each penalty kick video as many times as they judged necessary. For reliability, the final version of the OSPAF was used after implementation in Lince Plus software (Gabin et al., 2012; Soto et al., 2019). Lince Plus is free software that has been used by many researchers needing a tool to tag behaviors using video recordings, coding behaviors, and data registers (Soto et al., 2019). Dimensions and categories of OSPAF were coded, and the observations of the two observers were compared using this software. Criteria were entered with the full definition of the variable (i.e., Run up speed), and categories were coded with the initial letters (i.e., Fast = F and Slow = S), as illustrated in the figure below.

Figure 1. Lince plus interface



Pilot Study

The pilot study refers to a mini version of the full-scale study and the specific pre-testing of the particular research instrument, here the online questionnaire (Teijlingen and Hundley, 2001). The pilot study aimed to get formal feedback on variables for penalty kick analysis and to collect observable variables suggested by professionals in the area.

Main Study

The main study aimed to follow a systematic process to accumulate content validity and reliability evidence to adequately categorize and record behaviors of both penalty takers and goalkeepers during penalty kicks. A panel of experts answered the online survey. The level of concordance among experts for each variable proposed in the OSPAF was analyzed. A modified Delphi method was performed (Dalkey and Helmer, 1963; Hasson et al., 2000; Dayé, 2018). The process to achieve content validity for the OSPAF was adapted from Brewer and Jones (2002) and Fernandes et al. (2019), including content validity with experts and inter-and intra-Observer reliability assessment.

For concordance analysis (content validity) three dimensions were defined (Fitzpatrick, 1983; Fernandes et al., 2019):

Table 1. Concordance analysis in the main study

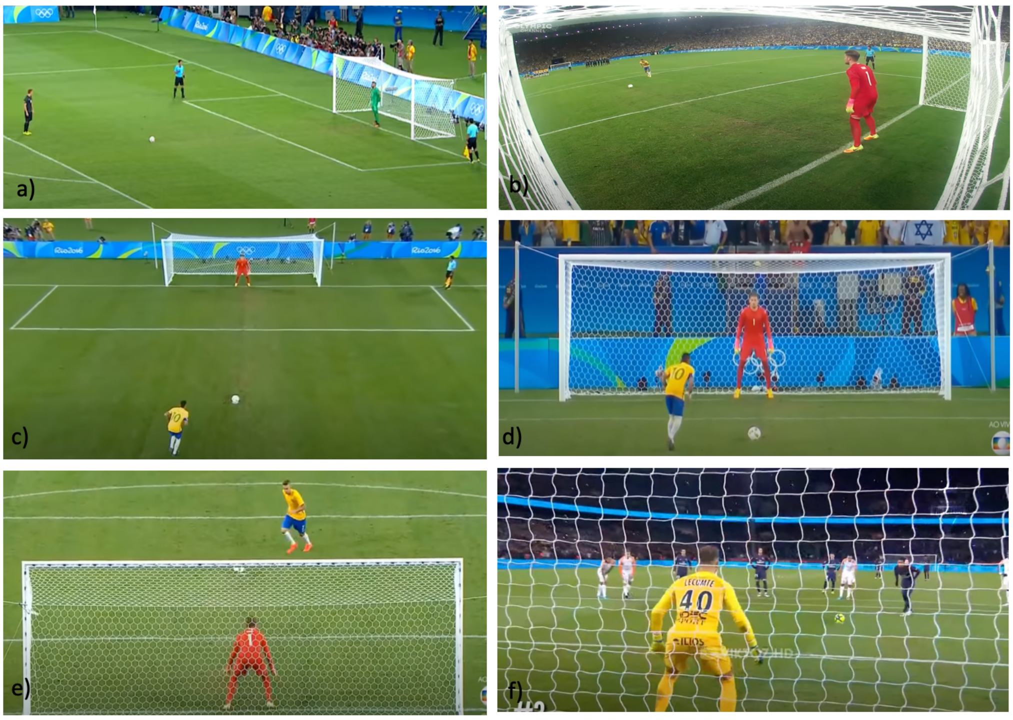
Dimension	Description	Question in the survey	Measurement
Agreement	degree of general acceptance of the variables to be included in the observational system	How is your level of agreement with the inclusion of the variable for penalty kick analysis in the proposed system?	Five-point Likert scale (Strongly disagree, Disagree, Neither disagree nor agree, Agree, Strongly agree)
Univocity	clarity domain of a definition	The definition of the variable is clear enough for understanding?	A binary scale (Yes or No)
Adequacy	level of pertinence and importance of criteria	What is the level of importance of the variable for the observational system?	A different five-point Likert scale (Very low, Low, Medium, High, Very high)

The verification of the reliability of OSPAF was made through the assessment of Cohen's kappa (κ) between observers (inter-observer agreement) and for the analysis of interpretative stability within one observer (intra-observer agreement). For the inter-observer agreement, apart from the analysis carried out by the principal researcher, a second researcher was trained to analyze the penalty kicks with OSPAF. After the training period, the two observers independently analyzed 40 randomly selected penalty kicks of the World Cups 2014 and 2018. Regarding the intra-observer agreement, the principal investigator performed the same analysis four weeks after the first analysis, thus minimizing task familiarity (Robinson and O'Donoghue, 2007), without conducting any analysis during this time, thus checking the temporal stability of the analysis (Aranda et al., 2019).

Video Study

Using an online questionnaire, 14 penalty kick videos from elite football, each from 7 different angles, were presented (Figure 2).

Figure 2. Penalty kick viewing angles



Legend: a) Lateral side; b) Lateral side with camera inside the goal; c) Behind the penalty taker aerial view; d) Behind the penalty taker pitch view; e) Behind the goalkeeper aerial view; f) Behind the goalkeeper pitch view; g) Rotational angle: this viewing is a rotational angle, which occurs when the penalty taker initiates the approach to the ball. The three images show some moments of the rotation. (Figures and videos are of public domain).

The methodology adopted in the present study is similar to Baranowski and Hecht (2017) (i.e., fifteen-second scenes were used as examples, and later on, a questionnaire was applied to gather feedback). The videos had a pixel resolution of 1,280 × 720. The experts should indicate the best viewing angles for penalty kick analysis. They were instructed to watch the videos on a computer or a notebook. The choice of angles was adapted from a field division into zones proposed by Garganta (1997) and previously used by Moraes et al. (2014). This division corresponds to the topographical division of the playing field, and its use ensured the establishment of spatial references for choosing the angles.

Besides, the experts were asked how many viewing angles were needed for penalty kick analysis, whether changing the viewing angle could influence the observer's analysis, and whether video quality is a fundamental prerequisite for standardizing penalty kick analysis using an observational system. The level of concordance among experts for the following domains was analyzed:

Table 2. Video study concordance analysis

Domain	**Question in the survey**	**Measurement**
Number of angles needed for penalty analysis	In your opinion, how many video angles are required for the evaluation of a penalty kick in observational studies?	Five-point Likert scale (1 video angle, combination of 2 video angles, combination of 3 video

		angles, combination of 4 video angles, combination of 5 or more video angles)
Influence of changing angles on the observer's analysis	In your opinion, changing the angle presented could influence the evaluation of penalty kicks by an observer?	A binary scale (1. Yes or 0. No)
Pre-requisite of video quality	In your opinion, the video quality is a prerequisite for penalty kick analysis in football?	A binary scale (1. Yes or 0. No)

5.2.3 Statistical Analysis

For descriptive analysis, mean and standard deviation were used. Aiken's V was calculated (Aiken, 1985) for content validity of the OSPAF variables and to evaluate the level of agreement of the experts according to the number of angles needed for penalty analysis; the influence of changing angles on the observer's analysis; and the pre-requisite of video quality. Aiken's V allows for quantifying the relevance of items expressed in Likert scales, according to the opinions of a group of experts. Its values vary between 0 and 1, with 1 indicating a perfect agreement among the judges. Previous studies have used the same coefficient to establish validity in observational instruments (Villarejo et al., 2014; Garcia-Santos and Ibanez, 2016; Fernandes et al., 2019; Ortega-Toro et al., 2019). The p level considered for Aiken's V was 0.05, and a 95% confidence interval was used. The score confidence interval was used to provide the expected accuracy of Aiken's V value (Randall et al., 2009). The calculation of Aiken's V is as follows:

$$V = \frac{\sum s}{n\,(c-1)}$$

Description: n = number of judges; c = highest value of Likert scale; s = r – l; r = the judgement given by a judge; l = lowest value of Likert scale

For each dimension (agreement, univocity, and adequacy), the criteria for the elimination or acceptance of the items were fixed in advance. The reference table proposed by Aiken (1985) for samples with n < 25 was used (number of rating categories: 5; V = 0.64, p < 0.05; number of rating categories: 2; V = 0.75,

$p < 0.05$). Consequently, variables with Aiken's V values below cut-off values of 0.64 in agreement or adequacy, or univocity below 0.75, were eliminated.

For intra- and inter-reliability of the OSPAF, Cohen's kappa was calculated. The interpretation of this coefficient was adopted as follows: $\kappa > 0.8$ very good; $0.6 < \kappa < 0.8$ good; $0.4 < \kappa < 0.6$ moderate; $0.2 < \kappa < 0.4$ fair; $\kappa < 0.2$ poor (Altman, 1991; O'Donoghue, 2009). For the specific objectives of this study only values $\kappa > 0.8$ were considered satisfactory (Lames, 1994).

To identify the best angle for penalty analysis, descriptive statistics were used. Microsoft Excel 2016 was used to calculate the values of Aiken's V and confidence interval; Lince Plus software to record the behaviors (Gabin et al., 2012; Soto et al., 2019). The Statistical Package for the Social Sciences software (IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.) was used to calculate Cohen's Kappa.

5.3 Experimental design: Paper 2

5.3.1 Dataset

The dataset consisted of 150 penalty kicks from the main European football leagues (Premier League, Ligue 1, Bundesliga, LaLiga, Serie A, and Champions League; seasons 2017 to 2020).

5.3.2 Procedures

The videos were recorded by TV broadcasters and were registered and analyzed post-event. All data was annotated by two experienced researchers using the OSPAF. Dimensions and categories of OSPAF were coded in Lince software (Gabin et al., 2012; Soto et al., 2019).

5.3.3 Statistical Analysis

Cohen's kappa (κ) was utilized to verify the reliability of OSPAF (inter-observers and intra-observer). Logistic regression (enter method) analyses were performed. The p level considered < 0.001 (95% confidence interval). All data were analyzed using JASP software (Team, 2020; JASP Version 0.14; Computer software).

5.4 Experimental design: Paper 3

5.4.1 Dataset

The dataset consisted of 34 penalty kicks from the main European football leagues (Premier League, Ligue 1, Bundesliga, LaLiga, Serie A, and Champions League; seasons 2017 to 2020).

5.4.2 Procedures

The researchers annotated all penalty kick data with the OSPAF (Pinheiro et al., 2021). Body orientation was analyzed using Openpose (CMU-Perceptual-Computing-Lab, 2017). The choice and analysis of the penalty kick video viewing angles were standardized, with a pixel resolution of 1280 x 720. The viewing angle used in the present study was the view behind the penalty taker (Figure 3). The confidence score, calculated by Openpose, was used to evaluate reliability (Sangüesa et al., 2019). In order to check the stability within the observation, every penalty kick was analyzed with the OpenPose twice. Retest reliability was utilized to check these repeated measurements.

Body pose detection and orientation

OpenPose (version 1.4.0) was installed from GitHub (CMU-Perceptual-Computing-Lab, 2017) and run with a notebook (Apple´s M1 Chip) under default settings. Orientation from pose used pre-trained models and 3D visualization techniques to obtain a first orientation estimation of each player. Once the pose is extracted for each player, the coordinates and confidence level associated with the body parts are stored to estimate the pose orientation. As a result, in the moving skeletal pictures generated by OpenPose, the skeleton marks are shown and overlapped well with the figure of players (Nakai et al., 2019). For technical details of pose models see Ramakrishna et al., 2014; Wei et al., 2016; Cao et al., 2017.

In the present paper, the orientation of a player's body was defined as the 2D rotation of the player's upper-torso around the vertical axis, which is assumed to coincide with the field projection of a normal vector placed in the center of their upper-torso, involving both shoulders and hip parts (Sangüesa et al., 2019). Especially in the case of the non-kicking foot, the hallux and the fifth toe of the support foot were used as the left-right pair to find the normal vector. Orientation was measured in degrees. For technical details of this methodological approach, see Sangüesa et al., 2019; 2020a; 2020b.

In the current study, two frames were analyzed. Firstly, when the penalty taker starts the run-up into the ball and, secondly, when he touches the ball (Figure 3). Then, the target variables for the penalty taker (non-kick foot orientation, hips, and shoulders) and the goalkeeper (anticipation movement (explained in detail below), right and left foot orientation) were extracted. There might be blurry frames and overlap of players. OpenPose could then fail to detect the main biometric body parts of the two players involved in this analysis. Therefore, in this case, the neighboring frames, in which biometric body parts can be detected, were used.

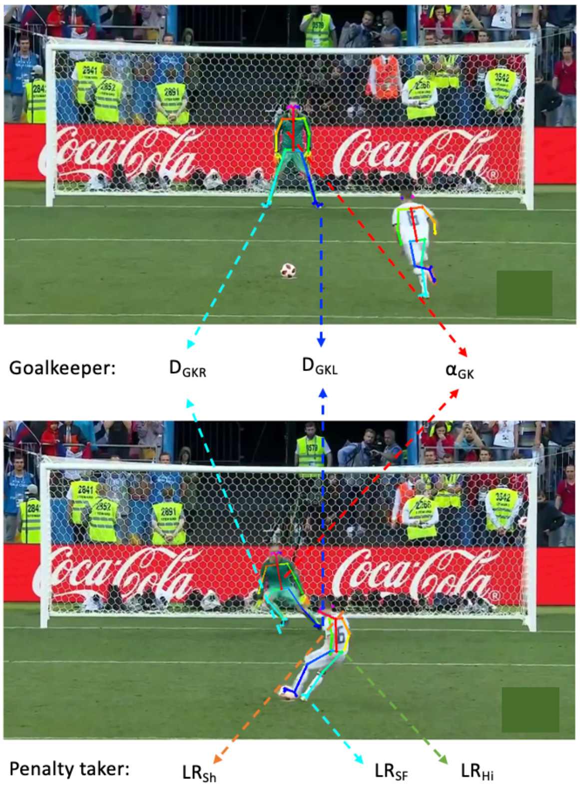
Once the pose was extracted for the goalkeeper and penalty taker, the direct linear transformation (DLT) algorithm (Hartley and Zisserman, 2003) was used to map the coordinate information of players into a 2D field with a homography, given the four field corners' coordinates in the image (or its projection out of the image in the non-visible cases). The homography was first computed based on four 2D to 2D point correspondences between the frame (Equation 1). From the output of OpenPose, the coordinates of the main upper-torso parts are found in the image domain; by mapping the left-right pair (either shoulders or hips) in the 2D field, a first insight of the player orientation is obtained. The player can be inclined towards the right (0-90o, 270-360o) or the left (90- 270o) side of the field.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \alpha H \begin{bmatrix} x \\ y \\ w \end{bmatrix} \text{, } where\ homography\ H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \qquad \text{(Equation 1)}$$

After that, the 2D field projections of the left-right (LR) pair of penalty taker’s shoulders, hips, and non-kicking foot (big toe and small toe) were calculated. All body parts' orientations could point to the left or right half, based on the angle system presented by Sangüesa et al. (2019; 2020a; 2020b). Based on the 2-D projection, LR-side Booleans ($LR_{Sh,}$ $LR_{Hi,}$ LR_{SF}: penalty kicker's shoulder, hips, and non-kicking foot orientation, respectively), angles (α_{Sh} $\alpha_{Hi,}$ α_{SF}: penalty kicker's shoulder, hip, and non-kicking foot, respectively) and confidences

(C_{Sh}, C_{Hi}, C_{SF}: confidence score of the orientation of penalty kicker's shoulder, hips and non-kicking foot, respectively) were obtained. The corresponding confidences are the average of OpenPose's player toes, shoulders, and hips confidences respectively. Figure 3 shows the output of the OpenPose on which the key biometric body parts of an individual are detected, illustrating the estimation process of orientation.

Figure 3. Penalty kick viewing angle, frames analyzed and process of pose estimation



Legend: The upper image corresponds to the moment when the penalty taker starts the run-up approaching the ball, and the down one corresponds to the moment when he touches the ball.The coordinates belonging to shoulders, hips, and non-kicking foot of penalty taker are mapped in a 2D field. LR-side Booleans (LR_{Sh}, LR_{Hi}, LR_{SF}: penalty kicker's shoulder, hips, and non-kicking foot orientation, respectively), angles (α_{Sh} α_{Hi}, α_{SF}: penalty kicker's shoulder, hip, and non-kicking foot, respectively) and confidences (C_{Sh}, C_{Hi}, C_{SF}: confidence score of the orientation of penalty kicker's shoulder, hips and non-kicking foot, respectively). Coordinates belonging to the neck and hip of the goalkeeper at these two moments are mapped to a vector and the angle is calculated (α_{GK}). D_{GKR} and D_{GKL}: the movement distance of right and left foot; C_{GK}: represent confidence score.

Anticipation movement of the goalkeeper

The anticipation movement of the goalkeeper in the penalty kick was defined here as to how far the goalkeeper moves between (1) the moment when the penalty taker starts the run-up approaching the ball and the (2) moment the penalty taker first touches the ball. In detail, the line formed by the connection between the goalkeeper's neck and the middle of the hip was used to depict the position status of the goalkeeper in these two moments. Furthermore, the Angle (αGK) between the two lines drawn from the two moments measures the anticipation movement of the goalkeeper. The confidence level for this measure (C_{GK}) was calculated by the average confidence scores of the neck and middle of the hip. This process is illustrated in figure 3.

The movement distance of the goalkeeper's left and right foot was also used to measure the anticipation movement. Left and right ankles were used to represent the left and right feet respectively, moreover, coordinate information together with metric Euclidean distance were used to depict the movement distance of the goalkeeper's feet, as shown in figure 3.

Ball speed

Ball speed was determined with the open-source software program Kinovea Motion Analysis (v0.8.15, Kinovea, France). This software has already been used in various studies analyzing penalty kicks (Hunter et al., 2018; Makaruk et al., 2020).

Notational analysis

A previously developed and validated observational system (OSPAF) for penalty analysis in elite football was also used in the present study (Pinheiro et al., 2021). The protocols for the use of observational systems were adopted (Aranda et al., 2019; Fernandes et al., 2019; Lames & Hansen, 2001). All the observable behaviors recorded are described in Pinheiro et al. (2021).

5.4.3 Statistical Analysis

For descriptive analysis, mean and standard deviation were used. Shapiro Wilk test was performed to verify data normality. The association level between the OSPAF variables with the penalty taker and goalkeeper strategy was determined with the use of the chi-square ($\chi 2$) test. The effect size was determined using the Cramer's V, and classified as weak (ES ≤ 0.2), moderate (0.2 < ES ≤ 0.6), and strong (ES > 0.6) (Cohen, 1988). The association level between the OpenPose variables with the penalty taker and goalkeeper strategy was determined with the use of the point-biserial correlation. Retest reliability was utilized to check the repeated measurements of OpenPose (Vilagut, 2014). Test-retest reliability coefficients (also called coefficients of stability) vary between 0 and 1, where 1: perfect reliability, ≥ 0.9: excellent reliability, ≥ 0.8 < 0.9: good reliability, ≥ 0.7 < 0.8: acceptable reliability, ≥ 0.6 < 0.7: questionable reliability, ≥ 0.5 < 0.6: poor reliability, < 0.5: unacceptable reliability, 0: no reliability (Lindstrom, 2010; Vogt, 2005). To identify which variables would be able to predict the penalty takers and goalkeeper strategy, logistic regression (*enter method*) analyses were performed. Dimensions and categories of OSPAF were coded in Lince software (Figure 1) (Gabin et al., 2012; Soto et al., 2019). Kappa levels of the OSPAF were 0.90 and 0.86 - intra and inter reliability (Pinheiro et al., 2021). The interpretation of this coefficient was adopted as follows: $\kappa > 0.8$ very good; $0.6 < \kappa < 0.8$ good; $0.4 < \kappa < 0.6$ moderate; $0.2 < \kappa < 0.4$ fair; $\kappa < 0.2$ poor (O'Donoghue, 2010; Altman, 1991). The level of statistical significance adopted was $\alpha = 0.05$, with a 95% confidence interval. All data were analyzed using JASP software (Team, 2020; JASP Version 0.14; Computer software).

6 RESULTS

6.1 Paper: Design and validation of an observational system

Pilot study

The first version of the new penalty kick analysis system was created, based on the collection of several variables from previous studies (Hughes and Wells, 2002; Timmis et al., 2014, 2018; Noël et al., 2015; Almeida et al., 2016; Comas et al., 2018). Characteristics were selected that are likely to distinguish the profile of successful or unsuccessful penalty kicks and strategies. Furthermore, contextual factors were included (e.g., location of the match; the result of the match at the time of the penalty kick; penalty kick during the normal time or extra time). This first step enabled the development of the first round of content validity of the proposed observational system, with the variables and their definitions.

Qualitative feedback was also gathered from the professionals about the conduct of the online questionnaire. The questionnaire was indicated by 65% of the experts as being very long and complex. In this way, the design of the questionnaire was adjusted for the main study, by repositioning the descriptions of the variables close to the answer box. The use of an online survey presented no problems and was considered a suitable tool for further expert participation. All the 27 variables proposed in the pilot study have been pointed out by the experts as relevant for the analysis of penalties in football. Aiken's *V* for Agreement ($p < 0.05$) ranged from 0.66 to 0.89 (cut-off: 0.64); for Adequacy ($p < 0.05$) from 0.64 to 0.83 (cut-off: 0.64) and for Univocity ($p < 0.05$) from 0.89 to 1.0 (cut-off: 0.75). Moreover, the experts suggested the variables ball speed, match importance, and initial goalkeeper posture, which were added to the observational system.

Main study

The results regarding content validity in the main study were obtained by calculating Aikens' *V* and the 95% confidence interval. Data are shown in 3. All observable variables of the OSPAF, presented in table 3, showed Aiken values above the cut-off value ($p < 0.05$).

Table 3. Aiken's V values of the observational system for penalty analysis in football (OSPAF)

	Agreement (5-scale) - 95%				Adequacy (5-scale) - 95%				Univocity (2-scale) - 95%			
Variable definition	**M**	***V***	**L**	V_p	**M**	***V***	**L**	V_p	**M**	***V***	**L**	V_p
Run up speed	4.10	0.77	0.67	0.71	4.00	0.75	0.65	0.73	0.90	0.90	0.70	0.92
Run up fluency	4.38	0.85	0.75	0.88	4.14	0.79	0.68	0.83	1	1	0.84	1
Run up approach angle	3.86	0.71	0.61	0.69	3.76	0.69	0.58	0.81	0.76	0.76	0.52	1
Number of steps	3.76	0.69	0.58	0.74	3.81	0.70	0.59	0.82	0.95	0.95	0.77	0.96
Kicking technique	4.38	0.85	0.75	0.78	4.29	0.82	0.72	0.83	1	1	0.84	1
Perceived ball speed	4.19	0.80	0.70	-	4.10	0.77	0.67	-	0.81	0.81	0.58	-
Foot used to kick	3.67	0.67	0.56	0.71	3.86	0.71	0.61	0.67	1	1	0.84	0.89
Non-kicking foot orientation	4.24	0.81	0.71	0.86	4.00	0.75	0.65	0.82	0.90	0.90	0.70	1
Penalty taker gaze behavior	4.38	0.85	0.75	0.89	4.43	0.86	0.76	0.81	0.81	0.81	0.58	1
Goalkeeper (GK) initial posture	3.90	0.73	0.62	-	3.86	0.71	0.61	-	0.95	0.95	0.77	-

Deception by the penalty taker	4.19	0.80	0.70	0.84	4.24	0.81	0.71	0.83	0.86	0.86	0.64	1
Anticipation movement of the goalkeeper at the ball contact point	4.29	0.82	0.72	0.83	4.24	0.81	0.71	0.78	0.90	0.90	0.70	0.95
Goalkeeper tactical action	4.14	0.79	0.68	0.72	4.24	0.81	0.71	0.83	0.90	0.90	0.70	0.89
Goalkeeper performance	4.76	0.94	0.87	0.77	4.71	0.93	0.85	0.69	1	1	0.84	0.90
Moment of the match	4.48	0.87	0.78	0.84	4.14	0.79	0.68	0.75	1	1	0.84	1
Location of the match (kicker point of view)	4.00	0.75	0.65	0.81	3.95	0.74	0.62	0.81	1	1	0.84	1
Momentary result (kicker point of view)	4.29	0.82	0.72	0.87	4.19	0.80	0.70	0.79	1	1	0.84	1
Momentary result (goalkeeper point of view)	4.24	0.81	0.71	0.83	4.10	0.77	0.67	0.80	0.95	0.95	0.77	1
Match Importance	4.38	0.85	0.75	-	4.14	0.79	0.68	-	0.90	0.90	0.70	-
Penalty kick direction	4.29	0.82	0.72	0.77	4.33	0.83	0.74	0.69	0.95	0.95	0.77	1
Penalty kick height	4.29	0.82	0.72	0.85	4.29	0.82	0.72	0.78	0.95	0.95	0.77	1
Penalty kick outcome	4.67	0.92	0.84	0.84	4,71	0.93	0.85	0.69	1	1	0.84	1
Penalty taker strategy	4.81	0.95	0.88	0.88	4.57	0.89	0.81	0.81	0.95	0.95	0.77	1
Goalkeeper strategy	4.52	0.88	0.79	0.87	4.33	0.83	0.74	0.79	0.90	0.90	0.70	1

Legend: M = median; *V* = Aiken´s *V* value of the main study; L = lower 95% confidence interval limit; V_p = Aiken´s *V* value of the pilot study. Cut-off for 5-scale ($p < 0.05$) = $V > 0.64$ (n=20); for 2-scale ($p<0.05$) = $V > 0.75$ (n=20); Agreement: degree of general acceptance of criteria; Adequacy: level of pertinence and importance for criteria and categories specific purpose; Univocity: clarity domain of a definition.

Unlike in the pilot study, the following variables have not achieved the minimum values for 5-scale ($p < 0.05$; $V < 0.64$, $n = 20$) and/or for 2-scale ($p < 0.05$; $V < 0.75$; $n = 20$) to be included in the final version of the OSPAF and were excluded. The table 4 shows the excluded variables.

Table 4. Variables, definitions, attribute levels and Aiken's V values of the excluded variables

Variable	**Definition**	**Attribute levels**	V_{AG}	V_{AD}	V_U
Run up type	The angle of the penalty kicker's run with the ball	Frontal or diagonal	0.52	0.54	0.72
Run up length	How long is the run-up of the penalty kicker	Short run-up or Long run-up	0.44	0.44	0.65
Swing behavior of the kicking leg	Kick leg balancing profile of the penalty kicker	Normal or delayed	0.78	0.73	0.60
Position of the arm opposite the kicking leg	Position of the penalty kicker's arm during the kick	Low Abduction, Perpendicular to the player's trunk or High Abduction	0.65	0.51	0.80
Preparation time	Time after the referee's whistle	> 3 seconds, < 3 seconds	0.50	0.53	0.90
Distraction by the goalkeeper	Indication if the goalkeeper has done any action to distract the kicker at any time during the penalty kick	Yes or No	0.81	0.76	0.70
Presence of advertisement behind the goal	Indication of if there are advertising boards behind the goal	Yes or No	0.39	0.44	1

* Results might be interested for studies with thematic focus where these variables play a role. Legend: V_{AG} = Aiken´s *V* value of agreement dimension; V_{AD} = Aiken´s *V* value of adequacy dimension; V_{U} = Aiken´s *V* value of univocity dimension.

The following table 5 presents the final version of the OSPAF and the operational definitions.

Table 5. Final version of the OSPAF

Variables	Definition	Attribute Levels
Run up speed	Running speed of the penalty kicker towards the ball	Fast or Slow
Run up fluency	Characteristic of the penalty kicker's run during the approach of the ball, with or without pauses.	Continuous Running or Running with pauses
Run up approach angle	Penalty kicker's running angle to the ball.	Frontal or diagonal
Number of steps	Number of steps of the penalty kicker until contact with the ball	1-3; 3-5 or +5
Kicking technique	The technique used by the penalty kicker to kick the ball	Side foot kick or Instep kick
Perceived ball speed	How hard is the ball kicked?	Powerful shot or Placed shot
Foot used to kick	Foot used by the penalty kicker to kick the ball	Right or Left
Non-kicking foot orientation	Spatial orientation of the penalty kicker's support foot	Same orientation as the final direction of the kick; or Different orientation as the final direction of the kick
Penalty taker gaze behavior	Gaze behavior of the kicker during the approach run.	Gaze at the ball or Not at the ball
Goalkeeper (GK) initial posture	Position of the body segments.	Arms raised; Arms down or Arms extended in a position perpendicular to the goalkeeper 's trunk
Deception by the penalty taker	Indication if the kicker has done any action to distract the goalkeeper during his or her run-up	Yes or No
Anticipation movement of the goalkeeper at the ball contact point	Action is performed parallel to the kicker's kick action.	No Movement; Partial movement (at least 1 body segment moved); or Full movement (>1 body segment moved)
Goalkeeper tactical action	General evaluation of the way the goalkeeper acted during the penalty shoot-out, to the anticipatory aspect	Try to guess the location of the shot; or Awaiting the penalty taker action
Goalkeeper performance	Evaluation of the goalkeeper's performance according to his movement and contact with the ball	0: GK made any final movement to the side of the goal opposite to the final ball location; 1: GK did not move from the center of the goal; 2: GK made a movement in the correct direction but did not dive and failed to make contact with the ball; 3: GK dived in the correct direction but failed to make contact with the ball; 4: GK dived in the correct direction and contacted the ball without saving it; or 5: GK successfully saved the kick
Moment of the match	Time of the match when the penalty will be taken	First half; Second Half or Extra time or Shoot out
Location of the match (kicker point of view)	Indication if the penalty kicker is from the home team, visitor, or if he plays on a neutral field.	Home, Neutral or Away
Momentary result (kicker point of view)	Result of the match (for the penalty kicker) at the moment the penalty was marked.	Winning, Drawing or Losing
Momentary result (GK point of view)	Result of the match (for the Goalkeeper) at the moment the penalty was marked.	Winning, Drawing or Losing
Match importance	Level of importance of the match for the team	Championship final match; Decisive knockout match; Group stage match; Early season game; Match in final stages of the season

Penalty kick direction	The direction of the ball on goal	Left; Center or Right
Penalty kick height	Height of the ball on goal	Upper; Center or Down
Penalty kick outcome	Result of the penalty kick	Goal; Saved by goalkeeper or Shot misses goal (wide, over or post)
Penalty taker strategy	Overall strategy perceived by the observer (Kuhn, 1988)	Goalkeeper Dependent; Unclear or Goalkeeper independent
Goalkeeper strategy	Overall strategy perceived by the observer (Kuhn, 1988)	Kicker Independent; Unclear or Kicker dependent

The table 6 shows the Kappa values for each of the variables of the observational system for penalty analysis in football (OSPAF). From the 24 items validated through the Aiken coefficient, 19 obtained Kappa values above 0.80 for intra-reliability, and the other 5 items values above 0.75. Regarding the inter-reliability, 17 items presented Kappa values above 0.80, the 7 remaining items had values above 0.70. Thus, Kappa values indicate a high level of reliability of the proposed new penalty kick analysis system.

Table 6. Cohen's Kappa for the OSPAF variables.

Variable	K (Intra-observer)	K (Inter-observers)
Run up speed	0.81	0.76
Run up fluency	1.00	0.80
Run up approach angle	0.85	0.80
Number of steps	0.89	0.82
Kicking technique	0.91	0.82
Perceived ball speed	0.84	0.79
Foot used to kick	1.00	1.00
Non kicking foot orientation	0.75	0.81
Penalty taker gaze behavior	0.78	0.78
Goalkeeper initial posture	0.84	0.84
Deception by the penalty taker	0.92	0.81
Anticipation movement of the GK at the ball contact point	0.86	0.78
Goalkeeper tactical action	0.77	0.70
Goalkeeper performance	0.86	0.83
Moment of the match	1.00	1.00
Location of the match (kicker point of view)	1.00	1.00
Momentary result (kicker point of view)	1.00	1.00
Momentary result (GK point of view)	1.00	1.00
Match Importance	1.00	1.00
Penalty kick direction	1.00	1.00
Penalty kick height	0.95	0.90
Penalty kick outcome	1.00	1.00
Penalty taker strategy	0.75	0.73
Goalkeeper strategy	0.79	0.75
Median value for Kappa	**0.90**	**0.86**

6.2 Paper 2: Empirical study based on the observational system

The OSPAF kappa values showed very good strength of agreement (0.95 and 0.92). The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, and the deception performed by the penalty taker were related to the goalkeeper dependent strategy ($\chi2$ (145) = 130.596, 86.5% correct classifications, $p < .001$). The model correctly classified 86.5% of cases. The goalkeeper tactical action guess was related to the goalkeeper independent strategy ($\chi2$ (148) = 137.680, 87.5% correct classifications, $p < .001$).

Based on this empirical study, OSPAF has proven to be a valid and reliable tool to distinguish between the strategies of the penalty taker and the goalkeeper in the penalty kick in elite football. The precise operational settings in the system allow it to be reproduced reliably. The results of this preliminary study demonstrate OSPAF as an instrument that integrates the main variables for penalty analysis in high-level football. The OSPAF may serve as a standard tool for observational investigations of penalties in football to make the results from different studies more comparable.

6.3 Paper 3: Body orientation analysis in penalty kicks

Descriptive data of all the OpenPose and OSPAF variables analyzed are presented as supplementary data (please see supplementary material: graph 1 and 2).

OpenPose confidence score and retest reliability

The mean confidence score of OpenPose measures was 0.80 ± 0.14. The confidence score per variable is shown in the Table 7.

Table 7. OpenPose confidence score

Player	Body orientation angle	Confidence score
Penalty taker	Non-kick foot orientation	0.51
	Shoulders	0.87
	Hips	0.85
Goalkeeper	Anticipation	0.87
	Left foot	0.84
	Right foot	0.83

Test-retest reliability values are presented in table 8.

Table 8. Test-retest reliability per variable

Player	Body orientation angle	*r*
Penalty taker	Non-kick foot orientation	0.924
	Shoulders	0.998
	Hips	0.991
Goalkeeper	Anticipation	0.998
	Left foot	0.953
	Right foot	0.961

p<0.05*

Influence variables on goalkeeper strategy

The association between all OpenPose and OSPAF variables with the goalkeeper's strategy was analyzed. Table 9 presents only the variables that presented association and the respective values.

Table 9. Association between OSPAF and OpenPose variables with the goalkeeper strategy

	OSPAF Variables	**χ2**	**p**	**Cramer´s V**
Goalkeeper strategy	Run up speed	4.875	< 0.05	0.354
	GK Tactical action	26.542	< 0.05	0.825
	OpenPose Variable	r_{pb}	**p**	
	Goalkeeper anticipation	0.959	< 0.05	

A logistic regression (*enter method*) was performed to investigate the relation between the goalkeeper´s tactical action and run up speed on the likelihood of the goalkeeper strategy. The logistic regression model was statistically significant, χ^2 (36) = 28.592, p < .001. The model correctly classified 84.6% of cases. The goalkeeper´s tactical action (awaiting) and run speed (slow) was related to a kicker dependent strategy. While including the correlated OpenPose variable (goalkeeper anticipation) in the model (χ^2 (35) = 49.648, p < .001), the accuracy is increased to 97.0%. Therefore, lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting) and run up speed (slow) were associated with a kicker dependent strategy.

Influence variables on penalty taker strategy

The association between all OSPAF and OpenPose variables with the penalty taker's strategy was analyzed. Table 10 presents only the variables that presented association and the respective values.

Table 10. Association between OSPAF and OpenPose variables with the penalty taker strategy

	OSPAF Variables	**χ2**	**p**	**Cramer´s V**
Penalty taker strategy	Run up speed	2.300	< 0.05	0.243
	Run up fluency	5.512	< 0.05	0.376
	Gaze behavior	22.224	< 0.05	0.755
	Deception	8.770	< 0.05	0.474
	OpenPose Variable	r_{pb}	**p**	
	Ball speed	0.927	< 0.05	

A logistic regression (*enter method*) was performed to investigate the relation between the correlated OSPAF variables (run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker and ball speed) on the likelihood of the goalkeeper dependent strategy. The logistic regression model was statistically significant, χ^2 (33) = 24.819, p < .001. The model correctly classified 97.1% of cases. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, the deception performed by the penalty taker and lower ball speed were related to a goalkeeper dependent strategy.

7 DISCUSSION

Three studies with different aims were conducted to approve and use an observational system for in-match penalty kick analysis: (1) study one designed and validated the observational system, including the investigation of the influence of the video footage; (2) study two used the expert-validated observational system for penalty kick analysis in football (OSPAF) to identify the relationship between a set of observable variables and the players’ strategy; and (3) the study three tested an innovative approach using the OSPAF, applying technological methods to analyze its variables, such as computer techniques for body pose estimation, and machine learning-based video analysis.

As the objectives of study 1 are oriented towards the construction and validation of an observation instrument, the results refer to the quality control of the data, focusing on the statistical procedures adopted, such as Aiken's V values, intra, and inter-observer agreement. The Aiken's V value, measured in the dimensions Agreement, Univocity and Adequacy, indicated that content validity was achieved (For 5-scale: $p < 0.05$ and $V > 0.64$; for 2-scale: $p < 0.05$ and $V > 0.75$). All variables included in OSPAF presented Aiken's V value above the cut-off threshold. Intra and inter-rater reliability showed very good strength of agreement (O'Donoghue, 2010).

The methodological design in study 1 containing a three-study concept (pilot, main and video study) made it possible to have a practical approach to the proposed instrument through the pilot study with high-level football coaches. The main study, including three different dimensions (i.e., agreement, adequacy, and univocity), ensured that only variables with a high level of concordance, clarity, and relevance were included in the final format of the instrument. The video study suggested a novelty by including viewing angle analysis in the development of the observational system. To the best of our knowledge, there is no other study yet that has included this type of specification for notational studies applied to penalty kick analysis in elite football. The description and standardization of the viewing angles and video quality allow the reproducibility of the instrument and reduce human error. Results indicate that for an optimum penalty kick analysis, a combination of at least three different viewing angles of the same penalty is recommended. The pitch-level viewing angle behind the penalty taker (d), aerial viewing angle behind the penalty taker (c), and pitch-level viewing angle behind the goalkeeper (e) are the best viewing angles for observation analysis, according to a panel of 20 experts in sports science. The change of viewing angles plays an important role in notational analysis, as it may influence the perception of the observer.

Once the OSPAF was completely validated, the next step was to use the system for penalty kicks analysis in competitive scenarios. A total of 150 penalty kicks from the main European football leagues were analyzed. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, and the deception performed by the penalty taker were related to the goalkeeper dependent strategy ($\chi 2$ (145) = 130.596, 86.5% correct classifications, $p < .001$). The goalkeeper tactical action guess was related to the goalkeeper independent strategy ($\chi 2$ (148) = 137.680, 87.5% correct classifications, $p < .001$). The results corroborate previous studies by showing that the run-up pattern is associated with the strategy of the penalty takers (Noël et al., 2015; Kuhn, 1988). Furthermore, the gaze behavior of the penalty taker endorses previous research (Noël and van der Kamp, 2012). One characteristic differentiating the penalty takers' profiles is the space-temporal pattern of gaze (Noël and van der Kamp, 2012). Penalty takers who use a keeper-dependent strategy spend more time looking at the goalkeeper throughout the run-up and kick execution than penalty takers who use a keeper-independent strategy. The deception strategies adopted by the penalty takers also corroborate previous research (Lopes et al., 2014). Players who use this strategy try to instigate the goalkeepers to jump to one side before the kicker shoots. Lopes et al. (2014) indicated that trying to deceive the goalkeeper, penalty takers can modify the predictive value of the local body kinematics to some extent, most particularly early in the approach.

Study 3 presents a novel approach adopting OpenPose measures combined with notional analysis (i.e., OSPAF) to analyze penalty kicks. The mean confidence score of the OpenPose variables was 0.80, and test-retest reliability showed excellent reliability (Lindstrom, 2010; Vogt, 2005). The selected body orientation angle (goalkeeper anticipation) measured through OpenPose, correlated significantly with the goalkeeper strategy. The prediction model of the goalkeeper's strategy had its accuracy increased when the variable goalkeeper anticipation was included. This finding corroborates the applicability of OpenPose to obtain the body orientation of professional football players during matches (Sangüesa et al., 2019).

Lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and the run-up speed of the penalty taker (slow) were associated with a kicker-dependent strategy of the goalkeeper. From a behavioral perspective, the present findings corroborate this dyadic interaction between the players in a penalty kick, as results showed that the goalkeeper strategy is influenced by the run-up speed of the penalty taker. In real competitions, penalty kicks are an interaction process, and the observable performance is rather the emergent result of this interaction process than the display of skills and abilities of the two parties (Lames and McGarry, 2007).

Even though the selected body angles measured through OpenPose did not associate significantly with the shooter strategy, the variables measured by OSPAF (i.e., run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed) were able to classify 97.1% of the penalty taker strategy correctly. These findings also corroborate the results of study 2 present in this thesis. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, the deception performed by the penalty taker, and lower ball speed were related to a goalkeeper dependent strategy. A possible explanation about the body angles measured through OpenPose could be that the biomechanical patterns of approaching the ball during the kick may vary from player to player, regardless of the strategy adopted. Previous research has shown that kicking from an approach angle of 45° and 60° may alter aspects of kick technique, such as enhancing pelvic rotation and thigh abduction of the kicking leg at impact (Scurr and Hall, 2009).

Among the limitations of this chapter, one can cite that the study was conducted via an extensive questionnaire. To validate the content of the observational system, the experts needed to evaluate it item by item. This required a larger workload. Another limitation was not to use a larger sample (e.g., entire season) in study 3, as it could bring practical applications and be more representative. Nevertheless, the OSPAF enables the differentiation of the technical and tactical behavior of the goalkeeper and penalty taker. The present instrument is a comprehensive observational system, which contains the most relevant variables for penalty kick analysis validated by experts. The inclusion of observable variables about the penalty taker actions, goalkeeper, context, and outcomes, makes the proposed model more complete than most of the others proposed previously. The current work presents an innovative approach to analyzing penalty kicks in football, combining notational analysis with OpenPose. Its integration with video specification allows this model to be used as a coaching tool to assess players' orientation under different penalty kicks, improving sports preparation against upcoming opponents.

8 CONCLUSION

The OSPAF evidenced content validity, inter-and intra-reliability for analyzing penalty kicks in football through a gold standard methodology for instrument validation. The present work concludes that the final instrument is adequate and consistent for analyzing successful and non-successful penalty kick patterns. Besides, an innovative approach in applying Openpose measures integrated with notational analysis to investigate the factors influencing the players' strategy in penalty kicks was tested. Results showed the applicability of OpenPose for in-match penalty kick analysis and an improvement in the prediction of the goalkeeper strategy using a body orientation variable (anticipation movement of the goalkeeper).

9 INTRODUCTION

9.1 Background

The use of electronic performance-tracking systems (ETPS) to obtain spatiotemporal data in football is becoming a generalized practice for performance analysis (Carrilho et al., 2020). Wearable EPTS devices, which include measurement technologies such as a Global (GPS) or Local (LPS) Positioning System, accelerometer, gyroscope, and magnetometer, have become prevalent in football training to control and improve performance (Medina et al., 2017) and to track match-play performance (IFAB, 2016).

Football players are subject to different types and amounts of load during training sessions (Pinheiro et al., 2018), intending to optimize their performance (Graham et al., 2018) in competition and reducing, as far as possible, the risk of injury (Gabbett, 2016). Therefore, the evaluation of the training load in general, and precisely that of the underlying training tasks, is indispensable both in terms of optimizing the players' conditional performances (avoiding under- or over-training through conditions that are very different from those of matches) and in preventing overtraining and injuries (Sangnier et al., 2018).

During the last few years, integrated physical training (athletic training with the ball in specific football game situations, including opponents and partners) has been used as an alternative to athletic training without the ball. This type of training relies in part on small-sided games (SSG) (Dellal et al., 2012; Hill-Haas et a., 2009; Hill-Haas et al., 2011; Reilly, 2004). Small-sided games (SSGs) are played on reduced pitch areas, often using modified rules and involving fewer players than traditional football. These games are less structured than traditional fitness training methods but are very popular training drills for players of all ages and levels (Hill-Haas et al., 2011).

9.2 Scientific problem

Although the external and internal loads of different SSG conditions are among the most studied performance parameters, some limitations remain in the international literature (Folgado et al., 2019; Clemente et al., 2019). The heterogeneity that exists among designs causes a lack of consistency about the design of SSG (Bujalance-Moreno et al., 2019), the level of ability of the players, the coach encouragement, and the rules (e.g., number and type of goals, number of touches in the ball) used by researchers (Praça et al., 2021). Such inconsistencies make it difficult to compare studies. Additionally, the main body of research in this area is conducted with young players (Bujalance-Moreno et al., 2019; Folgado et al., 2019), while studies conducted with senior players, especially with top-level professionals players, are less known (Clemente et al., 2019). Due to the multiple conditional demands of SSGs, there is no consensus on what variables, or game formats, better represent the game demands (Zurutuza et al., 2020).

SSGs does not precisely reproduce the real game context (Casamichana and Castellano, 2015; Halouani et al., 2014). In addition, these game-based tasks could exceed speed changes and reduced demands in high-intensity movements compared to official matches (Gómez-Carmona et al., 2018), potentially increasing injury risk (Rampinini et al., 2007). For this reason, previous research hypothesized that the demands of the SSGs are not related to competing demands, and the modification of the objective in game-based tasks will influence their requirements (Gómez-Carmona et al., 2018).

Given its importance in planning effective training loads, especially during the pre-season, more research is needed to clarify whether the manipulation of different SSG formats causes similar responses to competitive scenarios in elite football. Understanding the possible differences between SSGs and friendly matches would be helpful for coaches and practitioners dealing with training prescription in elite football, assisting with pre-season programming, better coping with training load management, and preventing muscular injuries (Buckthorpe et a., 2019).

10 OBJECTIVE

The present second chapter of this thesis is dedicated to additional research on performance analysis in elite football, aiming to investigate internal and external training load in different training formats and match-play demands.

10.1 Specific objective: Paper 4- Small-sided games *vs* match-play demand: a case study

This pilot study aimed to investigate the differences in external and internal load during pre-season training sessions with different SSGs and a friendly match in top-class professional football players.

11 PUBLICATION LIST

The following publication is presented in support of the chapter two of the thesis:

IV Pinheiro, G. de S., Quintão, R.C., Nascimento, V. B., Claudino, J. G., Alves, A. L., Costa, I.T., Costa, V.T. (2022). Small-sided games do not replicate all external and internal loads of a football match-play during pre-season: A case study. *International Journal of Sports Science & Coaching*. https://doi.org/10.1177/17479541211069935

12 MATERIALS AND METHODS

12.1 Ethical procedure

All participants signed the Free and Informed Consent Term. The anonymity of the participants was preserved throughout the process. The data were only involved in this study after the participants' agreement. It was obtained a consent letter from the club agreeing with the procedures. All research procedures were conducted according to the norms established by the National Health Council Resolution (466/2012) and the Declaration of Helsinki for human research. The project was approved by the Human Research Ethics Committee (485/10). All players were submitted to medical evaluations by the club's medical staff and presented adequate health status for professional football practice.

12.2 Subjects

An initial sample of 23 professional football players (who were already champions of the main competitions in Brazil, South America, Europe, and the

FIFA World Cup) was monitored in a competitive season across the national championship (38 matches). In an attempt to provide a representative profile for this study, we elected only to include players who were regular starters with ≥ 60% participation in the total matches of the season and who were not absent for more than 21 days due to injuries. Therefore, the final sample was composed of 9 male professional football players (25.11 ± 4.59 years, body mass 74.33 ± 8.3 kg, height 176.56 ± 7.94 cm). The group was composed of defenders, midfielders, and forwards. Goalkeepers (GK) were excluded from the analysis. These individuals were part of a first division team of Brazilian football, with professional experience in training and competitions of national and international level, recognized by the Brazilian Football Confederation (CBF) and South American Football Confederation (CONMEBOL).

12.3 Procedures

A descriptive comparative design was used to investigate possible differences between several pre-season sessions, including SSG formats and friendly match. The study was conducted over a full pre-season and lasted three weeks. Various studies used similar weekly training frequencies (Halouani et al., 2014). The following game formats: 4v4, 6v6, 7v7, 8v8, 10v10, and 14v14, were compared with one friendly match (FM) in terms of the activity profiles of the players. The different SSGs used in this study were similar to previous studies (Bujalance-Moreno et al., 2019; Hill-Haas et al., 2011; Owen et al., 2020; Sgrò et al., 2018). The pre-season training session's formats carried out with different SSGs are described in table 11.

Table 11. Description of the SSG´s

SSG	Pitch size	Area per player	SSG planning	Game description	Game purpose
4v4	32 x 20 m	80 m^2	4 sets x 7.5 min, 5 min of passive rest	Small SSG + GK	Free play
6v6	40 x 30 m	100 m^2	4 sets x 12.5 min, 5 min of passive rest	Small SSG + GK	Free play
7v7	52,5 x 68 m	255 m^2	4 sets x 10 min, 5 min of passive rest	Medium SSG + GK	Free play
8v8	35 x 40 m	87.5 m^2	4 sets x 7.5 min, 5 min of passive rest	Small SSG + GK	Ball possession

10v10	105 x 68 m	357 m^2	4 sets x 7.5 min, 5 min of passive rest	Large SSG + GK	Set piece training
10v10	105 x 68 m	357 m^2	4 sets x 7.5 min, 5 min of passive rest	Large SSG + GK	Free play
14v14	52,5 x 68 m	127.5 m^2	4 sets x 7.5 min, 5 min of passive rest	Small SSG + GK	Recreational training

Legend: SSG = small-sided game; GK = goalkeeper. The friendly match was played on a 105x68m pitch (area per player: 357m^2)

The pre-season sessions were performed in three consecutive weeks and were part of a regular training session. The 4v4 and 6v6 SSGs were performed on Tuesday and Thursday of the first week, respectively. The 7v7, 8v8, and 10v10 SSGs were performed on Tuesday, Thursday, and Saturday of the second week, respectively. The 10v10, 14v14 SSGs, and FM were performed on Tuesday, Thursday, and Saturday of the third week, respectively. We chose the same weekdays to minimize the influence of training loads over the weeks on players' physical responses; therefore, the recovery time was standardized. In the FM, the opponent was a professional football team, which competes in championships at a national level.

The players were familiarized with the use of these devices and with the SSG formats used. Coaches gave verbal encouragement to players during the training formats. Each team was composed of at least one defender, midfielder, and forward to allow teams to explore the physical and technical-tactical specificities of each playing position during the different SSGs (Praça et al., 2021). Considering that all players are part of a top-class professional team, we assumed that the homogeneity of the sample would not require an intentional team creation and opponent composition by the researchers. Therefore, the decision to create SSG´s teams and define opponents was coach-driven.

All SSGs were implemented immediately after a warm-up (15–30 min) containing preparatory activities such as moderate running, dynamic stretching, balance and agility exercises, and accelerations. This process helped to ensure similar conditions across all SSGs. The ball was always available by prompt replacement to maximize effective playing time, except in the 4v4 format, which was played in a particular arena surrounded by walls that continuously kept the ball in play. The number of touches to the ball was free for each player. Official FIFA-approved goals (7.32 × 2.44 m) were used. The offside rule was applied in all training formats, as well as in the match. On the rest periods, players were allowed to drink liquids ad libitum.

12.4 Data collection

The external load variables were obtained from portable GPS devices (GPSports SPI Pro X). According to the manufacturer, the GPS device has a sampling frequency of 15 Hz and includes a 100-Hz triaxial accelerometer. The manufacturer supplemented the GPS frequency to provide a sampling rate of 15 Hz.25 Each player used a special vest which enabled the device to be fitted to the upper part of his back. The use of the special GPS vest and the heart rate (HR) monitor has not influenced the player's performance, as the club uses these devices in training sessions and official matches. The GPS devices were activated 15 min before the beginning of each training session, following the manufacturer's instructions. The data were transferred to a computer and analyzed in the software Team AMS (R1 2016, Australia). These devices have been previously used in elite football.26 The registered variables were training session duration (min), average heart rate (bpm), total distance (m), distance covered per minute (m/min), the total number of accelerations >2.5 m/s^2, number of accelerations >2.5 m/s^2 per minute, average distance of accelerations (m), the average value of acceleration (m/s^2).

The monitored training sessions occurred in the morning, between 9:00 am, and 11:00 am, with sunny weather conditions and similar temperatures (24.3 ± 3.3°C), separated by an interval of 24 h between them. Most of the training formats were performed on the same field with natural grass. Only the 4v4 sided game was played on artificial grass. The official match was played in home condition, between 4:00 and 6:00 pm on natural grass with sunny conditions and similar temperatures (26.3 ± 2.7°C). Pauses in the training game formats were excluded from the analysis. In order to ensure the ecological validity of the data collected, the planning and execution of the training did not suffer interference from the researchers (Pinheiro et al., 2018). The athletes performed 2–3 strength training sessions per week during the study period. These sessions took place in the gym under the supervision of the club.

12.5 Statistical analysis

G*Power 3.1 software was used for the sample calculation. The hypothesized effect size assumed was 50%. A sample generalization power of up to 44% was achieved. A descriptive analysis of the data was performed, presenting the results in mean and standard deviation. Shapiro Wilk test was performed to verify data normality. As data presented normal distribution, parametric tests were performed. One-way ANOVA was performed to analyze the variance of all evaluated variables. The partial eta squared (η^2) has tested the effect size (ES) of ANOVA. The Ferguson's classification for the ES was used (Ferguson, 2009): no effect (ES < 0.04); minimum effect (0.04 < ES < 0.25); moderate effect (0.25 < ES < 0.64); and strong effect (ES > 0.64). The pairwise comparisons were tested with Cohen's d to analyze the effect size. The following classification to measure the magnitude of ES was used (Cohen, 1988): no effect (d < 0.41), minimum effect (0.41 < d < 1.15), moderate effect (1.15 < d < 2.70) and strong effect (d > 2.70). The level of statistical significance adopted was $\alpha = 0.05$. All data were analyzed using JASP software (Team, 2020; JASP Version 0.14; Computer software).

13 RESULTS

Descriptive statistics of duration, heart rate, total distance (m), distance per minute (m/min), the total number of accelerations >2.5 m/s^2, number of accelerations >2.5 m/s^2 per minute, average distance of accelerations (m) and average accelerations (m/s^2) are shown in supplementary material (please see Graph 3).

There was no difference in the training session duration ($p = 0.995$) across all training formats. Comparing all the pre-season training tasks, no differences were found just in the average accelerations (m/s^2) ($p = 0.128$). Moderate differences were found in number of accelerations >2.5 m/s^2 per minute ($\eta^2 = 0.396$, moderate effect; $p < 0.001$) and average distance of accelerations ($\eta^2 = 0.545$ moderate effect; $p < 0.001$). Strong differences were found in HR ($\eta^2 = 0.788$, large effect; $p < 0.001$), total distance ($\eta^2 = 0.797$, strong effect; $p < 0.001$), distance per minute ($\eta^2 = 0.775$ strong effect; $p < 0.001$), total number of accelerations >2.5 m/s^2 ($\eta^2 = 0.699$, strong effect; $p < 0.001$).

Tukey post hoc analysis revealed that the total number of accelerations >2.5 m/s^2 in the FM was higher than the 10v10 set piece ($d = 2.409$), 14v14 ($d = 2.229$) and 8v8 ($d = 2.049$). No difference was found comparing the match with the 10v10 sided game ($p = 0.265$), 4v4 ($p = 0.997$) and 7v7 ($p = 0.440$). Lower values were found comparing the FM with 6v6 ($d = 2.979$). The number of accelerations >2.5 m/s^2 per minute in the FM was higher than the 10v10 set piece ($d = 2.926$, strong effect). No differences were found comparing the FM with the 10v10 sided game ($p = 0.988$), the 14v14 ($p = 0.975$), the 4v4 ($p = 0.955$), the 6v6 ($p = 0.993$), the 7v7 ($p = 0.828$) and the 8v8 ($p = 0.998$). Average distance of accelerations in the FM was higher than the 10v10 set piece ($d = 2.188$), 14v14 ($d = 1.670$), 4v4 ($d = 3.046$) and 8v8 ($d = 1.443$). No difference was found comparing the FM with the 10v10 sided game ($p = 0.842$), the 6v6 ($p = 0.063$) and 7v7 ($p = 0.472$). HR in the FM was higher than the 10v10 set piece ($d = 3.734$), 10v10 sided game ($d = 2.071$), 14v14 ($d = 3.773$), 7v7 ($d = 2.853$) and the 8v8 ($d = 4.070$). No difference was found comparing the FM with the 4v4 ($p = 0.984$) and the 6v6 ($p = 0.985$). Total distance in the FM was higher than the 10v10 set piece ($d = 10.967$), 10v10 sided game ($d = 3.255$), 14v14 ($d = 8.839$), 4v4 ($d = 5.385$), 7v7 ($d = 2.116$) and the 8v8 ($d = 3.889$). No difference was found comparing the FM with the 6v6 ($p = 0.233$). Distance per minute in the FM was higher than the 10v10 set piece ($d = 8.680$), 10v10 sided game ($d = 2.090$),14v14 ($d = 4.390$), 4v4 ($d = 2.162$), 6v6 ($d = 3.521$), 7v7 ($d = 3.104$) and the 8v8 ($d = 1.998$). Further tables providing comparisons between the FM and the SSG´s are available as supplementary material.

14 DISCUSSION

The purpose of the present pilot study was to investigate the differences in external and internal load during pre-season training sessions carried out with different SSGs and a friendly match in top-class professional football players. Our main findings showed no differences only in the average value of accelerations (m/s^2) across different SSGs and the FM. However, there were differences in the total number of accelerations > 2.5 m/s^2 and number of accelerations > 2.5 m/s^2 per minute, average distance of accelerations (m), HR (bpm), total distance (m), and distance per minute (m/min). The current results shows that the external and internal loads differ across different SSGs in relation to the FM during the pre-season training sessions.

The publication of this paper provides evidence to reflect the widespread use of the SSGs in elite football practices and their practical implications. Notwithstanding the popularity of SSG's and their tactical benefits (Hill-Haas et al., 2011), it is unclear whether the common widely used SSG's reproduce the demands of competitive matches regarding physical performance (Casamichana et al., 2012) or if it may be effective to reduce the likelihood of muscular injuries (Buckthorpe et al., 2019). Evidence has emerged in recent years supporting the need to include sprint stimuli in players' preparation to minimize injury risk, in addition to the SSG practices. Furthermore, if players do not accumulate high chronic sprinting loads, they are more prone to "peaks" in acute sprinting loads during matches, which will ultimately lead them to an increased likelihood of muscular injury, especially hamstrings injuries (Buckthorpe et al., 2019). Therefore, players physical preparation with similar demands to matches is necessary (i.e., similar acceleration distances), as an inadequate exposure to specific training loads could lead to muscular injury.

The limitation of this study was to have a small sample size. However, even with such a sample size, a sample generalization of the data of up to 44% was achievable. Therefore, future studies are encouraged to analyze a larger number of players. Another limitation was not evaluating other constructs, such as perceived exertion, muscle soreness, and tactical demands. Nevertheless, these measures can complement the analysis of the SSG and the FM to understand better the differences. Despite this, the present study brings reflections and practical findings of the external and internal load of various pre-season training tasks carried out with different SSG formats and FM, to further develop our understanding of the training stimulus provided by football-specific training sessions during pre-season in professional football.

15 CONCLUSION

Pre-season training tasks with different SSG formats elicit different demands on elite professional football players regarding running, acceleration, and cardiovascular responses. Therefore, caution should be taken when selecting different SSGs, to improve the players' performance during the pre-season. This pilot study may help coaches learn whether proposed tasks underload, replicate, or overload the requirements of friendly match-play, which might be considered when scheduling training sessions.

16 GENERAL OUTLOOK

Performance analysis is primarily done to provide support for individual athletes as well as squads (O'Donoghue, 2010). Performance analysis is a discipline of sports science that presents a broad analytical approach. There has been an increase in research examining performance in set-plays (e.g., penalty kicks) in the last few years, providing valuable information on variables that influence their effectiveness (Sarmento et al., 2018). Meanwhile, there is growing research on comparing physical and physiological demands between different training formats and matches (Gómez-Carmona et al., 2018).

The present thesis (chapter one) provides a new instrument for the comprehensive analysis of penalty kicks in elite football. Standard statistical requirements for the validation of the system presented were met without exception. Furthermore, an empirical study was carried out to test the instrument with data from elite leagues and proved capable of identifying strategies adopted by players. An additional study used OSPAF applying computational techniques for body pose estimation and machine learning-based video analysis to analyze its variables. Chapter two provides practical applications on the comparability of the internal and external load of different training formats concerning the match-play demands in elite football. The current finds may help coaches learn whether proposed tasks underload, replicate, or overload the requirements of friendly match-play, which might be considered when planning training sessions.

The major challenge for future research remains on modeling strategies in the penalty kicks by direct evaluation using a large sample, therefore providing insights for training and competition. Secondly, investigate the physical, physiological, and psychophysiological demands of an entire season in elite football and compare them with the internal and external loads of different training formats to avoid the increased risk of muscular injuries.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychology Measurement*. 45, 131–142. doi: 10.1177/0013164485451012

Almeida, C. H., Volossovitch, A., & Duarte, R. (2016). Penalty kick outcomes in UEFA club competitions (2010 to 2015): The roles of situational, individual and performance factors. *International Journal of Performance Analysis in Sport*, 16(2), 508–522. https://doi.org/10. 1080/24748668.2016.11868905

Altman, D. G. (1991). Practical Statistics for Medical Research. Boca Raton, FL: CRC press.

Anguera, M. T., Blanco, A., and Losada, J. L. (2001). Observational Designs, a key issue in the observational methodology process (in Spanish). *Metodología de las Ciencias del Comportamiento*. 3, 135–161.

Aranda, R., González-Ródenas, J., López-Bondia, I., Aranda-Malavés, R., Tudela-Desantes, A., and Anguera, M. T. (2019). "REOFUT" as an observation tool for tactical analysis on offensive performance in soccer: mixed method perspective. *Frontiers in Psychology.* 10:1476. doi: 10.3389/fpsyg.2019.01476

Avugos, S., Azar, O.H., Sher, E., Gavish, N., Bar-Eli, M. (2020). The Right-Oriented Bias in Soccer Penalty Shootouts. *Journal of Behavioral and Experimental Economics*; 89:101546. https://doi.org/10.1016/j.socec.2020.101546

Bar-Eli, M., Azar, O. H., Ritov, I., Keidar-Levin, Y., & Schein, G. (2007). Action bias among elite soccer goalkeepers: The case of penalty kicks. *Journal of Economic Psychology*: 28(5), 606-621. doi: DOI 10.1016/j.joep.2006.12.001Bar

Bowtell, M., King, M. A., Pain, M. T. G. (2009). Analysis of the keeper-dependent strategy in the soccer penalty kick. *International Journal of Sports Science and Engineering*. 2009: 3, 93- 102.

Brewer, C., and Jones, L. (2002). A five-stage process for establishing contextually valid systematic observation instruments: the case of Rugby Union. *Sport Psychology*. 16, 138–159. doi: 10.1123/tsp.16.2.138

Brinkschulte, M., Furley, P., and Memmert, D. (2020). English football players are not as bad at kicking penalties as commonly assumed. *Scientific Reports*. 10:7027. doi: 10.1038/s41598-020-63889-6

Buckthorpe, M., Wright, S., Bruce-Low, S., Nanni, G., Sturdy, T., Gross, A. S., Bowen, L., Styles, B., Della Villa, S., Davison, M., & Gimpel, M. (2019). Recommendations for hamstring injury prevention in elite football: translating research into practice. *British journal of sports medicine*, 53(7), 449–456. https://doi.org/10.1136/bjsports-2018-099616

Bujalance-Moreno, P., Latorre-Román, P.Á., García-Pinillos, F. (2019). A systematic review on small-sided games in football players: acute and chronic adaptations. *Journal of Sports Sciences*; 37: 921.

Cao, Z., Simon, T., Wei, S., & Sheikh, Y. (2017). Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1302-1310.

Casamichana, D. & Castellano, J. (2015). The Relationship Between Intensity Indicators in Small-Sided Soccer Games. *Journal of Human Kinetics*,46(1) 119-128. https://doi.org/10.1515/hukin-2015-0040

Casamichana, D., Castellano, J., & Castagna, C. (2012). Comparing the physical demands of friendly matches and small-sided games in semiprofessional soccer players. *Journal of strength and conditioning research*, *26*(3), 837–843. https://doi.org/10.1519/JSC.0b013e31822a61cf

Carrilho, D., Santos Couceiro, M., Brito, J., Figueiredo, P., Lopes, R. J., & Araújo, D. (2020). Using Optical Tracking System Data to Measure Team Synergic Behavior: Synchronization of Player-Ball-Goal Angles in a Football Match. *Sensors*, *20*(17), 4990. MDPI AG. http://dx.doi.org/10.3390/s20174990

Chiappori, P. A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: the case of penalty kicks in soccer. *American Economic Review*. 92:4. doi: 10.1257/00028280260344678

Clemente, F. M., Sarmento, H., Rabbani, A., Van Der Linden, C., Kargarfard, M., & Costa, I. T. (2019). Variations of external load variables between medium- and large-sided soccer games in professional players. *Research in sports medicine* (Print), 27(1), 50–59. https://doi.org/10.1080/15438627.2018.1511560

Cobb, N. M., Unnithan, V., and McRobert, A. P. (2018). The validity, objectivity, and reliability of a soccer-specific behaviour measurement tool. *Science and Medicine in Football*. 2, 1–7. doi: 10.1080/24733938.2017.1423176

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2. Auflage). Hillsdale, NJ: Erlbaum.

Comas, G. N., Bardavío, J. S., Nuviala, R., and Miguel, D. F. (2018). Design of an observational instrument for the assessment of penalties in football and analysis of the results obtained (In Spanish). *Journal of Sport Psychology*. 27, 189–199.

Dellal, A., Chamari, K., Pintus, A., Girard, O., Cotte, T., & Keller, D. (2008). Heart rate responses during small-sided games and short intermittent running training in elite soccer players: a comparative study. *Journal of strength and conditioning research*, 22(5), 1449–1457. https://doi.org/10.1519/JSC.0b013e31817398c6

Dalton, K., Guillon, M., and Naroo, S. A. (2015). An analysis of penalty kicks in elite football post 1997. *International Journal of Sports Sciences and Coaching*. 10, 815–827. doi: 10.1260/1747-9541.10.5.815

Dalkey, N., and Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*. 9, 351–515. doi: 10.1287/mnsc.9.3.458

Dayé, C. (2018). How to train your oracle: the Delphi method and its turbulent youth in operations research and the policy sciences. Soc. Stud. Sci. 48, 846–868. doi: 10.1177/0306312718798497

Dicks, M., Button, C., and Davids, K. (2010). Availability of advance visual information constrains association-football goalkeeping performance during penalty kicks. *Perception* 39, 1111–1124. doi: 10.1068/p6442

Dicks, M., Davids, K., and Button, C. (2009). Representative task designs for the study of perception and action in sport. *International Journal of Sport Psychology.* 40, 506–524.

Ellis, L. and Ward, P. (2021). The effect of a high-pressure protocol on penalty shooting performance, psychological, and psychophysiological response in professional football: A mixed methods study, *Journal of Sports Sciences*, DOI: 10.1080/02640414.2021.1957344

Fariña, R. A., Fábrica, G., Tambusso, P. S., Alonso, R. (2013). Taking the goalkeeper's side in association football penalty kicks. *International Journal of Performance Analysis in Sport,* 13(1), 96–109. https://doi.org/10.1080/24748668.2013.11868634

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532–538. https://doi.org/10.1037/a0015808

Fernandes, T., Camerino, O., Garganta, J., Pereira, R., and Barreira, D. (2019). Design and validation of an observational instrument for defence in soccer based on the Dynamical Systems Theory. *International Journal of Sports Sciences and Coaching*. 14, 138–152. doi: 10.1177/1747954119827283

Fitzpatrick, A. R. (1983). The meaning of content validity. *Applied Psychology Measurement.* 7, 3–13. doi: 10.1177/014662168300700102

Folgado, H., Bravo, J., Pereira, P., & Sampaio, J. (2019). Towards the use of multidimensional performance indicators in football small-sided games: the effects of pitch orientation. *Journal of sports sciences*, 37(9), 1064–1071. https://doi.org/10.1080/02640414.2018.1543834

Furley, P., Dicks, M., Jordet, G. (2020). The psychology of soccer. Chapter: The Psychology of Penalty Kicks - The Influence of Emotions on Penalty Taker and Goalkeeper Performance. Routledge, 1st Edition.

Gabbett, T. J. (2016). The training-injury prevention paradox: should athletes be training smarter and harder? *British Journal of Sports Medicine*. 50, 273–280. doi: 10.1136/bjsports-2015-095788

Gabin, B., Camerino, O., Anguera, M. T., and Castañer, M. (2012). Lince: multiplatform sport analysis software. Procedia Soc. Behav. Sci. 46, 4692–4694. doi: 10.1016/j.sbspro.2012.06.320

Garcia-Santos, D., and Ibanez, S. J. (2016). Design and validation of an observational instrument to basketball referee evaluation (IOVAB). *SPORT TK-Revista EuroAmericana de Ciencias del Deporte*. 5, 15–26. doi: 10.6018/264601

Garganta, J. (1997). Tactical Modelling of the Football Match: Study of the Organisation of the Offensive Phase in High Performance Teams. Thesis, in portuguese. Porto: University of Porto

Gómez-Carmona, C., Gamonales, J., Pino-Ortega, J., & Ibáñez, S. (2018). Comparative Analysis of Load Profile between Small-Sided Games and Official

Matches in Youth Soccer Players. *Sports*, 6(4), 173. MDPI AG. http://dx.doi.org/10.3390/sports6040173

Göral, K. (2016). FIFA U-20 world cup 2013: analysis and evaluation of goals scored. *RBFF Brazilian Magazine of Futsal and Football*, 8(28), 29–38. https://dialnet.unirioja.es/servlet/ articulo?codigo=5343411

Graham, S. R., Cormack, S., Parfitt, G., and Eston, R. (2018). Relationships between model estimates and actual match-performance indices in professional Australian footballers during an in-season macrocycle. *International Journal of Sports Physiology and Performance*. 13, 339–346. doi: 10.1123/ijspp.2017-0026

Gürkan, O., Ertetik, G., & Müniroglu, S. (2017). Analysis of goals scored in UEFA champions league by the time periods. *International Journal of Science Culture and Sport*, 5(3), 140–147. https://doi.org/10.14486/IntJSCS660

Halouani, J., Chtourou, H., Gabbett, T., Chaouachi, A., & Chamari, K. (2014). Small-sided games in team sports training: a brief review. *Journal of strength and conditioning research*, 28(12), 3594–3618. https://doi.org/10.1519/JSC.0000000000000564

Hartley, R., & Zisserman, A. (2004). Multiple View Geometry in Computer Vision (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511811685

Hasson, F., Keeney, S., and McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*. 32, 1008–1015. doi: 10.1046/j.1365-2648.2000.t01-1-01567.x

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based nursing*, 18(3), 66–67. https://doi.org/10.1136/eb-2015-102129

Higueras-Herbada, A., Lopes, J. E., Travieso, D., Ibáñez-Gijón, J., Araújo, D., and Jacobs, D. M. (2020). Height after side: goalkeepers detect the vertical direction of association-football penalty kicks from the ball trajectory. *Frontiers in Psychology*. 1:311. doi: 10.3389/fpsyg.2020.00311

Hill-Haas, S. V., Dawson, B., Impellizzeri, F. M., & Coutts, A. J. (2011). Physiology of small-sided games training in football: a systematic review. *Sports medicine (Auckland, N.Z.)*, *41*(3), 199–220. https://doi.org/10.2165/11539740-000000000-00000

Hill-Haas, S, Coutts, A, Rowsell, G, and Dawson, B. (2009). Generic versus small-sided game training in soccer. *International Journal of Sports Medicine* 30: 636–642.

Horn, M., de Waal, S., & Kraak, W. (2020). In-match penalty kick analysis of the 2009/10 to 2018/19 English Premier League competition. *International Journal of Performance Analysis in Sport.* 21(1),139–155. doi:10.1080/24748668.2020.1855052

Hughes, M., and Wells, J. (2002). Analysis of penalties taken in shoot-outs. *International Journal of Performance Analysis in Sport* 2, 55–72. doi: 10.1080/24748668.2002.11868261

Hunter, A. H., Murphy, S. C., Angilletta, M. J., Jr, & Wilson, R. S. (2018). Anticipating the Direction of Soccer Penalty Shots Depends on the Speed and Technique of the Kick. *Sports (Basel, Switzerland)*, *6*(3), 73. https://doi.org/10.3390/sports6030073

IFAB. Amendments to the Laws of the Game 2015/16. Available online: https://resources.fifa.com/mm/document/affederation/ifab/02/60/91/38/circular_log_amendments_2015_v 1.0_en_neutral.pdf (accessed on 20 January 2022).

Instat (2020). Penalties: Ultimate guidelines; Massive research of penalties by Instat. Available online at: https://instatsport.com/football/article/penalty_research (accessed January 25, 2022).

Jamil, M., Littman, P., Beato, M. (2020). Investigating inter-league and inter-nation variations of key determinants for penalty success across European football, *International Journal of Performance Analysis in Sport*, 20:5, 892-907, DOI: 10.1080/24748668.2020.1794720

Jordet, G., Hartman, E., Visscher, C., and Lemmink, K. A. P. M. (2007). Kicks from the penalty mark in soccer: the roles of stress, skill, and fatigue for kick outcomes. *Journal of Sports Sciences*. 25, 121–129. doi: 10.1080/02640410600624020

Kuhn, W. (1988). “Penalty-kick strategies for shooters and goalkeepers,” in Science and Football, eds. T. Reilly, K. Lees, and K. Davids (London: E and FN Spon), 489–492.

Lames, M. (1994). Systematische Spielbeobachtung (in German, Systematic game observation). Münster: Philippka-Sportverlag.

Lames, M. (2006). Modelling the interaction in game sports - Relative phase and moving correlations. *Journal of Sports Science and Medicine*. 5, 556–560.

Lames, M., and McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, 7, 62–79. doi: 10.1080/24748668.2007.11868388

Lames, M., and Hansen, G. (2001). Designing observational systems to support top-level teams in game sports. *International Journal of Performance Analysis in Sport* 1, 85–91. doi: 10.1080/24748668.2001.11868251

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London* B, 352, 1231 – 1239.

Li, Y., Alexander, M., Glazebrook, C., and Leiter, J. (2015). Prediction of kick direction from kinematics during the soccer penalty kick. 3, 1–7. *International Journal of Kinesiology & Sports Science*. doi:10.7575/aiac.ijkss.v.3n.4p.1.

Lindstrom, D. (2010). Schaum’s Easy Outline of Statistics, Second Edition (Schaum’s Easy Outlines) 2nd Edition. McGraw-Hill Education

Lopes, J. E., Araújo, D., Duarte, R., Davids, K., and Fernandes, O. (2012). Instructional constraints on movement and performance of players in the penalty kick. *International Journal of Performance Analysis in Sport* 12, 311–345. doi: 10.1080/24748668.2012.11868602

Lopes, J. E., Jacobs, D. M., Travieso, D., & Araújo, D. (2014). Predicting the lateral direction of deceptive and non-deceptive penalty kicks in football from the kinematics of the kicker. *Human Movement Science*, 36, 199–216. doi:10.1016/j.humov.2014.04.004

Mackenzie, R., and Cushion, C. (2013). Performance analysis in football: a critical review and implications for future research. *Journal of Sports Sciences*. 31, 639–676. doi: 10.1080/02640414.2012.746720

Makaruk, H., Porter, J.M., Bodasińska, A., Palmer, S. (2020). Optimizing the penalty kick under external focus of attention and autonomy support instructions, *European Journal of Sport Science*, 20:10, 1378-1386, DOI: 10.1080/17461391.2020.1720829

Medina, D.; Pons, E.; Gomez, A.; Guitart, M.; Martin, A.; Vazquez-Guerreo, J.; Camenforte, I.; Carles, B.; Font, R. (2017). Are There Potential

Safety Issues Concerning the Safe Usage of Electronic Personal Tracking Devices? The Experience of a Multi-sport Elite Club. *International Journal of Sports Physiology and Performance*. 4, 1–12.

Memmert, D., Noël, B.: The Penalty Kick: The Psychology of Success. Meyer & Meyer Sport, London (2020)

Memmert, D., Hüttermann, S., Hagemann, N., Loffing, F., and Strauss, B. (2013). Dueling in the penalty box: evidence-based recommendations on how shooters and goalkeepers can win penalty shootouts in soccer. *International Review of Sport and Exercise Psychology*. 6, 209–229. doi: 10.1080/1750984X.2013.811533

Moraes, E. L., Cardoso, F., and Teoldo, I. (2014). Analysis of the offensive standards of the Spanish national football team at the 2010 FIFA World Cup® in relation to the “status” of the match (In portuguese). *Brazilian Journal of Physical Education and Sport.* 28, 3. doi: 10.1590/1807-55092014000300361

Morya, E., Ranvaud, R., & Pinheiro, W. M. (2003). Dynamics of visual feedback in a laboratory simulation of a penalty kick. *Journal of Sports Sciences*, 21, 87 – 95.

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in sports and active living*, 2, 50. https://doi.org/10.3389/fspor.2020.00050

Nakai, M., Tsunodam Y., Hayashi, H., Murakoshi, H. (2019). Prediction of Basketball Free Throw Shooting by OpenPose. In: Kojima K., Sakamoto M., Mineshima K., Satoh K. (eds) New Frontiers in Artificial Intelligence. JSAI-isAI 2018. Lecture Notes in Computer Science, vol 11717. Springer, Cham. https://doi.org/10.1007/978-3-030-31605-1_31

Navarro, M., van der Kamp, J., Ranvaud, R., and Savelsbergh, G. J. (2013). The mere presence of a goalkeeper affects the accuracy of penalty kicks. *Journal of Sports Sciences*. 31, 921–929. doi: 10.1080/02640414.2012.762602

Nevill, A., Atkinson, G., and Hughes, M. (2008). Twenty-five years of sport performance research in the Journal of Sports Sciences, *Journal of Sports Sciences*, 26:4, 413 — 426

Noël, B., and van Der Kamp, J. (2012). Gaze behaviour during the soccer penalty kick: an investigation of the effects of strategy and anxiety. *International Journal of Sport Psychology*. 43, 326–345. doi: 10.7352/IJSP.2012.43.326

Noël, B., Furley, P., van der Kamp, J., Dicks, M. and Memmert, D. (2015). The development of a method for identifying penalty kick strategies in association football. *Journal of Sports Sciences*. 33(1), 1-10.

Noël B, van der Kamp J and Klatt S (2021). The Interplay of Goalkeepers and Penalty Takers Affects Their Chances of Success. *Frontiers in Psychology*. 12:645312. doi: 10.3389/fpsyg.2021.645312

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory (pp. 1 – 18). New York: Plenum Press.

O'Donoghue, P. (2009). Research Methods for Sports Performance Analysis. London: Routledge. doi: 10.4324/9780203878309

O'Donoghue, P. (2010). Research Methods for Sports Performance Analysis. London: Routledge.

Ortega-Toro, E., García-Angulo, A., Giménez-Egido, J. M., García-Angulo, F. J., and Palao, J. M. (2019). Design, validation, and reliability of an observation instrument for technical and tactical actions of the offense phase in soccer. *Frontiers in Psychology*. 1:22. doi: 10.3389/fpsyg.2019.00022

Owen, A. L., Newton, M., Shovlin, A., & Malone, S. (2020). The Use of Small-Sided Games as an Aerobic Fitness Assessment Supplement Within Elite Level Professional Soccer. *Journal of human kinetics*, *71*, 243–253. https://doi.org/10.2478/hukin-2019-0086

Paterson, G., van der Kamp, J., and Savelsbergh, G. (2020). Moving advertisements systematically affect gaze behavior and performance in the soccer penalty kick. *Frontiers in Sports and Active Living*. 1:69. doi: 10.3389/fspor.2019.00069

Pinheiro, G. S., Nascimento, V. B., Dicks, M., Costa, V. T., & Lames, M. (2021). Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF). *Frontiers in Psychology*, 12, 661179. https://doi.org/10.3389/fpsyg.2021.661179

Pinheiro, G.S., Quintao, R.C., Custodio, I., Casamichana, D., Couto, B.P. (2018). Analysis of relationships between different training load monitoring tools in elite U-20 soccer. *Human Movement*. 19(5)special/issue:52–59; doi: https://doi.org/10.5114/hm.2018.81014

Praça, G.M., Andrade, A.G.P, Bredt, S.G.T, Moura, F.A., Moreira, P.E.D. (2021) Progression to the target vs. regular rules in Soccer small-sided Games, *Science and Medicine in Football*, DOI: 10.1080/24733938.2020.1869811

Preciado, M., Anguera, M. T., Olarte, M., & Lapresa, D. (2019). Observational Studies in Male Elite Football: A Systematic Mixed Study Review. *Frontiers in Psychology*, 10, 2077. https://doi.org/10.3389/fpsyg.2019.02077

Ramakrishna, V., Munoz, D., Hebert, M., Bagnell, J.A., and Sheikh, Y. (2014). Pose Machines: Articulated pose estimation via inference machines. In European Conference on Computer Vision, pages 33-47. Springer.

Randall, D. P., Peter, R., and Giacobbi, J. (2009). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*. 8, 213–225. doi: 10.1207/s15327841mpee0804_3

Rampinini, E., Impellizzeri, F. M., Castagna, C., Abt, G., Chamari, K., Sassi, A., & Marcora, S. M. (2007). Factors influencing physiological responses to small-sided soccer games. *Journal of Sports Sciences*, 25(6), 659–666. doi:10.1080/02640410600811858

Reilly, T., Cabri, J., & Araújo, D. (2005). Science and football V: The proceedings of the Fifth World Congress on Science and Football (1. Aufl.). London: Routledge.

Reilly, T and White, C. (2004). Small-sided games as an alternative to interval training for soccer players. *Journal of Sports Sciences* 22: 559.

Robinson, G., and O'Donoghue, P. (2007). A weighted kappa statistic for reliability testing in performance analysis in sport. *International Journal of Performance Analysis in Sport.* 7, 12–19. doi: 10.1080/24748668.2007.11868383

Sangnier, S., Cotte, T., Brachet, O., Coquart, J., and Tourny, C. (2018). Planning training workload in football using small-sided games' density. *Journal of Strength and Conditioning Research*. 33, 2801–2811. doi: 10.1519/JSC.0000000000002598

Sangüesa, A.A., Haro, G., Ballester, C., Martín, A. (2019). Head, Shoulders, Hip and Ball... Hip and Ball! Using Pose Data to Leverage Football Player Orientation. *Barça sports analytics summit*. Available on: https://barcainnovationhub.com/es/event/barca-sports-analytics-summit-2019/

Sangüesa, A.A., Martín, A., Fernández, J., Ballester, C., & Haro, G. (2020a). Using Player's Body Orientation to Model Pass Feasibility in Soccer. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 3875-3884.

Sangüesa, A.A., Martín, A., Fernandez, J., Rodríguez, C., Haro, G., and Ballester, C. (2020b). Always Look on the Bright Side of the Field: Merging Pose and Contextual Data to Estimate Orientation of Soccer Players. arXiv e-prints, page arXiv:2003.00943.

Sarmento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N., and Leitão, J. C. (2014). Match analysis in football: a systematic review. *Journal of Sports Sciences*. 32, 1831–1843. doi: 10.1080/02640414.2014.898852

Sarmento, H., Clemente, F. M., Araújo, D., Davids, K., McRobert, A., & Figueiredo, A. (2018). What Performance Analysts Need to Know About Research Trends in Association Football (2012-2016): A Systematic Review. *Sports medicine* (Auckland, N.Z.), 48(4), 799–836. https://doi.org/10.1007/s40279-017-0836-6

Savelsbergh, G. J. P., Versloot, O., Masters, R., and van der Kamp, J. (2010). Saving Penalties, Scoring Penalties. eds. I. Renshaw, K. Davids, and G. J. P. Savelsbergh (London: Routledge), 57–68.

Scurr, J., & Hall, B. (2009). The effects of approach angle on penalty kicking accuracy and kick kinematics with recreational soccer players. *Journal of Sports Science & Medicine.*, *8*(2), 230–234.

Slutter, M.W.J., Thammasan, N. and Poel, M. (2021). Exploring the Brain Activity Related to Missing Penalty Kicks: An fNIRS Study. *Frontiers in Computer Science*. 3:661466. doi: 10.3389/fcomp.2021.661466

Sgrò, F, Bracco, S, Pignato, S, Lipoma, M. (2018). Small-sided games and technical skills in soccer training: systematic review and implications for sport and physical education practitioners. *Journal of Sports Sciences* ; 6: 9–19. DOI: 10.17265/2332-7839/2018.01.002

Soto, A., Camerino, O., Iglesias, X., Anguera, M. T., and Castañer, M. (2019). LINCE PLUS: research software for behavior video analysis. *Apunts Educación Física y Deportes.* 137, 149–153. doi: 10.5672/apunts.2014-0983.es.(2019/3).137.11

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. *Journal of Graduate Medical Education*. 3, 119–120. doi: 10.4300/jgme-d-11-00075.1

Timmis, M. A., Piras, A., and van Paridon, K. N. (2018). Keep your eye on the ball; the impact of an anticipatory fixation during successful and unsuccessful soccer penalty kicks. *Frontiers in Psychology*. 1:58. doi: 10.3389/fpsyg.2018.02058

van der Kamp, J. (2006). A field simulation study of the effectiveness of penalty kick strategies in soccer: late alterations of kick direction increase errors and reduce accuracy. *Journal of Sports Sciences*. 24, 467–477. doi: 10.1080/02640410500190841

Teijlingen, E.V., & Hundley, V.A. (2002). The importance of pilot studies. Nursing standard (Royal College of Nursing (Great Britain) : 1987), 16 40, 33-36.

Vilagut G. (2014) Test-Retest Reliability. In: Michalos A.C. (eds) Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0753-5_3001

Villarejo, D., Ortega, E., Gómez, M. Á, and Palao, J. M. (2014). Design, validation, and reliability of an observational instrument for ball possessions in rugby union. *International Journal of Performance Analysis in Sport*. 14, 955–967. doi: 10.1080/24748668.2014.11868771

Vogt, W.P. (2005). Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences. SAGE.

Wei, S., Ramakrishna, V., Kanade, T., and Sheikh, Y. (2016). Convolutional pose machines. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p. 4724-4732,

Weigelt, M., and Memmert, D. (2012). Goal-side selection in soccer penalty kicking when viewing natural scenes. *Frontiers in Psychology*. 3:312. doi: 10.3389/fpsyg.2012.00312

Wood, G., Jordet, G., and Wilson, M. R. (2015). On winning the “lottery”: psychological preparation for football penalty shoot-outs. *Journal of Sports Sciences*. 33, 1758–1765. doi: 10.1080/02640414.2015.1012103

White, S., and O’Donoghue, P. (2013). “Factors influencing penalty kick success in soccer,” in Science and football VII: The Proceedings of the Seventh

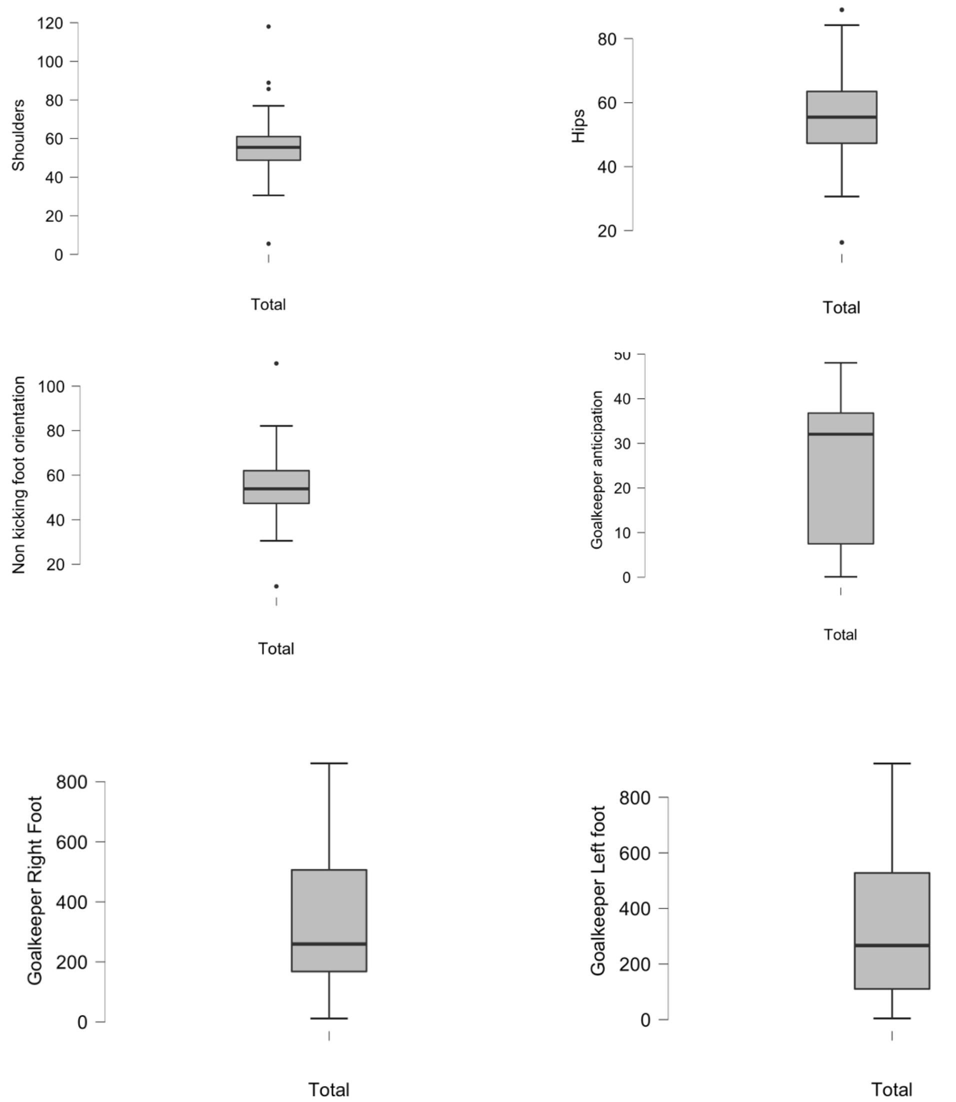
World Congress on Science and Football, eds H. Nunome, B. Drust, and B. Dawson (London: Routledge), 237–242.

Wunderlich, F., Berge, F., Memmert, D., and Rein, R. (2020). Almost a lottery: the influence of team strength on success in penalty shootouts. *International Journal of Performance Analysis in Sport* 20, 1–13. doi: 10.1080/24748668.2020.1799171

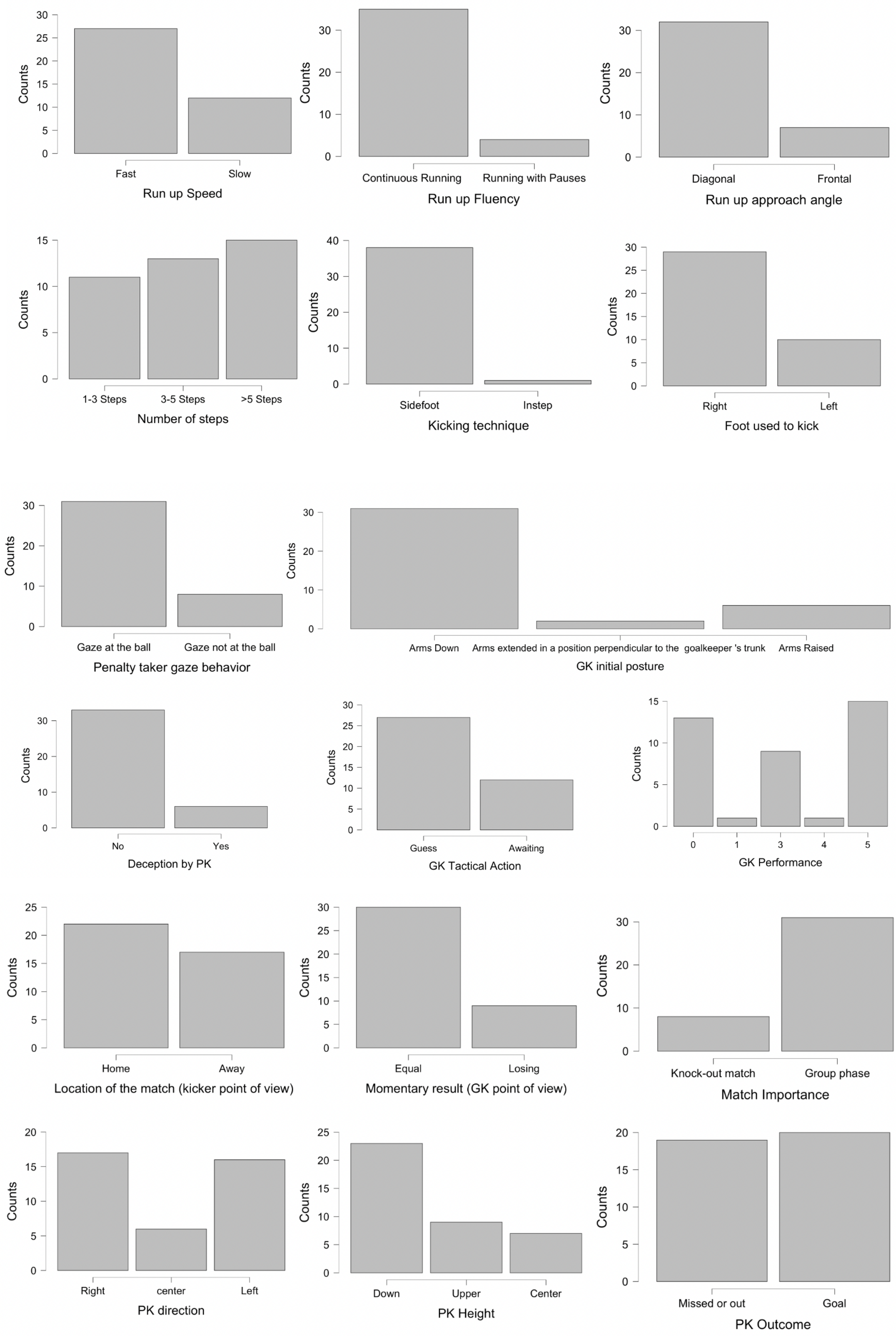
Zurutuza, U., Castellano, J., Echeazarra, I., Guridi, I. and Casamichana, D. (2020). Selecting Training-Load Measures to Explain Variability in Football Training Games. *Frontiers in Psychology*. 10:2897. doi: 10.3389/fpsyg.2019.02897

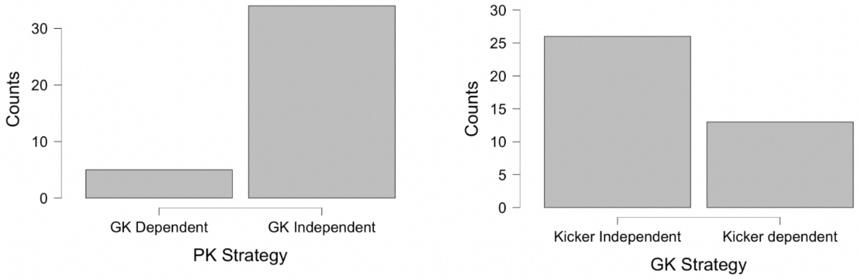
SUPPLEMENTARY MATERIAL

Paper I. Graph 1. Descriptive data of OpenPose variables

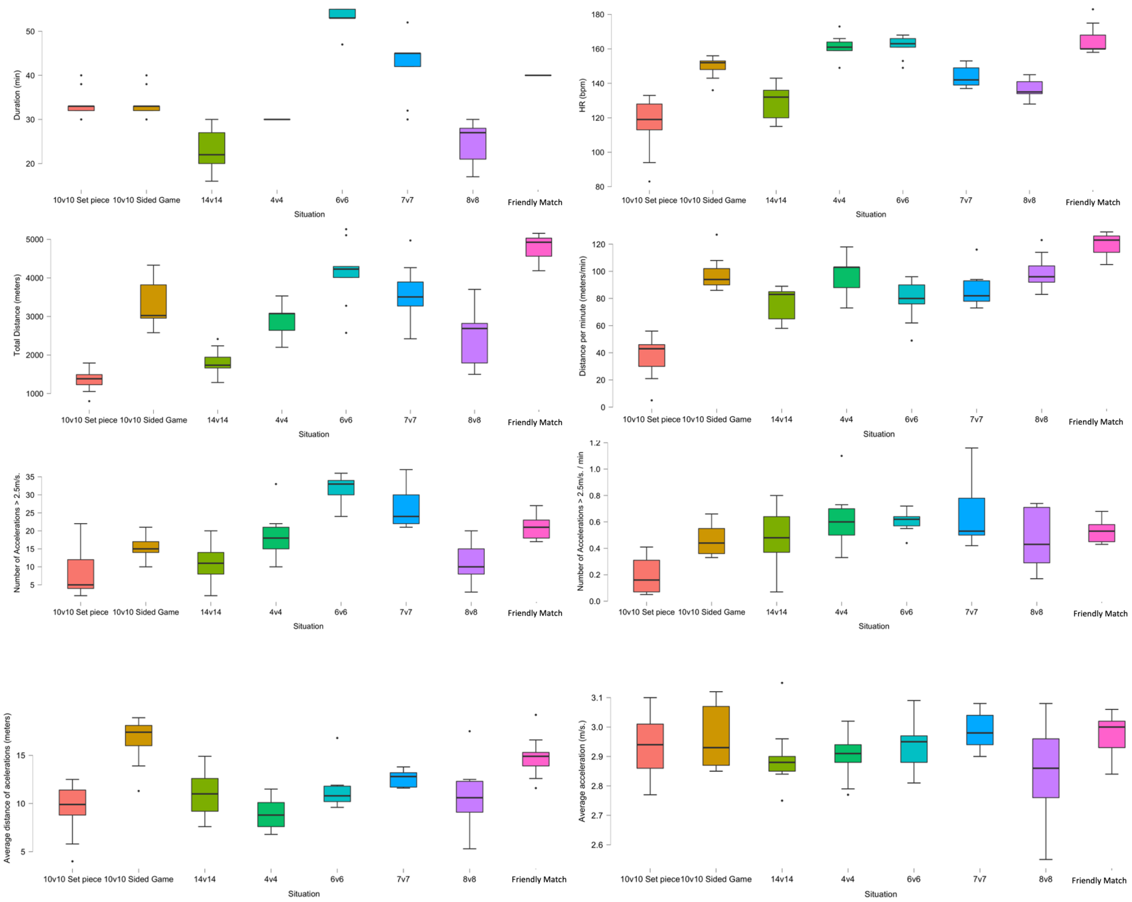


Paper I. Graph 2. Distribution of OSPAF variables





Paper IV. Graph 3. Descriptive statistics of duration, heart rate, total distance (m), distance per minute (m/min), the total number of accelerations >2.5 m/s^2, number of accelerations >2.5 m/s^2 per minute, average distance of accelerations (m) and average accelerations (m/s^2)



Paper 4. Supplementary tables

Table 1. Comparison of the physical and physiological demands of the Friendly match (FM) and the SSG 4v4

	HR (bpm)		Total distance (m)		Distance per minute (m/min)		N° accelerations > 2.5m/s^2		N° accelerations > 2.5m/s^2 per minute		Average Distance accelerations (m)	
	FM	**4v4**	**FM**	**4v4**	**FM**	**4v4**	**FM**	**4v4**	**FM**	**4v4**	**FM**	**4v4**
mean	165.44	161.56	4,807.22	2,901.89	120.22	96.78	21.33	18.89	0.54	0.63	14.89	8.94
sd	8.52	6.44	311.58	391.71	7.71	13.26	3.43	6.49	0.08	0.22	2.2	1.66
P	0.984		< 0.001		0.005		0.997		0.975		< 0.001	
d	-		5.385		2.162		-		-		3.046	

Legend: sd = Standard deviation; *P* = P value; d = Cohen's d (Effect size)

Table 2. Comparison of the physical and physiological demands of the SSG 6v6, 7v7 and 8v8

	HR (bpm)			Total distance (m)			Distance per minute (m/min)			N° accelerations > 2.5m/s^2			N° accelerations > 2.5m/s^2 per minute			Average Distance accelerations (m)		
	6v6	**7v7**	**8v8**	**6v6**	**7v7**	**8v8**	**6v6**	**7v7**	**8v8**	**6v6**	**7v7**	**8v8**	**6v6**	**7v7**	**8v8**	**6v6**	**7v7**	**8v8**
mean	161.56	144.33	136.11	4,135.44	3,602.22	2,465.67	78.22	86.11	98.89	31.89	26.56	11.67	0.6	0.66	0.48	11.47	12.61	10.62
sd	6.48	6.08	5.6	825.66	742.77	792.44	15	13.49	12.96	3.66	6.06	5.72	0.07	0.24	0.22	2.17	0.8	3.56
P	0.985	< 0.001	< 0.001	0.233	0.002	< 0.001	< 0.001	< 0.001	0.015	< 0.001	0.440	0.002	0.993	0.828	0.998	0.063	0.472	0.006
d	-	2.853	4.070	-	2.116	3.889	3.521	3.104	1.998	2.979	-	2.049	-	-	-	-	-	1.443

Legend: sd = Standard deviation; *P* = P value; d = Cohen's d (Effect size)

Table 3. Comparison of the physical and physiological demands of the Friendly match (FM) and the SSG 10v10 sided, 10v10 set piece and 14v14

	HR (bpm)				Total distance (m)				Distance per minute (m/min)				N° accelerations > 2.5m/s²				N° accelerations > 2.5m/s² per minute				Average Distance accelerations (m)			
	FM	**10v10***	**10v10****	**14v14**	**FM**	**10v10***	**10v10****	**14v14**	**FM**	**10v10***	**10v10****	**14v14**	**FM**	**10v10***	**10v10****	**14v14**	**FM**	**10v10***	**10v10****	**14v14**	**FM**	**10v10***	**10v10****	**14v14**
mean	165.44	149.78	115.56	129.89	4,807.22	3,726.33	1,372.89	1,788.44	120.22	97.78	36.44	76.77	21.33	15.33	8.55	10.55	0.54	0.46	0.19	0.45	14.89	16.49	9.33	11.05
sd	8.52	6.47	16.87	10.25	311.58	587.71	314.91	369.23	7.71	13.08	15.63	11.68	3.43	3.31	6.67	5.91	0.08	0.11	0.13	0.23	2.2	2.45	2.84	2.39
P		0.011	0.001	0.001		0.001	0.001	0.001		0.012	0.001	0.001		0.265	0.001	0.002		0.988	0.004	0.975		0.842	0.001	0.023
d		2.071	3.734	1.670		3.255	10.967	8.839		2.090	8.680	4.390		-	2.409	2.229		-	2.926	-		-	2.188	1.670

Legend: sd = Standard deviation; *P* = P value; d = Cohen's d (Effect size); *10v10 sided game; **10v10 set piece

APPENDIX

frontiers in Psychology

METHODS
published: 28 May 2021
doi: 10.3389/fpsyg.2021.661179

Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF)

Guilherme de Sousa Pinheiro[1], Vitor Bertoli Nascimento[2], Matt Dicks[3], Varley Teoldo Costa[4] and Martin Lames[1]*

[1] Chair for Performance Analysis and Sport Informatics, Department of Sports and Health Sciences, Technical University of Munich, Munich, Germany, [2] Motor control and Biomechanics, Department of Sports, State University of Londrina, Londrina, Brazil, [3] School of Sports, Health and Exercise Sciences, University of Portsmouth, Portsmouth, United Kingdom, [4] UFMG Soccer Science Center, Department of Sports, Federal University of Minas Gerais, Belo Horizonte, Brazil

OPEN ACCESS

Edited by:
Miguel-Angel Gomez-Ruano,
Polytechnic University of Madrid,
Spain

Reviewed by:
Bingnan Gong,
Capital Institute of Physical Education and Sports, China
Cesar Mendez,
Rey Juan Carlos University, Spain

***Correspondence:**
Guilherme de Sousa Pinheiro
guilherme.pinheiro@tum.de

Specialty section:
This article was submitted to Movement Science and Sport Psychology,
a section of the journal Frontiers in Psychology

Received: 30 January 2021
Accepted: 04 May 2021
Published: 28 May 2021

Citation:
Pinheiro GS, Nascimento VB, Dicks M, Costa VT and Lames M (2021) Design and Validation of an Observational System for Penalty Kick Analysis in Football (OSPAF). Front. Psychol. 12:661179. doi: 10.3389/fpsyg.2021.661179

The analysis of penalty kick has played an important role in performance analysis. The study aims are to get formal feedback on the relevance of variables for penalty kick analysis, to design and validate an observational system; and to assess experts' opinion on the optimum video footage in penalty kick analysis. A structured development process was adopted for content validity, reliability and agreement on video usage. All observational variables included in OSPAF showed Aiken's V values above the cut-off (for 5-scale $V > 0.64$; for 2-scale = $V > 0.75$; $p < 0.05$). Cohen's Kappa resulted in mean intra- and inter-rater reliability values of 0.90 and 0.86, respectively. It is recommended to combine at least three different viewing angles ($V = 0.90$; $p = 0.006$) with standardization of video quality ($V = 0.95$; $p = 0.006$). Changing the viewing angles may influence the observer perception ($V = 0.86$; $p = 0.006$). The aerial and pitch-level viewing angle behind the penalty taker and pitch-level viewing angle behind the goalkeeper were indicated as most appropriate for observational analysis ($V = 0.97$; $p = 0.01$). The OSPAF met all requirements of instrument validation. It may be recommended as basis of future observational systems on penalty kicks.

Keywords: performance analysis, observational methodology, content validity, Aiken's V, elite football

INTRODUCTION

Penalty kicks play a decisive role in the outcome of a match in competitive football (Makaruk et al., 2020; Paterson et al., 2020). UEFA introduced penalty shoot-outs to major tournaments in 1976 (FIFA followed in 1978) as a means of deciding matches in the knockout phase of major tournaments when the score is a draw at the end of the match. Many important competitions were decided by penalties, for example, in the 2018 World Cup, 25% of the matches in the knockout stages were decided by penalty shoot-outs. In the big five European leagues, the average number of penalties per game since the 2017/18 season is 0.31 (Instat, 2020). It has been estimated that about 30–40% of the goals are scored from set plays (i.e., penalty kicks, free kicks, corners, and throw-ins) (Fariña et al., 2013; Sarmento et al., 2018). Among these set plays, the penalty is the situation with the highest chance of scoring a goal. The penalty situation can be considered as the peak of high-pressure performance in elite football (Wood et al., 2015; Brinkschulte et al., 2020). Notably, researchers have shown a significant interest in uncovering factors that affect success in the penalty kick (Paterson et al., 2020).

To further improve the performance of football players, a large amount of data is produced in professional leagues that provide many options to analyze games and identify critical factors for success (Lepschy et al., 2020). Penalty kicks have been mainly analyzed in two contexts: First, in a laboratory or other non-game controlled settings (video-simulation and in-situ experimental conditions), aiming at the analysis of perceptual-motor and cognitive aspects of performance (e.g., Dicks et al., 2010; Lopes et al., 2012; Weigelt and Memmert, 2012; Navarro et al., 2013); and second, in real match situations, enabling the identification of prominent factors that affect both players' performances and the penalty kick outcome using mainly observational methods (e.g., Chiappori et al., 2002; Jordet et al., 2007; White and O'Donoghue, 2013; Horn et al., 2021). While in the first context, a common theoretically motivated focus has been developed to enhance the representative design of methods used to examine the expertise of penalty takers and goalkeepers (Dicks et al., 2009), in the second, researchers have attempted to improve data collection procedures based on game video analysis (Almeida et al., 2016).

From a behavioral perspective, the penalty outcome depends, above all, on the emerging results of the "penalty taker—goalkeeper" dyadic interaction (Lopes et al., 2012; Almeida et al., 2016). Reviews on performance analysis have suggested that the future of game analysis in football requires the building of observational instruments and analytical procedures that integrate the study of criteria related to the interactions between opponents (Mackenzie and Cushion, 2013; Sarmento et al., 2014). Notational analyses are scientific procedures that reveal the occurrence of perceivable behaviors, allowing them to be formally recorded and quantified (Anguera et al., 2001). To support top-level teams two purposes of game observation are predominant: preparation against a future opponent and the optimization of training (Lames and Hansen, 2001; Lames and McGarry, 2007). The analysis, therefore, aims to describe the participants' behavior during real competitive scenarios, concerning different tactical, technical, and performance aspects (Vázquez-Diz et al., 2019).

Although observational studies are frequently used and their utility in different contexts has been widely proven, there are concerns regarding the information related to validity and reliability concerning the processes of systematic observation (Chacón-Moscoso et al., 2018). As a prerequisite for any performance analysis research that uses a novel system or instrument, the repeatability and accuracy of this new tool and the validity of performance indicators it provides should be tested before collecting and analyzing players and teams' performances (O'Donoghue, 2014; Gong et al., 2019).

Validity and reliability are important criteria for any scientific measurement. Validity is generally referred to as the ability of a measurement tool to reflect what it is designed to measure, and usually, for performance analysis instruments, it can be determined through expert coaches' opinions in each sports category (O'Donoghue, 2009; AERA et al., 2014). Reliability is the consistency of a measure and is a part of the evidence of validity (Sullivan, 2011; Heale and Twycross, 2015). It refers to the reproducibility of values of a test, assay, or other measurements in repeated trials by the same individuals (intra-observer reliability) and repeatability over different observers (inter-observer reliability) (O'Donoghue, 2009; Gong et al., 2019). Because the human observer is the measurement instrument in observation, it is highly recommended to establish content validity and reliability of notational systems and observational instruments to reduce the error caused by human subjectivity (Lames, 1994; O'Donoghue, 2009; Taherdoost, 2016; Cobb et al., 2018).

Previous research has used a development process to evidence content validity of an observational instrument (Brewer and Jones, 2002; Fernandes et al., 2019). This process includes several sub-stages, such as literature review, instrument development, the establishment of content validity with experts, observer training, pilot study, inter-observer reliability, and intra-observer reliability assessment. This suggested development process has guided the present study.

Several studies have investigated the penalty kick strategies in football (e.g., Savelsbergh et al., 2005; van der Kamp, 2006; Bowtell et al., 2009; Vega Marcos et al., 2010; Zuo et al., 2010; Lopes et al., 2012; Timmis et al., 2014; Furley et al., 2017; Sarmento et al., 2018; Makaruk et al., 2020), however, only a few studies proposed instruments to examine this set piece. For example, Comas et al. (2018) created a system based on an observational methodology with to analyze the direction of the ball at a penalty kick in football. These authors pointed out a relationship between the spatial position of the support foot and the opposite arm to the shooting foot with the direction of the ball on the penalty kick, both for right-footed and left-footed players. Noël et al. (2015) developed a method for identifying the determinants of penalty kick strategies in a controlled simulated situation, before evaluating penalty kick performances using video footage from competitive matches. They included 12 variables in this observational system. A logistic regression model identified three variables (attention to the goalkeeper, run-up fluency, and kicking technique) that in combination could predict kick strategy in 92% of the penalties. However, one possible limitation is that the penalty takers followed a script denoting whether they use a keeper independent or dependent strategy and therefore the design did not reflect the unfolding of an interaction between the two players. In real competitions, penalty kicks are an interaction process, and the observable performance is rather the emergent result of this interaction process than the display of skills and abilities of the two parties (Lames, 2006; Lames and McGarry, 2007). Future research needs to address how these factors affect the validity of the instrument.

Performance indicators play a key role in contemporary sports analytics (Sampaio and Leite, 2013). In training and coaching, these metrics play an important role mostly as starting point for a more in-depth qualitative game analysis (Lames and Hansen, 2001; Carling et al., 2014). There are several studies analyzing penalties in field settings (Dalton et al., 2015; Wood et al., 2015; Brinkschulte et al., 2020; Higueras-Herbada et al., 2020; Wunderlich et al., 2020; Horn et al., 2021). Most of them do not aim at giving a detailed description of the actions of the shooter and goalkeeper, but focus more on statistical results (e.g., quotes for scoring and saving penalties of different kinds). Despite the extensive coverage in the literature of penalty kicks in elite football and methods developed for the analysis of penalty taker actions (Noël et al., 2015), there is no scientific consensus concerning observational variables to use for the analysis of both goalkeeper and penalty taker actions. The development of this instrument enables the collection of data using systematic observation. Differentiating between penalty kick patterns would be of both scientific and applied interest. This would allow researchers to identify determinants of successful kicks (e.g., patterns of gaze) especially under high pressure, as well as facilitating future comparisons between investigations on this topic. Also, practitioners in professional football could distinguish penalty kick strategies and so inform coaching, training, and scouting. OSPAF may serve as a standard tool for observational investigations of penalties in football to make the results from different studies more comparable. This would allow for replications of studies to track for example long-time trends and also for comparisons between different settings (e.g., countries, leagues, age groups, gender). Therefore, a methodological design containing three studies (pilot study, main study, and video requirements study) was carried out. The aim of (1) the pilot study was to get formal feedback on variables for penalty kick analysis suggested by professionals in the area; (2) the main study aimed at designing and validating an observational system applied to in-match penalty kick analysis; and (3) the video study served to evaluate the influence of the video footage (i.e., viewing angles, number of angles and video quality) on penalty kick analysis through an observational system.

MATERIALS AND METHODS

Pilot Study

Participants in the pilot study were four sports scientists and three high-level football coaches (43.32 ± 15.48 years). The inclusion criteria established for forming part of the panel of sports scientists were: (1) postgraduate master in sports sciences or Ph.D. in sports sciences, (2) to have had at least 3 years experience as a university researcher in sports sciences, (3) experience in performance analysis research (final master's thesis, doctoral thesis or scientific publication); and for high-level football coaches were: (1) graduate in physical activity and/or sport sciences, (2) have an official license as a football coach, (3) more than 3 years as a football coach in a team of an official competition. They evaluated and provided judgment on the instrument's variables. All participants provided informed consent after details of the study were communicated in written form before participation in the study. All procedures performed in the study were in strict accordance with the Declaration of Helsinki as well as the ethical standards of the Technical University of Munich.

The pilot study refers to a mini version of the full-scale study, as well as the specific pre-testing of the particular research instrument, here the online questionnaire (Van Teijlingen and Hundley, 2001). The pilot study aimed to was to get formal feedback on variables for penalty kick analysis and to collect observable variables suggested by professionals in the area. A survey, developed in Google Forms, was used to assess the content validity of the proposed observational system (Fernandes et al., 2019). A link for the online survey was emailed to the participants. They were instructed to answer the questionnaire on a computer or notebook, and there was no time limit to answer the questions.

Main Study

A panel of 20 experts (41.85 ± 13.96 years), from Brazil, England, Germany, Israel, Netherlands, Romania, and Spain, who met the following criteria: (1) Ph.D. in sports sciences, and (2) experience of publishing in penalty kick research was contacted and voluntarily agreed to participate. More detailed characteristics about the experts were collected, such as sports biography and open items on the experts' general judgment on each criterion. All participants provided informed consent after details of the study were communicated in written form before participation in the study. All procedures performed in the study were in strict accordance with the Declaration of Helsinki as well as the ethical standards of the Technical University of Munich.

The main study aimed to design and validate an observational system applied to in-match penalty kicks analysis and to follow a systematic process to accumulate evidence of content validity and reliability to adequately categorize and record behaviors of both penalty takers and goalkeepers during penalty kicks. The process to achieve content validity for the OSPAF is described below in different stages, adapted from Brewer and Jones (2002) and Fernandes et al. (2019):

Content Validity With Experts

The panel of experts answered the survey in web format and the level of concordance among experts for each of the variables proposed in the OSPAF was analyzed. A modified Delphi method was performed (Dalkey and Helmer, 1963; Hasson et al., 2000; Dayé, 2018). For concordance analysis three dimensions were defined (Fitzpatrick, 1983; Fernandes et al., 2019):

- Agreement: the degree of general acceptance of the variables to be included in the observational system. The question in the survey was: How is your level of agreement with the inclusion of the variable for penalty kick analysis in the proposed system? A five-point Likert scale (Strongly disagree, Disagree, Neither disagree nor agree, Agree, Strongly agree) was utilized.
- Univocity: clarity domain of a definition; a binary scale (Yes or No). The question in the survey was: The definition of the variable is clear enough for understanding?
- Adequacy: level of pertinence and importance of criteria. The question in the survey was: What is the level of importance of the variable for the observational system? A different five-point Likert scale (Very low, Low, Medium, High, Very high) was applied (Jamieson, 2004).

Inter- and Intra-Observer Reliability

The verification of the reliability of OSPAF was made through the assessment of Cohen's kappa (κ) between observers (inter-observer agreement) and for the analysis of interpretative stability within one observer (intra-observer agreement). For the inter-observer agreement, apart from the analysis carried out by the main researcher, a second researcher was trained in the analysis of the penalty kicks with OSPAF. After the training period, the two observers independently analyzed 40 randomly selected penalty kicks of the World Cups 2014 and 2018. Regarding the intra-observer agreement, the principal investigator performed the same analysis 4 weeks after the first analysis thus minimizing task familiarity (Robinson and O'Donoghue, 2007), without conducting any type of analysis during this time, thus checking the temporal stability of the analysis (Aranda et al., 2019).

Video Study

Participants and Procedures

The same panel of 20 experts as in the main study participated also in this third study, aiming to evaluate the influence of the viewing angles for penalty kick analysis through an observational system. Using an online questionnaire, 14 penalty kick videos from elite football each from 7 different angles were presented (**Supplementary Figure 1**). The methodology adopted in the present study is similar to Baranowski and Hecht (2017) (i.e., fifteen-second scenes used as examples, and later on a questionnaire was applied to gather feedback). The videos had a pixel resolution of 1,280 × 720. The experts should indicate which were the best viewing angles for penalty kick analysis. They were instructed to watch the videos on a computer or a notebook.

The choice of angles was adapted from a division of the field into zones proposed by Garganta (1997) and previously used by Moraes et al. (2014). This corresponds to the topographical division of the playing field, and its use ensured the establishment of spatial references for choosing the angles. Experts could watch each penalty kick video as many times as they judged necessary.

Besides, the experts were asked about how many viewing angles were needed for penalty kick analysis; whether changing the viewing angle could influence the observer's analysis, and whether video quality is a basic prerequisite for standardizing penalty kick analysis using an observational system.

The panel of experts answered the survey in web format and the level of concordance among experts for the following domains were analyzed:

- The number of angles needed for penalty analysis: The question in the survey was: In your opinion, how many video angles are required for the evaluation of a penalty kick in observational studies? A five-point Likert scale (1 video angle, combination of 2 video angles, combination of 3 video angles, combination of 4 video angles, combination of 5 or more video angles) was utilized.
- Influence of changing angles on the observer's analysis: The question in the survey was: In your opinion, changing the angle presented could influence the evaluation of penalty kicks by an observer? A binary scale (1. Yes or 0. No) was used.
- Pre-requisite of video quality: The question in the survey was: In your opinion, the video quality is a prerequisite for penalty kick analysis in football? A binary scale (1. Yes or 0. No) was used.

Instrument

For the pilot and main study, a survey with two different versions developed in Google Forms to assess content validity with the experts. For the video study another online survey, containing penalty kick videos (i.e., 2016 Olympics, World Cups between 2010 and 2018, and major European leagues from 2015 to 2020) was utilized. For reliability, the final version of the OSPAF was used after implementation in Lince Plus software (Gabin et al., 2012; Soto et al., 2019). Lince Plus is free software that has been used by many researchers needing a tool to tag behaviors using video recordings, coding behaviors, and data register (Soto et al., 2019). Dimensions and categories of OSPAF were coded and the observations of the two observers were compared using this software. Criteria were entered with the full definition of the variable (i.e., Run up speed), and categories were coded with the initials letters (i.e., Fast = F and Slow = S), as illustrated in the figure below.

Statistical Analysis

For descriptive analysis, mean and standard deviation were used. Aiken's V was calculated (Aiken, 1985) for content validity of the OSPAF variables and to evaluate the level of agreement of the experts according to the number of angles needed for penalty analysis; the influence of changing angles on the observer's analysis; and the pre-requisite of video quality. Aiken's V allows for quantifying the relevance of items expressed in Likert scales, according to the opinions of a group of experts. Its values vary between 0 and 1, with 1 indicating a perfect agreement among the judges. Previous studies have used the same coefficient to establish validity in observational instruments (Villarejo et al., 2014; Garcia-Santos and Ibanez, 2016; Fernandes et al., 2019; Ortega-Toro et al., 2019). The p level considered for Aiken's V was 0.05 and a 95% confidence interval was used. The score confidence interval was used to provide the expected accuracy of Aiken's V value (Randall et al., 2009). The calculation of Aiken's V is as follows:

$$V \quad \frac{\sum s}{n\,(c-1)}$$

Description: n = number of judges; c = highest value of Likert scale; $s = r - l$; r = the judgement given by a judge; l = lowest value of Likert scale.

For each dimension (agreement, univocity, and adequacy), the criteria for the elimination or acceptance of the items were fixed in advance. The reference table proposed by Aiken (1985) for samples with $n < 25$ was used (number of rating categories: 5; $V = 0.64$, $p < 0.05$; number of rating categories: 2; $V = 0.75$, $p < 0.05$). Consequently, variables with Aiken's V values below cut-off values of 0.64 in agreement or adequacy, or univocity below 0.75, were eliminated.

For intra- and inter-reliability of the OSPAF, Cohen's kappa was calculated. The interpretation of this coefficient was adopted as follows: $\kappa > 0.8$ very good; $0.6 < \kappa < 0.8$ good; $0.4 < \kappa < 0.6$ moderate; $0.2 < \kappa < 0.4$ fair; $\kappa < 0.2$ poor (Altman, 1991; O'Donoghue, 2009). For the specific objectives of this study only values $\kappa > 0.8$ were considered satisfactory (Lames, 1994).

To identify the best angle for penalty analysis, descriptive statistics were used. Microsoft Excel 2016 was used to calculate the values of Aiken's V and confidence interval; Lince Plus software to record the behaviors (Gabin et al., 2012; Soto et al., 2019). The Statistical Package for the Social Sciences software (IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.) was used to calculate Cohen's Kappa.

RESULTS

Pilot Study

The first version of the new penalty kick analysis system was created, based on the collection of several variables from previous studies (Hughes and Wells, 2002; Timmis et al., 2014, 2018; Noël et al., 2015; Almeida et al., 2016; Furley et al., 2017; Comas et al., 2018). Characteristics were selected that are likely to distinguish the profile of successful or unsuccessful penalty kicks and strategies. Furthermore, contextual factors were included (e.g., location of the match; the result of the match at the time of the penalty kick; penalty kick during the normal time or extra time). This first step enabled the development of the first round of content validity of the proposed observational system, with the variables and their definitions.

Qualitative feedback was also gathered from the professionals about the conduct of the online questionnaire. The questionnaire was indicated by 65% of the experts as being very long and complex. In this way, the design of the questionnaire was adjusted for the main study, by repositioning the descriptions of the variables close to the answer box. The use of an online survey presented no problems and was considered a suitable tool for further expert participation.

All the 27 variables proposed in the pilot study have been pointed out by the experts as relevant for the analysis of penalties in football. Aiken's V for Agreement ($p < 0.05$) ranged from 0.66 to 0.89 (cut-off: 0.64); for Adequacy ($p < 0.05$) from 0.64 to 0.83 (cut-off: 0.64) and for Univocity ($p < 0.05$) from 0.89 to 1.0 (cut-off: 0.75). Moreover, the experts suggested the variables ball speed, match importance, and goalkeeper initial posture, which were added to the observational system.

Main Study

The results regarding content validity in the main study were obtained by calculating Aikens' V and the 95% confidence interval. Data are shown in **Supplementary Table 1**. All observable variables of the OSPAF, presented in **Supplementary Table 1**, showed Aiken values above the cut-off value ($p < 0.05$).

Unlike in the pilot study, the following variables have not achieved the minimum values for 5-scale ($p < 0.05$; $V < 0.64$, $n = 20$) and/or for 2-scale ($p < 0.05$; $V < 0.75$; $n = 20$) to be included in the final version of the OSPAF and were excluded. **Supplementary Table 2** shows the excluded variables.

The following **Supplementary Table 3** presents the final version of the OSPAF and the operational definitions.

Supplementary Table 4 shows the Kappa values for each of the variables of the observational system for penalty analysis in football (OSPAF). From the 24 items validated through the Aiken coefficient, 19 obtained Kappa values above 0.80 for intra-reliability, and the other 5 items values above 0.75. Regarding the inter-reliability, 17 items presented Kappa values above 0.80, the 7 remaining items had values above 0.70. Thus, Kappa values indicate a high level of reliability of the proposed new penalty kick analysis system.

Video Analysis Study on Optimum Video Footage

The preferred angles for observational analysis of penalty kicks, indicated by the panel of experts in the present study, are shown in **Supplementary Figure 2**.

It was presented to the experts 14 penalty videos with 7 different viewing angles. 71.4% of the experts indicated the angle c (Behind the penalty taker aerial view), 18.2% the angle d (Behind the penalty taker pitch view), and 10.4% the angle e (Behind the goalkeeper aerial view). The experts agreed on the following methodological requirements: to analyze the penalty in the game through an observational system it is necessary to combine at least 3 different viewing angles of the same penalty ($V = 0.90$; $p = 0.006$). The change of the viewing angles can influence the analysis of the observer ($V = 0.86$; $p = 0.006$). Moreover, the standardization of video quality is a prerequisite for notational analysis ($V = 0.95$; $p = 0.006$).

DISCUSSION

Three studies with different aims were conducted to approve an observational system for in-match penalty kick analysis: (1) the pilot study to get formal feedback on variables for penalty kick analysis suggested by professionals in the area; (2) the main study to design and validate the observational system; and (3) the video study to evaluate the influence of the video footage.

The first version of the observational system was created and validated with practitioners. All 27 proposed observational variables were considered relevant. Also, the practitioners suggested the inclusion of the following variables ball speed, match importance, and goalkeeper initial posture. Technical adjustments have been made about the conduct of the online questionnaire to better understand the proposed questions, e.g., repositioning of questions, explanations, and videos. As study 1 was considered a pilot, it provided a point of discussion for further studies in the present research. One of the advantages of conducting a pilot study is that it might give warning about where the main research project could fail, where research protocols may not be followed, or whether proposed methods or instruments are inappropriate or too complicated. They fulfill a range of important functions and can provide valuable insights for other researchers (Van Teijlingen and Hundley, 2001).

Since the study objectives are oriented to the construction of an observation instrument, the results refer to the quality control of the data, focused on the Aiken's V values, intra and inter-observer agreement. The Aiken's V value, measured in the dimensions Agreement, Univocity and Adequacy, indicated that content validity was achieved (For 5-scale: $p < 0.05$ and $V > 0.64$; for 2-scale: $p < 0.05$ and $V > 0.75$). The methodological rigor adopted and discussed in the study provides sports scientists, coaches, and professionals involved in football an instrument capable of assessing important indicators in a penalty kick in elite football. Although some variables were indicated as valid for penalty kick analysis in the pilot study (i.e., run-up type, run up length, swing behavior of the kicking leg, position of the arm opposite the kicking leg, preparation time, distraction by the goalkeeper, presence of advertisement behind the goal), in the main study they did not reach the minimum values to be included in the final version of the observational system. These findings might be partially explained by the fact that the panel of experts participating in the main study was larger ($n = 20$) than in the pilot study ($n = 7$). Besides, the panel in the main study had substantial scientific research experience in performance analysis. This may lead to a more detailed analysis of the research design.

The kick strategy of the penalty taker is commonly distinguished as being either goalkeeper-dependent or goalkeeper-independent (Kuhn, 1988). Adopting the "goalkeeper-independent" strategy the penalty taker has a pre-established plan about the direction of the kick and ignores any action of the goalkeeper during the preparatory period (run-up). Alternatively, using the "goalkeeper-dependent" strategy the kicker intends to take advantage of the goalkeeper's anticipatory action. During the run-up, the penalty taker tries to obtain information from the actions of the goalkeeper in an attempt to anticipate which side the goalkeeper will dive. The analysis of the kick strategies has been investigated about numerous factors, such as spatiotemporal (e.g., run-up, ball speed; Kuhn, 1988; Noël et al., 2015), foot orientation (e.g., the direction of the supporting foot; Li et al., 2015), perceptual (e.g., visual search behaviors; Noël et al., 2015), individual (e.g., footedness; Instat, 2020), psychological (e.g., team status, kick importance). Although all variables included in OSPAF presented Aiken's V value above the cut-off threshold, few variables showed lower values (i.e., number of steps, foot used to kick, and run-up approach angle). This low rating contrasts previous findings which indicate that the foot used to kick can reveal cues for penalty shooting analyses. The dominant foot height and the dominant foot angle is also correlated with the height of the shooting in a penalty kick (Higueras-Herbada et al., 2020). Other authors have indicated that the footedness may influence the outcome of the penalty kick (Almeida et al., 2016). Besides that, the run-up approach angle has shown to be important to predict kick directions (Li et al., 2015). On the contrary high ratings were found to other variables, such as run-up fluency, kicking technique, gaze behavior, goalkeeper performance, the moment of the match, match importance, penalty kick outcome, penalty taker, and goalkeeper strategy. The high ratings confirmed previous findings (Kuhn, 1988; van der Kamp, 2006; Noël and van der Kamp, 2012; White and O'Donoghue, 2013; Li et al., 2015; Noël et al., 2015; Almeida et al., 2016). Studies have shown that the run-up pattern differed between strategies. For penalty takers with the keeper-independent strategy (Kuhn, 1988), the run-up seems to be more fluent, and the total run-up and last step distance is longer than for kicks with the keeper-dependent strategy (Noël et al., 2015). According to Kuhn (1988), ball speed could also distinguish strategies. One characteristic that differentiates the penalty takers' profiles is the space-temporal pattern of gaze (Noël and van der Kamp, 2012). Those authors found that those penalty takers who use a keeper-dependent strategy spend more time looking at the goalkeeper throughout the run-up and kick execution than penalty takers who use a keeper-independent strategy. The prevalence of penalty kick strategies can also be mediated by personal or situational factors, including a player's skill or the importance of the kick (Noël et al., 2015).

For intra-rater-reliability, the lowest value found was kappa = 0.75 for the variables non-kicking foot orientation and penalty taker strategy. For inter-rater reliability, the lowest value found was kappa = 0.70 for the variable goalkeeper tactical action. Both values are still good strength of agreement (O'Donoghue, 2010). The low value in the variable non-kicking foot orientation can be explained by the small size of the object of interest compared to the large volume that contains the necessary elements for recording a penalty kick. Although a minimum level of video pixel resolution is required as inclusion criterion for this study, this is not sufficient for precise observations. The variable goalkeeper's tactical action requires the interpretation of behavior and consequently involves a large amount of subjectivity. Despite having clear definitions in the proposed system, judgments here may be influenced by the sporting experience of each observer or the former playing position (e.g., goalkeeper). The same does hold for the other variables with low kappa values (e.g., strategy assessment). Typically, these variables require a subjective interpretation but are so far only accessible with observational methods. However, the values regarding the level of intra- and inter-observer agreement reached in this study showed that the instrument is reliable when used by trained observers, meeting the minimum thresholds proposed by Altman (1991).

A novelty of this study is the inclusion of viewing angle analysis and video standardization. To the best of our knowledge, there is no other study yet that has included this type of specification for notational studies applied to penalty kick analysis in elite football. The description and standardization of the viewing angles and video quality allow the reproducibility of the instrument and reduce human error. Results indicate that for an optimum penalty kick analysis a combination of at least 3 different viewing angles of the same penalty is recommended. The pitch-level viewing angle behind the penalty taker (d), aerial viewing angle behind the penalty taker (c), and pitch-level viewing angle behind the goalkeeper (e) are the best viewing angles for observation analysis according to a panel of 20 experts in sports science. The change of viewing angles plays an important role in notational analysis, as it may influence the perception of the observer. Depending on the positioning and setting of cameras, recordings may literally provide a different view or perspective of human activity, confirming, complementing, or contrasting what the researchers themselves can see Todd et al. (2007) and LeBaron et al. (2018). Additionally, the standardization of video quality is indicated as a prerequisite for notational analysis. Observational studies in football, specifically in the penalty kick situation using video analysis should describe the pattern of viewing angles presentation, the quantity of angles, and video quality, since these settings may have a direct influence on the perception of the observer. The lack of standardization and indication of this information may compromise the analysis and comparisons of different observational studies. However, it is worth mentioning that the choice of the viewing angles might depend on the research question.

The present methodological design containing a three-study concept made it possible to have a practical approach to the proposed instrument through the pilot study with high-level football coaches. The main study, including three different dimensions (i.e., agreement, adequacy, and univocity) ensured that only variables with a high level of concordance, clarity, and relevance were included in the final format of the instrument. Observational systems, such as the present one, are an important methodology to investigate the structure of sports, guide the coaching and training process, design tactical and technical plans, and develop training methods (Lames and Hansen, 2001; Sarmento et al., 2010; Anguera et al., 2011; Villarejo et al., 2014). Effective observational instruments require high validity and reliability standards, for both the design process and their usefulness for gathering data from competitions (O'Donoghue et al., 2017). A major concern is an extent to which the content of each item of the scale reflects the content domain intended to be measured by the item (Randall et al., 2009). In the present study, based on the judgments of 20 experts, only variables that presented an Aiken's *V* value above the cut-off value were included in the OSPAF.

The validation of observational variables for penalty shooting may provide a general description of its technical execution, which allows for detecting the shooter's and the goalkeeper's strategy based on the behavioral variables studied. These variables, once validated, could be employed to complement and validate subjective strategy judgments. Only a small number of performance analysis studies have examined the validation process of observational instruments applied to penalty kicks (e.g., Noël et al., 2015; Comas et al., 2018). The present study controls content validity and reliability (Aiken's *V* and observer agreement) through observable and measurable variables, providing a more holistic and contextual analysis rather than the up to now more analytic and reductionist approaches.

Future research with the application of the OSPAF is needed to identify penalty kick strategies and the relationship between the variables in the system itself. One actual key question is the strategy assessment in in-match penalties, as it is necessary to modeling strategies by direct assessment, and further validation of this model by "soft" observational variables contained in OSPAF. A larger representative study in different leagues and female football can contribute to the identification of the successful and failure profile in penalty kicks across different levels. Additional studies can use the OSPAF applying technological methods to analyze its variables, such as gaze analysis by tracking instruments, computer techniques for body pose estimation, and machine learning-based video analysis.

Despite the possible limitation that the study was conducted via an extensive questionnaire, the OSPAF enables the differentiation of the technical and tactical behavior of the goalkeeper and penalty taker. The present instrument is a comprehensive observational system, which contains the most relevant variables for penalty kick analysis validated by experts. The inclusion of observable variables about the penalty taker actions, goalkeeper, context and outcomes, makes the proposed model more complete than most of the others proposed previously. Reliability has been examined per variable; this standard is sometimes not met in other studies, where the assessment is done for the system as a whole (e.g., Noël et al., 2015). In OSPAF, reliability analysis was conducted through the state of the art for validation of observational instruments introducing Aiken's *V* statistics specially designed for measuring the agreement of several judges. The high methodological rigor of this study consolidates the OSPAF as an instrument that integrates the main variables for penalty analysis in top-level football. Also, evidence for standardization of viewing angles and video quality is presented. Football coaches and match analysts of all levels can use the methodological framework of OSPAF to evaluate and record the penalty kick performance profile of their players throughout the season and using this information for adjusting and improving the coaching process. The final version of the OSPAF is included as a stand-alone supplementary resource, that can be downloaded by researchers and practitioners. For future observational studies of penalties, it is recommended to use OSPAF as a starting point and to add variables specific to the new topic under scrutiny.

CONCLUSION

The OSPAF evidenced content validity, inter- and intra-reliability for analyzing penalty kicks in football, through the use of a gold standard methodology for instrument validation. The present study concludes that the final instrument is adequate and consistent for analyzing successful and non-successful penalty kick patterns. Common statistical requirements for the validation of the system presented were met without exception. There are clear operational definitions in the system, and it can be reproduced reliably. The literature gains a validated tool capable of promoting reliable penalty analysis in elite football and provides new guidelines on the standardization of videos in notational systems.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be available from the authors upon request.

ETHICS STATEMENT

Written informed consent has not been obtained from the individual(s) for the publication of any images or potentially identifiable data included in this article, as the data is publically available.

AUTHOR CONTRIBUTIONS

GP and ML contributed to the conception and design of the study. GP and VN organized the database and performed the statistical analysis. GP, MD, VC, and ML wrote the first draft of the manuscript. GP, ML, and MD wrote sections of the manuscript. All authors contributed to the revision of the manuscript, read, and approved the presented version.

FUNDING

Funding as a doctoral scholarship was provided by the Katholische Akademische Ausländerdienst (KAAD); the publication costs were supported by the Technical University of Munich (TUM) Publishing Fund.

ACKNOWLEDGMENTS

We would like to thank all the experts who participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.661179/full#supplementary-material

REFERENCES

AERA, APA, and National Council on Measurement in Education and Joint Committee on Standards for Educational and Psychological Testing (2014). *Standards for Educational and Psychological Testing*. Washington, DC: AERA.

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educ. Psychol. Meas.* 45, 131–142. doi: 10.1177/0013164485451012

Almeida, C. H., Volossovitch, A., and Duarte, R. (2016). Penalty kick outcomes in UEFA club competitions (2010-2015): the roles of situational, individual and performance factors. *Int. J. Perform. Anal. Sport* 16, 508–522. doi: 10.1080/24748668.2016.11868905

Altman, D. G. (1991). *Practical Statistics for Medical Research*. Boca Raton, FL: CRC press.

Anguera, M. T., Blanco, A., and Losada, J. L. (2001). Observational Designs, a key issue in the observational methodology process (in Spanish). *Metodol Cienc. Comport.* 3, 135–161.

Anguera, M. T., Villaseñor, A. B., Hernández-Mendo, A., and López, J. L. L. (2011). Observational designs: fit and application in sport psychology (In Spanish). *Cuad de Psicol del Deporte* 11, 63–76.

Aranda, R., González-Ródenas, J., López-Bondia, I., Aranda-Malavés, R., Tudela-Desantes, A., and Anguera, M. T. (2019). "REOFUT" as an observation tool for tactical analysis on offensive performance in soccer: mixed method perspective. *Front. Psychol.* 10:1476. doi: 10.3389/fpsyg.2019.01476

Baranowski, A. M., and Hecht, H. (2017). Effect of camera angle on perception of trust and attractiveness. *Empir. Stud. Arts* 36, 90–100. doi: 10.1177/0276237417710762

Bowtell, M., King, M., and Pain, M. (2009). Analysis of the keeper-dependent strategy in the soccer penalty kick. *Int. J. Sports Sci. Eng.* 3, 93–102.

Brewer, C., and Jones, L. (2002). A five-stage process for establishing contextually valid systematic observation instruments: the case of Rugby Union. *Sport Psychol.* 16, 138–159. doi: 10.1123/tsp.16.2.138

Brinkschulte, M., Furley, P., and Memmert, D. (2020). English football players are not as bad at kicking penalties as commonly assumed. *Sci. Rep.* 10:7027. doi: 10.1038/s41598-020-63889-6

Carling, C. H., Wright, C., Lee, J. N., and Bradley, P. (2014). Comment on 'Performance analysis in football: a critical review and implications for future research'. *J. Sports Sci.* 32, 2–7. doi: 10.1080/02640414.2013.807352

Chacón-Moscoso, S., Sanduvete-Chaves, S., Anguera, M. T., Losada, J. L., Portell, M., and Lozano-Lozano, J. A. (2018). Preliminary checklist for reporting observational studies in sports areas: content validity. *Front. Psychol.* 9:291. doi: 10.3389/fpsyg.2018.00291

Chiappori, P. A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: the case of penalty kicks in soccer. *Am. Econ. Rev.* 92:4. doi: 10.1257/00028280260344678

Cobb, N. M., Unnithan, V., and McRobert, A. P. (2018). The validity, objectivity, and reliability of a soccer-specific behaviour measurement tool. *Sci. Med. Football.* 2, 1–7. doi: 10.1080/24733938.2017.1423176

Comas, G. N., Bardavío, J. S., Nuviala, R., and Miguel, D. F. (2018). Design of an observational instrument for the assessment of penalties in football and analysis of the results obtained (In Spanish). *J. Sport Psychol.* 27, 189–199.

Dalkey, N., and Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Manage Sci.* 9, 351–515. doi: 10.1287/mnsc.9.3.458

Dalton, K., Guillon, M., and Naroo, S. A. (2015). An analysis of penalty kicks in elite football post 1997. *Int. J. Sports Sci. Coach.* 10, 815–827. doi: 10.1260/1747-9541.10.5.815

Dayé, C. (2018). How to train your oracle: the Delphi method and its turbulent youth in operations research and the policy sciences. *Soc. Stud. Sci.* 48, 846–868. doi: 10.1177/0306312718798497

Dicks, M., Button, C., and Davids, K. (2010). Availability of advance visual information constrains association-football goalkeeping performance during penalty kicks. *Perception* 39, 1111–1124. doi: 10.1068/p6442

Dicks, M., Davids, K., and Button, C. (2009). Representative task designs for the study of perception and action in sport. *Int. J. Sport Psychol.* 40, 506–524.

Fariña, A. R., Gabriel Fábrica, P., Tambusso, S., and Alonso, R. (2013). Taking the goalkeeper's side in association football penalty kicks. *Int. J. Perform. Anal. Sport* 13, 96–109. doi: 10.1080/24748668.2013.11868634

Fernandes, T., Camerino, O., Garganta, J., Pereira, R., and Barreira, D. (2019). Design and validation of an observational instrument for defence in soccer based on the Dynamical Systems Theory. *Int. J. Sports Sci. Coach.* 14, 138–152. doi: 10.1177/1747954119827283

Fitzpatrick, A. R. (1983). The meaning of content validity. *Appl. Psychol. Meas.* 7, 3–13. doi: 10.1177/014662168300700102

Furley, P., Noël, B., and Memmert, D. (2017). Attention towards the goalkeeper and distraction during penalty shootouts in association football: a retrospective analysis of penalty shootouts from 1984 to 2012. *J. Sports Sci.* 35, 873–879. doi: 10.1080/02640414.2016.1195912

Gabin, B., Camerino, O., Anguera, M. T., and Castañer, M. (2012). Lince: multiplatform sport analysis software. *Procedia Soc. Behav. Sci.* 46, 4692–4694. doi: 10.1016/j.sbspro.2012.06.320

Garcia-Santos, D., and Ibanez, S. J. (2016). Design and validation of an observational instrument to basketball referee evaluation (IOVAB). *Sport TK* 5, 15–26. doi: 10.6018/264601

Garganta, J. (1997). *Tactical Modelling of the Football Match: Study of the Organisation of the Offensive Phase in High Performance Teams.* thesis, in portuguese. Porto: University of Porto.

Gong, B., Cui, Y., Gai, Y., Yi, Q., and Gómez, M. A. (2019). The validity and reliability of live football match statistics from champdas master match analysis system. *Front. Psychol.* 1:1339. doi: 10.3389/fpsyg.2019.01339

Hasson, F., Keeney, S., and McKenna, H. (2000). Research guidelines for the Delphi survey technique. *J. Adv. Nurs.* 32, 1008–1015. doi: 10.1046/j.1365-2648.2000.t01-1-01567.x

Heale, R., and Twycross, A. (2015). Validity and reliability in quantitative studies. *Evid. Based Nurs.* 18, 66–67. doi: 10.1136/eb-2015-102129

Higueras-Herbada, A., Lopes, J. E., Travieso, D., Ibáñez-Gijón, J., Araújo, D., and Jacobs, D. M. (2020). Height after side: goalkeepers detect the vertical direction of association-football penalty kicks from the ball trajectory. *Front. Psychol.* 1:311. doi: 10.3389/fpsyg.2020.00311

Horn, M., Waal, S., and Kraak, W. (2021). In-match penalty kick analysis of the 2009/10 to 2018/19 English Premier League competition. *Int. J. Perform Anal. Sport* 21, 1–17. doi: 10.1080/24748668.2020.1855052

Hughes, M., and Wells, J. (2002). Analysis of penalties taken in shoot-outs. *Int. J. Perform. Anal. Sport* 2, 55–72. doi: 10.1080/24748668.2002.11868261

Instat (2020). *Penalties: Ultimate guidelines; Massive research of penalties by Instat*. Available online at: https://instatsport.com/football/article/penalty_research (accessed May 03, 2021).

Jamieson, S. (2004). Likert scales: how to (ab)use them. *Med. Educ.* 38, 1217–1218. doi: 10.1111/j.1365-2929.2004.02012.x

Jordet, G., Hartman, E., Visscher, C., and Lemmink, K. A. P. M. (2007). Kicks from the penalty mark in soccer: the roles of stress, skill, and fatigue for kick outcomes. *J. Sports Sci.* 25, 121–129. doi: 10.1080/02640410600624020

Kuhn, W. (1988). "Penalty-kick strategies for shooters and goalkeepers," in *Science and Football*, eds T. Reilly, A. Lees, K. Davids, and W. J. Murphy (London: E and FN Spon), 489–492.

Lames, M. (1994). *Systematische Spielbeobachtung (in German, Systematic game observation)*. Münster: Philippka-Sportverlag.

Lames, M. (2006). Modelling the interaction in game sports - Relative phase and moving correlations. *J. Sports Sci. Med.* 5, 556–560.

Lames, M., and Hansen, G. (2001). Designing observational systems to support top-level teams in game sports. *Int. J. Perform. Anal. Sport* 1, 85–91. doi: 10.1080/24748668.2001.11868251

Lames, M., and McGarry, T. (2007). On the search for reliable performance indicators in game sports. *Int. J. Perform. Anal. Sport* 7, 62–79. doi: 10.1080/24748668.2007.11868388

LeBaron, C., Jarzabkowski, P., Pratt, M. G., and Fetzer, G. (2018). An introduction to video methods in organizational research. *Organ. Res. Methods* 21, 239–260. doi: 10.1177/1094428117745649

Lepschy, H., Wäsche, H., and Woll, A. (2020). Success factors in football: an analysis of the German Bundesliga. *Int. J. Perform. Anal. Sport* 20, 1–15. doi: 10.1080/24748668.2020.1726157

Li, Y., Alexander, M., Glazebrook, C., and Leiter, J. (2015). Prediction of kick direction from kinematics during the soccer penalty kick. *Int. J. Kin. Sports Sci.* 3, 1–7.

Lopes, J. E., Araújo, D., Duarte, R., Davids, K., and Fernandes, O. (2012). Instructional constraints on movement and performance of players in the penalty kick. *Int. J. Perform. Anal. Sport* 12, 311–345. doi: 10.1080/24748668.2012.11868602

Mackenzie, R., and Cushion, C. (2013). Performance analysis in football: a critical review and implications for future research. *J. Sport Sci.* 31, 639–676. doi: 10.1080/02640414.2012.746720

Makaruk, H., Porter, J. M., Bodasińska, A., and Palmer, S. (2020). Optimizing the penalty kick under external focus of attention and autonomy support instructions. *Eur. J. Sport Sci.* 20, 1378–1386. doi: 10.1080/17461391.2020.1720829

Moraes, E. L., Cardoso, F., and Teoldo, I. (2014). Analysis of the offensive standards of the Spanish national football team at the 2010 FIFA World Cup® in relation to the "status" of the match (In portuguese). *Rev. Bras. Educ. Fís. Esporte.* 28, 3. doi: 10.1590/1807-55092014000300361

Navarro, M., van der Kamp, J., Ranvaud, R., and Savelsbergh, G. J. (2013). The mere presence of a goalkeeper affects the accuracy of penalty kicks. *J. Sports Sci.* 31, 921–929. doi: 10.1080/02640414.2012.762602

Noël, B., and van der Kamp, J. (2012). Gaze behaviour during the soccer penalty kick: an investigation of the effects of strategy and anxiety. *Int. J. Sport Psychol.* 43, 326–345.

Noël, B., Furley, P., van der Kamp, J., Dicks, M., and Memmert, D. (2015). The development of a method for identifying penalty kick strategies in association football. *J. Sports Sci.* 33, 1–10. doi: 10.1080/02640414.2014.926383

O'Donoghue, P. (2009). *Research Methods for Sports Performance Analysis*. London: Routledge. doi: 10.4324/9780203878309

O'Donoghue, P. (2010). *Research Methods for Sports Performance Analysis*. London: Routledge.

O'Donoghue, P., and Holmes, L. (2014). *Data Analysis in Sport*. London: Routledge.

O'Donoghue, P., Holmes, L., and Robinson, G. (2017). *Doing a Research Project in Sport Performance Analysis*. Oxon: Routledge.

Ortega-Toro, E., García-Angulo, A., Giménez-Egido, J. M., García-Angulo, F. J., and Palao, J. M. (2019). Design, validation, and reliability of an observation instrument for technical and tactical actions of the offense phase in soccer. *Front. Psychol.* 1:22. doi: 10.3389/fpsyg.2019.00022

Paterson, G., van der Kamp, J., and Savelsbergh, G. (2020). Moving advertisements systematically affect gaze behavior and performance in the soccer penalty kick. *Front. Sports Act. Living* 1:69. doi: 10.3389/fspor.2019.00069

Randall, D. P., Peter, R., and Giacobbi, J. (2009). Applying a score confidence interval to Aiken ' s item content-relevance index. *Meas. Phys. Educ. Exerc. Sci.* 8, 213–225. doi: 10.1207/s15327841mpee0804_3

Robinson, G., and O'Donoghue, P. (2007). A weighted kappa statistic for reliability testing in performance analysis in sport. *Int. J. Perform. Anal. Sport* 7, 12–19. doi: 10.1080/24748668.2007.11868383

Sampaio, J., and Leite, N. (2013). "Performance indicators in game sports," in *Routledge Handbook of Sports Performance Analysis*, eds T. McGarry, P. O'Donoghue, and J. Sampaio (London: Routledge), 115–126.

Sarmento, H., Anguera, T., Campaniço, J., and Leitão, J. (2010). Development and validation of a notational system to study the offensive process in football. *Medicina* 46, 401–407. doi: 10.3390/medicina46060056

Sarmento, H., Clemente, F. M., Araújo, D., Davids, K., McRobert, A., and Figueiredo, A. (2018). What performance analysts need to know about research trends in association football (2012–2016): a systematic review. *Sports Med.* 28, 799–836. doi: 10.1007/s40279-017-0836-6

Sarmento, H., Marcelino, R., Anguera, M. T., CampaniÇo, J., Matos, N., and LeitÃo, J. C. (2014). Match analysis in football: a systematic review. *J. Sports Sci.* 32, 1831–1843. doi: 10.1080/02640414.2014.898852

Savelsbergh, G. J. P., Van der Kamp, J., Williams, A. M., and Ward, P. (2005). Anticipation and visual search behaviour in expert soccer goalkeepers. *Ergonomics* 48, 11–14. doi: 10.1080/00140130500101346

Soto, A., Camerino, O., Iglesias, X., Anguera, M. T., and Castañer, M. (2019). LINCE PLUS: research software for behavior video analysis. *Apunts Educ. Fís. Deportes* 137, 149–153. doi: 10.5672/apunts.2014-0983.es.(2019/3).137.11

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. *J. Grad. Med. Educ.* 3, 119–120. doi: 10.4300/jgme-d-11-00075.1

Taherdoost, H. (2016). Validity and reliability of the research instrument: how to test the validation of a questionnaire/survey in a research. *Int. J. Acad. Res.* 5, 28–36. doi: 10.2139/ssrn.3205040

Timmis, M. A., Piras, A., and van Paridon, K. N. (2018). Keep your eye on the ball; the impact of an anticipatory fixation during successful and unsuccessful soccer penalty kicks. *Front. Psychol.* 1:58. doi: 10.3389/fpsyg.2018.02058

Timmis, M. A., Turner, K., and Van Paridon, K. N. (2014). Visual search strategies of soccer players executing a power vs. placement penalty kick. *PLoS One* 9: e115179. doi: 10.1371/journal.pone.0115179

Todd, J. T., Thaler, L., Dijkstra, T. M. H., Koenderink, J. J., and Kappers, A. M. L. (2007). The effects of viewing angle, camera angle, and sign of surface curvature on the perception of three-dimensional shape from texture. *J. Vis.* 7, 1–16. doi: 10.1167/7.12.9

van der Kamp, J. (2006). A field simulation study of the effectiveness of penalty kick strategies in soccer: late alterations of kick direction increase errors and reduce accuracy. *J. Sports Sci.* 24, 467–477. doi: 10.1080/02640410500190841

Van Teijlingen, E. R., and Hundley, V. (2001). The importance of pilot studies. *Soc. Res.* 35, 33–63. doi: 10.7748/ns.16.40.33.s1

Vázquez-Diz, J. A., Morillo-Baro, J. P., Reigal, R. E., Morales-Sánchez, V., and Hernández-Mendo, A. (2019). Design and validation of an observational instrument for goalkeepers in beach handball. *Cuad. de Psicol. Deporte* 19, 135–146. doi: 10.6018/cpd.368901

Vega Marcos, R., Ruiz Barquín, R., Valle, S., and Valle Díaz, M. (2010). Tendencia de acción de porteros de fútbol profesional: el caso de los penaltis. *Cuad. de Psicol. del.* 10, 23–32.

Villarejo, D., Ortega, E., Gómez, M. Á., and Palao, J. M. (2014). Design, validation, and reliability of an observational instrument for ball possessions in rugby union. *Int. J. Perform. Anal. Sport* 14, 955–967. doi: 10.1080/24748668.2014.11868771

Weigelt, M., and Memmert, D. (2012). Goal-side selection in soccer penalty kicking when viewing natural scenes. *Front. Psychol.* 3:312. doi: 10.3389/fpsyg.2012.00312

White, S., and O'Donoghue, P. (2013). "Factors influencing penalty kick success in soccer," in *Science and football VII: The Proceedings of the Seventh World Congress on Science and Football*, eds H. Nunome, B. Drust, and B. Dawson (London: Routledge), 237–242.

Wood, G., Jordet, G., and Wilson, M. R. (2015). On winning the "lottery": psychological preparation for football penalty shoot-outs. *J. Sports Sci.* 33, 1758–1765. doi: 10.1080/02640414.2015.1012103

Wunderlich, F., Berge, F., Memmert, D., and Rein, R. (2020). Almost a lottery: the influence of team strength on success in penalty shootouts. *Int. J. Perform. Anal. Sport* 20, 1–13. doi: 10.1080/24748668.2020.1799171

Zuo, G., Li, Z., Gong, D., and Sun, R. (2010). "Penalty kick strategy using two-player collaboration method," in *Proceedings of the 2010 International Conference on Intelligent Computing and Integrated Systems* (Piscataway, NJ: IEEE), 705–708. doi: 10.1109/ICISS.2010.5655525

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Penalty kicks in elite football: identifying factors related to the player strategy

Conference Paper · September 2021

3 authors:

Guilherme Pinheiro
Technische Universität München

Varley Teoldo da Costa
Federal University of Minas Gerais

Martin Lames
Technische Universität München

INTERNATIONAL JOURNAL OF SPORT AND EXERCISE PSYCHOLOGY
2021, VOL. 19, NO. S1, S1–S539
https://doi.org/10.1080/1612197X.2021.1982479

Routledge
Taylor & Francis Group

Check for updates

International Journal of Sport and Exercise Psychology (IJSEP)

International Society of Sport Psychology (ISSP) 15th World Congress Proceeding

Taipei, Taiwan

September 30 – October 4, 2021

Robert J. Schinke and Yu-Kai Chang

Contents

World Congress Presidents and Committees

ISSP President Robert Schinke, Laurentian University, Canada
World Congress President Yu-Kai Chang, National Taiwan Normal University, Taiwan

ISSP Scientific Committees

Chair: Robert Schinke, Laurentian University Canada, Canada
Young-Ho Kim, Seoul National University of Science and Technology, South Korea
Artur Poczwardowski, University of Denver, United States
Tatiana Ryba, University of Jyväskylä, Finland
Thomas Schack, Bielefeld University, Germany
Gangyan Si, Hong Kong Sports Institute, China

Local Scientific Committees

Chair: Yu-Kai Chang, National Taiwan Normal University

Sport Psychology

Feng-Tzu Chen, China Medical University
Li-Kang Chi, National Taiwan Normal University
Yi-Hsiang Chiu, Chinese Culture University
Ya-Wen Hsu, National Chiayi University
Chung-Ju Huang, University of Taipei
San-Fu Kao, National Tsing-Hua University
Shih-Chun Kao, Purdue University
Chu-Min Liao, National Taiwan Sport University
Jing-Horng Lu, Chinese Culture University
Chiao-Lin Nien, National Taiwan University of Sport
Chun-Chih Wang, Soochow University

Exercise Psychology

Yu-Kai Chang, National Taiwan Normal University
Chien-Heng Chu, National Taiwan Normal University
I-Hua Chu, Kaohsiung Medical University
Chiao-Ling Hung, National Taiwan University
Tsung-Min Hung, National Taiwan Normal University
Po-Wen Ku, National Changhua University of Education
Chia-Liang Tsai, National Cheng Kung University
Chun-Hao Wang, National Cheng Kung University

Motor Learning and Control

Chih-Hui Chang, National Kaohsiung Normal University
Ching-Er Lin, National Taichung University of Education
Ju-Han Lin, National Dong Hwa University
Su-Yen Liu, National Chung Cheng University
Yeou-The Liu, National Taiwan Normal University

ISSP Executive Committee

ISSP RECOGNITION

Former Sites of the ISSP World Congress

2021 Taipei, Taiwan
2017 Seville, Spain
2013 Beijing, China
2009 Marrakesh, Morocco
2005 Sydney, Australia
2001 Skiathos, Greece
1997 Netanya, Israel
1993 Lisbon, Portugal
1989 Singapore
1985 Copenhagen, Denmark
1981 Ottawa, Canada
1977 Prague, Czechoslovakia
1973 Madrid, Spain
1968 Washington D.C, USA
1965 Rome, Italy

President

2017-2021 Robert Schinke
2013-2017 Gangyan Si
2009-2013 Sidónio Serpa
2005-2009 Dieter Hackfort
2001-2005 Keith Henschen
1997-2001 Gershon Tenembaum
1993-1997 Atsushi Fujita
Dennis Glencross
1989-1993 Robert Singer
1985-1989 Robert Singer
1981-1985 Miroslav Vanek
1977-1981 Miroslav Vanek
1973-1977 Miroslav Vanek
1968-1973 Ferruccio Antonelli
1965-1968 Ferruccio Antonelli

Distinguished International Sport Psychologist Award

2021 Natalia Stambulova
2017 Dieter Hackfort
Keith Henschen
2013 Gershon Tenenbaum
2005 Jurgen Nitsch
2001 Bruce C. Ogilvie
1997 Herrmann Rieder
Robert Singer
1993 Erna Geron
1989 Ferruccio Antonelli
Miroslav Vanek

Honor Award

2021 Jitendra Mohan
2017 Bola Ikulayo
Athanasios Papaioannou
2013 Guido Schilling
2009 Yannis Zervas
2001 Dieter Hackfort
Terry Orlick
1997 Glyn Roberts
Gershon Tenenbaum
1993 Atsushi Fujita
John Salmela
1989 Filip Genov
Dorothy Harris
Paul Kunath
Robert Singer
Harold Thomas Anthony Whiting

Developing Scholar Award

2021 Garry Kuan
Noora Ronkainen
2017 Amy Theresa Blodgett
Chun-Qing Zhang
2013 Yu-Kai Chang
Kristoffer Henriksen
2009 Tatiana Ryba
2005 Hulya Asci
Sian Beilock
2001 Andrew M. Williams
1997 Britton Brewer
Judy van Raalte

Developing Practitioner Award

2021 Soek Hau

Fellow

2020 Jean Cote

ABSTRACT
Self-management is defined as people's capacity to effectively monitor, control, and manage thoughts, emotions, and behaviours that could facilitate goal accomplishment and has shown to facilitate in enhancing one's psychological state and performance in sports. Martial arts practitioners are found employing self-management during training and competition to assist in monitoring, evaluating, and controlling the mental, emotional, technical, and physical dimensions of sports performance toward desired goals particularly in enhancing their psychological state. However, despite the multitude of studies verifying the relationships between martial arts athletes' self-management and psychological outcomes, no attempt has been made to systematically investigate these variables and consolidate the findings in the literature through a meta-analytic approach. This study examined the relationship between self-management and exercise self-confidence, satisfaction, and commitment in both modern and traditional martial arts among Korean practitioners using meta-analysis approach. It also investigated the level of sports participation and different martial arts sports as potential moderating variables. In total, 22 published Korean studies yielded 299 individual effect sizes and were included in the final meta-analytic pool. The included studies featured boxing, judo, fencing, wrestling, archery, kumdo, taekwondo, and Ssireum-a Korean traditional martial art with similarities with judo and wrestling. The Comprehensive Meta-Analysis (CMA) version II program was used to accurately estimate the effect sizes from the selected studies. Results revealed that the relationships between self-management and exercise satisfaction and self-confidence were moderate, whereas the relationship between self-management and exercise commitment was large. The effect of the training dimension of self-management was found to be the largest on exercise commitment while that of the mental dimension was second largest on exercise commitment. The influence of self-management on - athletic satisfaction was highest in Judo athletes, exercise commitment was largest in Ssireum, and exercise self-confidence was highest in wrestling. Finally, the use of self-management strategies was relatively less effective for elite practitioners compared with non-elite participants. The findings underscore the effectiveness of self-management strategies to enhance Korean martial arts practitioners' exercise self-confidence, commitment, and satisfaction. It also highlights the importance of the promotion of interventions and educational programs on how to incorporate/employ self-management in athletes' sports training. Other findings about self-management and psychological states of Martial arts players and their implications will be discussed further.

KEYWORDS
self-regulation; Korean traditional sports; martial arts; athlete's self-management questionnaire; psychological state

Penalty kicks in elite football: identifying factors related to the player strategy

Guilherme de Sousa Pinheiro[a], Varley Teoldo Costa[b] and Martin Lames[a]

[a]Technical University of Munich; [b]Federal University of Minas Gerais

ABSTRACT

Introduction: The analysis of penalty kick performance in football has played an important role in sports science (Paterson et al., 2020). There are two main strategies for taking a penalty, the keeper-independent strategy, and the keeper-dependent strategy (Kuhn, 1988). One main characteristic that differentiates the two strategies is the Spatio-temporal pattern of gaze (Noël & van der Kamp, 2012). Strategic, anticipatory, attention, and perception-based factors can also influence the success of the penalty kick (Memmert & Noel, 2020). Several studies have investigated the penalty kick strategies, but the results are limited due to the experimental design, which did not reflect the unfolding of an interaction between the two players. Objective: This study aimed to identify the relationship between a set of observable variables and the players' strategy using an expert-validated observational system for penalty kick analysis in football (OSPAF). Methods: The dataset consists of 150 penalty kicks from the main European football leagues (Premier League, Ligue 1, Bundesliga, LaLiga, Serie A, and Champions League; seasons 2017 to 2020). The videos were recorded by TV broadcasters and were registered and analyzed post-event. All data was annotated by two experienced researchers using the OSPAF. Cohen's kappa (κ) was utilized to verify the reliability of OSPAF (inter-observers and intra-observer). Logistic regression (enter method) analyses were performed. Dimensions and categories of OSPAF were coded in Lince software (Gabin et al., 2012; Soto et al., 2019). The p level considered < 0.001 (95% confidence interval). All data were analyzed using JASP software (Team, 2020; JASP Version 0.14; Computer software). Results: The OSPAF kappa values showed very good strength of agreement (0.95 and 0.92). The run-up speed slow, run up fluency running with pauses, penalty taker gaze behavior not at the ball, and the deception performed by the penalty taker were related to the goalkeeper dependent strategy (χ^2 (145) = 130.596, 86.5% correct classifications, $p < .001$). The model correctly classified 86.5% of cases. The goalkeeper tactical action guess was related to the goalkeeper independent strategy (χ^2 (148) = 137.680, 87.5% correct classifications, $p < .001$). Conclusion: OSPAF variables (e.g. run-up speed and fluency, penalty taker gaze behavior, and deception by penalty taker) were able to classify correctly 86.5% of cases on the likelihood of the penalty taker strategy. OSPAF may serve as a standard tool for observational investigations of penalties in football to make the results from different studies more comparable.

KEYWORDS
attention; performance analysis; football psychology; gaze behavior; deception

Lessons from an emerging practitioner working in foreign settings

Saqib Deen

Md Performance Psychology

ABSTRACT

This oral presentation is a reflexive account, which explores the personal journey of an emerging practitoner who practices in a foreign setting, on track for the Tokyo Olympics. The oral presentation will encapsulate a unique experience sheds a light on interational sport psychology, in which the practitioner who hails from the UK, and who is of a mixed African heritage, now operates a full-time consultancy in South East Asia in a highly diverse cultural setting. It is hoped that this oral presentation will provide a unique take on practicing sport science abroad outside of a practitioners home country, which is of a timely discussion given the ever increasing rates of globalisation.The first section of the presentation will include an introduction to the practitioner, his demographics (ethnicity, nationality, race) and journey into sport psychology. This will be quickly followed up by the practitioners development of philosophy and practicing competencies, such as training in REBT/CBT and exploring methods beyond traditional mental skills training and utilising organisational psychological principles. The presentation will then explore feelings of being a fraud, as is common with emerging practitioners (aligned to the Dunning Kruger effect), and how true practitioning confidence was built on the back of many initial failures and mistakes in practice. This includes issues with cultural assimilation and difficulties assimilating in well-established teams. Following from this, the practitioner will discuss how lessons taken from the ground up, support from supervisors, and adjusting RECBT to fit the culture assisted in producing real and tangible outcomes, and how practice accounted for diversity in this sense. This section would further explain how the practitioner began to adopt a higher reliance on learning foreign languages, reading microexpressions, and heightened awareness of body language; as these strategies were used to take into account the present cultural factors (such as collectivism and passive communication styles which are highly prelevant in this setting). Paying close attention to these factors and adopting a high amount of constant reflection have resulted in the practitioner evolving and having a much higher success rate with clients in terms of outcome and relationship management. The presentation will then move on to how sport psychology was implemented at the first major games for the practitioner (the Tokyo Olympics), and what the next areas of advancement are for the practitioner to grow and develop to a more successful practice.

KEYWORDS
applied sport psychology; cultural sport psychology; south east asia; reflexive sport psychology; contextual sport psychology

Effect of Emotional Freedom Technique (EFT) on heart rate, blood pressure and performance in national level shooters

Shivam Dwivedi[a], Akshita Sekhon[a] and Bhawna Chauhan[b]

[a]National Sports University, India; [b]Research Scholar

frontiers in Sports and Active Living

ORIGINAL RESEARCH
published: 10 March 2022
doi: 10.3389/fspor.2022.818556

Body Pose Estimation Integrated With Notational Analysis: A New Approach to Analyze Penalty Kicks Strategy in Elite Football

Guilherme de Sousa Pinheiro[1]*, Xing Jin[2], Varley Teoldo Da Costa[3] and Martin Lames[1]

[1] Performance Analysis and Sports Informatics, Technical University, Munich, Germany, [2] Department of Informatics, Technical University of Munich, Munich, Germany, [3] UFMG Soccer Science Center, Department of Sports Science, Federal University of Minas Gerais, Belo Horizonte, Brazil

Body orientation of football players has proven to be an informative resource related to successful penalty kicks. OpenPose is one of the most popular open-source pose estimation technologies. This study aims: (i) to verify whether OpenPose can detect relevant body orientation angles from video data of penalty kicks in elite football and (ii) to investigate the relationship between these body angles and observable behaviors analyzed *via* an observational system for penalty kick analysis in football (OSPAF) with the penalty taker and goalkeeper strategy. A total of 34 penalty videos, with standardized viewing angle, from the main European leagues (2017–2020) were analyzed. Relevant body orientation variables were selected for penalty kicks analysis and were extracted from video data through OpenPose technique. The OSPAF, previously validated by experts, was used. The mean confidence score of OpenPose measures was 0.80 ± 0.14. OpenPose Retest reliability values was 0.976 ± 0.03. Logistic regressions were performed to investigate the relationship between OpenPose investigated variables (penalty taker: shoulder, hips, and nonkicking foot orientation; goalkeeper: right and left foot, anticipation), observable behaviors (OSPAF variables), and the strategy (penalty taker: goalkeeper dependent or independent; goalkeeper: shooter dependent or independent) in penalty kicks. The selected body orientation angle (goalkeeper anticipation) measured through OpenPose correlated significantly with the goalkeeper strategy. The prediction model of the goalkeeper's strategy had its accuracy increased to 97% when the variable goalkeeper anticipation was included [$\chi^2_{(35)} = 49.648$, $p < 0.001$]. Lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and run up speed (slow) were associated with a kicker-dependent strategy. Regarding the penalty taker, the selected body angles measured through OpenPose did not associate significantly with the shooter strategy. Body orientation analysis by using OpenPose has shown sufficient reliability and provides practical applications for analyzing the strategies adopted by goalkeepers in penalty kicks in elite football.

Keywords: body orientation, performance analysis, OSPAF, OpenPose, human movement, motion capture, soccer analytics

OPEN ACCESS

Edited by:
Zimi Sawacha, University of Padua, Italy

Reviewed by:
Matteo Zago, University of Milan, Italy
Fabiola Spolaor, University of Padua, Italy

Correspondence:
Guilherme de Sousa Pinheiro
guilherme.pinheiro@tum.de

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: *19 November 2021*
Accepted: *26 January 2022*
Published: *10 March 2022*

Citation:
Pinheiro GS, Jin X, Costa VTD and Lames M (2022) Body Pose Estimation Integrated With Notational Analysis: A New Approach to Analyze Penalty Kicks Strategy in Elite Football. Front. Sports Act. Living 4:818556. doi: 10.3389/fspor.2022.818556

INTRODUCTION

The analysis of penalty kick performance in football has played an important role in sports analytics (Paterson et al., 2020; Noël et al., 2021; Pinheiro et al., 2021a,b). Over the past 30 years, there have been several scientific studies that identify the motivational-, strategic-, anticipatory-, attention-, and perception-based factors that can mean a successful or failed penalty kick (Memmert and Nöel, 2020). Recent research focusing on the technical dynamics of penalty kicks has also identified multiple key variables that can differentiate the players strategy (Pinheiro et al., 2021b) and enhance the overall chances of scoring a penalty kick (Jamil et al., 2020). The importance of the optimal performance of both the rival players during the penalty kick is paramount, especially since the introduction of the penalty shoot-out in major competitions to determine which team progresses after a drawn match (Fariña et al., 2013).

One prerequisite to increase the probability of successful performance is the implementation of the suitable penalty kick strategy (van der Kamp, 2006). Previous research has identified two main strategies for taking a penalty (Kuhn, 1988; van der Kamp, 2006). First, the keeper-independent strategy, where the kicker selects the target location to shoot toward before the run-up and does not attend to the actions made by the goalkeeper during the run-up. The decision of where to aim depends on the penalty taker's kicking preference (Noël et al., 2015). On the contrary, in the keeper-dependent strategy, the kicker tries to obtain information from the goalkeeper's reactions during the run-up. Nevertheless, the outcome of a penalty is determined by an interaction between the shooter's strategy (e.g., technique, speed) and the goalkeeper's strategy (Hunter et al., 2018; Pinheiro et al., 2021b). The optimal strategy depends on the keeper's behavior and the relative benefits of speed, accuracy, and unpredictability within each situation. Regarding the goalkeeper strategy, there are two approaches: the dependent and independent penalty takers. The goalkeeper who behaves according to the first group defines his movement based on the actions of the penalty taker. The second type of goalkeeper is the one who risks jumping to a corner independently of the kicker's movement (Kuhn, 1988).

The analysis of the penalty kick strategies has been investigated about numerous factors (Noël et al., 2015; Pinheiro et al., 2021b). Noël et al. (2015) developed a method for investigating penalty taker strategies, based on a controlled simulated situation. In a noncompetitive setting, youth players were instructed to take penalty kicks adopting either a keeper-independent or keeper-dependent strategy. Based on this setting, an observational system was developed to evaluate penalty kick performances by using video footage from competitive matches. Those authors identified that attention to the goalkeeper, run-up fluency, and kicking technique in combination could predict kick strategy in 92% of the penalties. However, one possible limitation is that the penalty takers followed a script denoting whether they use a keeper-independent or keeper-dependent strategy and, therefore, the design created differed very importantly from the match situation (Pinheiro et al., 2021a). Besides that, it remains unclear whether the young players disposed of a sufficient skill level to execute both the strategies with the same quality. To address the interaction process in professional football and provide a valid instrument, (Pinheiro et al., 2021b) developed an observational system for penalty kick analysis in football (OSPAF). The OSPAF met all the requirements of instrument validation.

Body orientation has been indicated as a key factor under covering the success in penalty kicks (Li et al., 2015). However, it is a yet little explored area in penalty kick analytics. There is a need within human movement sciences for a markerless motion capture system, which is easy to use and sufficiently accurate to evaluate motor performance (Nakano et al., 2020). OpenPose method adopts unique top-down position recognition by using deep learning and also the unique algorithm as affiliation recognition of body parts by Part Affinity Fields (PAFs) to detect the two-dimensional (2D) pose of multiple people in images (Nakai et al., 2019). OpenPose can recognize skeletons of multiple players in real-time, by using a simple web camera. Given a video or image, OpenPose estimates a total of 25 biometric human body parts (e.g., right knee, left knee, and right foot). The output of the algorithm is in the form of 25 × 3 vector for each individual, where the first two columns of the vector stand for the x-y coordinate of key points in the field domain, while the third column represents the confidence score. This method has shown high-level accuracy on multiple public benchmarks, being efficient for multiperson pose estimation (Cao et al., 2017). Zago et al. (2020) confirmed the feasibility of tracking kinematics by using OpenPose. OpenPose-based markerless motion capture can be used for human movement science with an accuracy of 30 mm or less (Nakano et al., 2020). Despite several studies in this area, key gaps remain, including a lack of research by using OpenPose to detect relevant body orientation angles in field settings and based on sports broadcasts such as penalty kicks from TV videos.

Sangüesa et al. (2019) had previously applied OpenPose to estimate the body orientation of football players from video data during match play. Those authors indicated that a time-based set of player orientations might detect specific situations where orientation is crucial in the match. Recently, Sangüesa et al. (2020a) used a player's body orientation to model pass feasibility in football. The inclusion of the orientation data estimated directly from video frames by using pose models, into a passing model, has proved to be a key feature in the decision-making process of players and is strictly correlated to the play outcome. In another study, Sangüesa et al. (2020b) mapped body pose parts (e.g., shoulders and hips) in a 2D field by combining OpenPose with a super-resolution network and merging the obtained estimation with contextual information (ball position). Results have been validated with players held electronic performance and tracking systems devices, obtaining a median error of 27° per player.

Notation analysis has been widely used to examine the technical properties of football performance through recording behavior incidence (Lames and Hansen, 2001; Hughes and Bartlett, 2004; Sarmento et al., 2014; Casal et al., 2017; Pinheiro et al., 2021b). In the recent years, there has been a vertiginous evolution in the match analysis methods, mainly motivated by the emergence of automatic registration procedures, which allows the immediate acquisition of a large amount of data related to the positioning of the players with the game (Castellano et al., 2014). The rise of sports analytics has provided a new set of metrics and statistics that can serve coaches to evaluate the player (Sangüesa et al., 2019). Nevertheless, one limitation is that one method does not entirely supply all the necessary information. There is, therefore, a need to use multimethod approach to solve sports analytics problems, analyzing variables by using different methods (Aranda et al., 2019). Methodology designs that combine different study approaches (e.g., observational and biomechanical/method that produce body angles), also known as mixed methods (Preciado et al., 2019), tend to provide a deeper understanding and reliability of the studied phenomenon (i.e., penalty kicks).

The influence variables on penalty kicks success are extensively studied (Jamil et al., 2020; Memmert and Nöel, 2020; Paterson et al., 2020; Noël et al., 2021; Pinheiro et al., 2021b). (Pinheiro et al., 2021b) recommended that future studies could use the OSPAF, applying technological methods to analyze its variables, such as computer techniques for body pose estimation and machine learning-based video analysis. To the best of our knowledge, no study has used OpenPose to detect relevant body orientation angles in penalty kicks in elite football from TV broadcast. Therefore, the aims of this study are: (i) to verify whether OpenPose can detect relevant body orientation angles from video data of penalty kicks in elite football and (ii) to investigate the relationship between these body angles and observable behaviors analyzed via OSPAF (Pinheiro et al., 2021b) with the penalty taker and goalkeeper strategy.

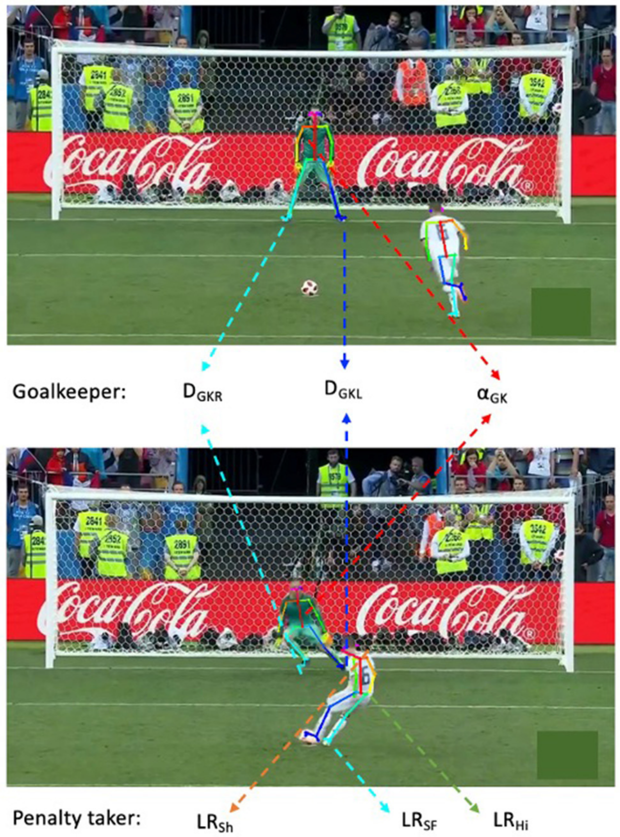


FIGURE 1 | Penalty kick viewing angle, frames analyzed, and process of pose estimation. The upper image corresponds to the moment when the penalty taker starts the run-up approaching the ball, and the down one corresponds to the moment when he touches the ball. The coordinates belonging to shoulders, hips, and non-kicking foot of penalty taker are mapped in a 2D field. LR-side Booleans (LR_{Sh}, LR_{Hi}, LR_{SF}: penalty kicker's shoulder, hips, and non-kicking foot orientation, respectively), angles (α_{Sh} α_{Hi}, α_{SF}: penalty kicker's shoulder, hip, and non-kicking foot, respectively) and confidences (C_{Sh}, C_{Hi}, C_{SF}: confidence score of the orientation of penalty kicker's shoulder, hips and non-kicking foot, respectively). Coordinates belonging to the neck and hip of the goalkeeper at these two moments are mapped to a vector and the angle is calculated (α_{GK}). D_{GKR} and D_{GKL}, the movement distance of right and left foot; C_{GK}, represent confidence score.

MATERIALS AND METHODS

Sample

The dataset consists of 34 penalty kicks from the main European football leagues (Premier League, Ligue 1, Bundesliga, LaLiga, Serie A, and Champions League; seasons 2017–2020). The videos were recorded from TV broadcasters and were registered and analyzed postevent. As the video recordings were public, confidentiality was not an issue and authorization was not required from the players observed or their representatives. The procedures performed in this study were in strict accordance with the Declaration of Helsinki as well as with the ethical standards of the Technical University of Munich.

Methodological Design

All the penalty kick data were annotated by the researchers with the OSPAF (Pinheiro et al., 2021b). Body orientation was analyzed by using OpenPose (CMU-Perceptual-Computing-Lab, 2017). The choice and analysis of the penalty kick video viewing angle was standardized (Pinheiro et al., 2021b), with a pixel resolution of 1,280 × 720. The viewing angle used in this study was the view behind the penalty taker (**Figure 1**). The confidence score, calculated by OpenPose, was used to evaluate reliability (Sangüesa et al., 2019). In order to check the stability within the observation, every penalty kick was analyzed with OpenPose twice. Retest reliability was utilized to check these repeated measurements.

Body Pose Detection and Orientation

OpenPose (version 1.4.0) was installed from GitHub (CMU-Perceptual-Computing-Lab, 2017) and run with a notebook (Apple's M1 Chip) under default settings. Orientation from pose used pretrained models and three-dimensional (3D) vision techniques to obtain a first orientation estimation of each player. Once the pose is extracted for each player, the coordinates and confidence level associated with the body parts are stored to estimate the pose orientation. As a result, in the moving skeletal pictures generated by OpenPose, the skeleton marks are shown and overlapped well with the figure of players (Nakai et al., 2019). For technical details of pose models, see Ramakrishna et al. (2014), Wei et al. (2016), and Cao et al. (2017).

In this study, the orientation of a player's body was defined as the 2D rotation of the player's upper torso around the vertical axis, which is assumed to coincide with the field projection of a normal vector placed in the center of their upper torso, involving both the shoulders and hip parts (Sangüesa et al., 2019). Especially in the case of the non-kicking foot, the hallux and the fifth toe of the support foot were used as the left-right (LR) pair to find the normal vector. Orientation was measured in degrees. For technical details of this methodological approach, see Sangüesa et al. (2019, 2020a,b).

In this study, two frames were analyzed. First, when the penalty taker starts the run-up into the ball and, second, when he touches the ball (**Figure 1**). Then, the target variables for the penalty taker (nonkick foot orientation, hips, and shoulders) and the goalkeeper [anticipation movement (explained in detail below) and right and left foot orientation] were extracted. There might be blurry frames and overlap of players. OpenPose could then fail to detect the main biometric body parts of the two players involved in this analysis; therefore, in this case, the neighboring frames, in which biometric body parts can be detected, were used.

Once the pose was extracted for the goalkeeper and penalty taker, the direct linear transformation (DLT) algorithm (Hartley and Zisserman, 2004) was used to map the coordinate information of players into a 2D field with a homography, given the 4 field corners' coordinates in the image (or its projection out of the image in the nonvisible cases). The homography was first calculated based on four 2D-to-2D point correspondences between the frames (Equation 1). From the output of OpenPose, the coordinates of the main upper-torso parts are found in the image domain; by mapping the LR pair (either shoulders or hips) in the 2D field, a first insight of the player orientation is obtained. The player can be inclined toward the right (0–90 and 270–360°) or the left (90–270°) side of the field.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \alpha H \begin{bmatrix} x \\ y \\ w \end{bmatrix}, \textit{where homography } H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (1)$$

After that, the 2D field projections of the LR pair of penalty taker's shoulders, hips, and nonkicking foot (big toe and small toe) were calculated. All the body parts' orientations could point to the left or right half, based on the angle system presented by Sangüesa et al. (2019, 2020a,b). Based on the 2D projection, LR-side Booleans ($LR_{Sh,}$ $LR_{Hi,}$ and LR_{SF}: penalty kicker's shoulder, hips, and non-kicking foot orientation, respectively), angles (α_{Sh} α_{Hi}, and α_{SF}: penalty kicker's shoulder, hip, and non-kicking foot, respectively), and confidences (C_{Sh}, C_{Hi}, and C_{SF}: confidence score of the orientation of penalty kicker's shoulder, hip, and non-kicking foot, respectively) were obtained. The corresponding confidences are the average of OpenPose's player toes, shoulders, and hips confidences, respectively. **Figure 1** shows the output of OpenPose on which the key biometric body parts of an individual are detected, illustrating the estimation process of orientation.

Anticipation Movement of the Goalkeeper

The anticipation movement of the goalkeeper in the penalty kick was defined here as to how far the goalkeeper moves between: (1) the moment when the penalty taker starts the run-up approaching the ball and the (2) moment when the penalty taker first touches the ball. In detail, the line formed by the connection between the goalkeeper's neck and the middle of the hip was used to depict the position status of the goalkeeper in these two moments. Furthermore, the angle (αGK) between the two lines drawn from the two moments measures the anticipation movement of the goalkeeper. The confidence level for this measure (C_{GK}) was calculated by the average confidence scores of the neck and middle of the hip. This process is given in **Figure 1**.

The movement distance of the goalkeeper's left and right foot was also used to measure the anticipation movement. Left and right ankles were used to represent the left and right feet, respectively; moreover, coordinate information together with metric Euclidean distance was used to depict the movement distance of the goalkeeper's feet, as shown in **Figure 1**.

Ball Speed

Ball speed was determined with the open-source software program Kinovea motion analysis (version 0.8.15, Kinovea, France). This software has already been used in various studies analyzing penalty kicks (Hunter et al., 2018; Makaruk et al., 2019).

Notational Analysis

A previously developed and validated observational system (OSPAF) for penalty analysis in elite football was also used in this study (Pinheiro et al., 2021b). The protocols for the use of observational systems were adopted (Lames and Hansen, 2001; Aranda et al., 2019; Fernandes et al., 2019). All the observable behaviors recorded are shown in **Table 1**.

Data Analysis

For descriptive analysis, mean and SD were used. The Shapiro–Wilk test was performed to verify data normality. The association level between the OSPAF variables with the penalty taker and goalkeeper strategy was determined with the use of the chi-squared (χ^2) test. The effect size was determined by using the Cramer's V and classified as weak (ES $\leq$ 0.2), moderate (0.2 $<$ ES $\leq$ 0.6), and strong (ES $>$ 0.6) (Cohen, 1988). The association level between OpenPose variables with the penalty taker and goalkeeper strategy was determined with the use of the point-biserial correlation. Retest reliability was utilized to check the repeated measurements of OpenPose (Vilagut, 2014). Test-retest reliability coefficients (also called coefficients of stability) vary between 0 and 1, where 1: perfect reliability, $\geq$ 0.9: excellent reliability, $\geq$ 0.8 $<$ 0.9: good reliability, $\geq$ 0.7 $<$ 0.8: acceptable reliability, $\geq$ 0.6 $<$ 0.7: questionable reliability, $\geq$ 0.5 $<$ 0.6: poor reliability, $<$ 0.5: unacceptable reliability, and 0: no reliability (Vogt, 2005; Lindstrom, 2010). To identify which variables would be able to predict the penalty takers and goalkeeper strategy, the logistic regression (*enter method*) analyses were performed. Dimensions and categories of OSPAF were coded in Lince software (**Figure 2**; Gabin et al., 2012;

TABLE 1 | OSPAF variables.

Variables	Definition	Attribute levels
Run up speed	Running speed of the penalty kicker toward the ball	Fast or slow
Run up fluency	Characteristic of the penalty kicker's run during the approach of the ball, with or without pauses.	Continuous running or running with pauses
Run up approach angle	Penalty kicker's running angle to the ball.	Frontal or diagonal
Number of steps	Number of steps of the penalty kicker until contact with the ball	1–3; 3–5; or +5
Kicking technique	The technique used by the penalty kicker to kick the ball	Side foot kick or instep kick
Foot used to kick	Foot used by the penalty kicker to kick the ball	Right or left
Penalty taker gaze behavior	Gaze behavior of the kicker during the approach run.	Gaze at the ball or not at the ball
Goalkeeper (GK) initial posture	Position of the body segments.	Arms raised; arms down or arms extended in a position perpendicular to the goalkeeper 's trunk
Deception by the penalty taker	Indication if the kicker has done any action to distract the goalkeeper during his or her run-up	Yes or no
Goalkeeper tactical action	General evaluation of the way the goalkeeper acted during the penalty shoot-out, to the anticipatory aspect	Try to guess the location of the shot; or awaiting the penalty taker action
Goalkeeper performance	Evaluation of the goalkeeper's performance according to his movement and contact with the ball	0: GK made any final movement to the side of the goal opposite to the final ball location; 1: GK did not move from the center of the goal; 2: GK made a movement in the correct direction but did not dive and failed to make contact with the ball; 3: GK dived in the correct direction but failed to make contact with the ball; 4: GK dived in the correct direction and contacted the ball without saving it; or 5: GK successfully saved the kick
Moment of the match	Time of the match when the penalty will be taken	First half; second half or extra time or shoot out
Location of the match (kicker point of view)	Indication if the penalty kicker is from the home team, visitor, or if he plays on a neutral field.	Home, neutral or away
Momentary result (kicker point of view)	Result of the match (for the penalty kicker) at the moment the penalty was marked.	Winning, drawing or losing
Momentary result (GK point of view)	Result of the match (for the Goalkeeper) at the moment the penalty was marked.	Winning, drawing or losing
Match importance	Level of importance of the match for the team	Championship final match; decisive knockout match; group stage match; early season game; match in final stages of the season
Penalty kick direction	The direction of the ball on goal	Left; center or right
Penalty kick height	Height of the ball on goal	Upper; center or down
Penalty kick outcome	Result of the penalty kick	Goal; saved by goalkeeper or Shot misses goal (wide, over or post)
Penalty taker strategy	Overall strategy perceived by the observer (6)	Goalkeeper dependent; unclear or goalkeeper independent
Goalkeeper strategy	Overall strategy perceived by the observer (6)	Kicker independent; unclear or kicker dependent

Soto et al., 2019). Kappa levels of the OSPAF were 0.90 and 0.86—intra- and interreliability (Pinheiro et al., 2021b). The interpretation of this coefficient was adopted as follows: $\kappa > 0.8$: very good; $0.6 < \kappa < 0.8$: good; $0.4 < \kappa < 0.6$: moderate; $0.2 < \kappa < 0.4$: fair; and $\kappa < 0.2$: poor (Altman, 1991; O'Donoghue, 2010). The level of statistical significance adopted was $\alpha = 0.05$, with a 95% CI. All the data were analyzed by using JASP software (JASP Team, 2021; Computer software; JASP Version 0.14).

RESULTS

Descriptive data of all the OpenPose and OSPAF variables analyzed are presented as **Supplementary Material**.

OpenPose Confidence Score and Retest Reliability

The mean confidence score of OpenPose measures was 0.80 ± 0.14. The confidence score per variable is shown in **Table 2**.

Test-retest reliability values are shown in **Table 3**.

Influence Variables on Goalkeeper Strategy

The association between all the OpenPose and OSPAF variables with the goalkeeper's strategy was analyzed. **Table 4** presents only the variables that presented association and the respective values.

A logistic regression (*enter method*) was performed to investigate the relationship between the goalkeeper's tactical action and run-up speed on the likelihood of the goalkeeper strategy. The logistic regression model was statistically

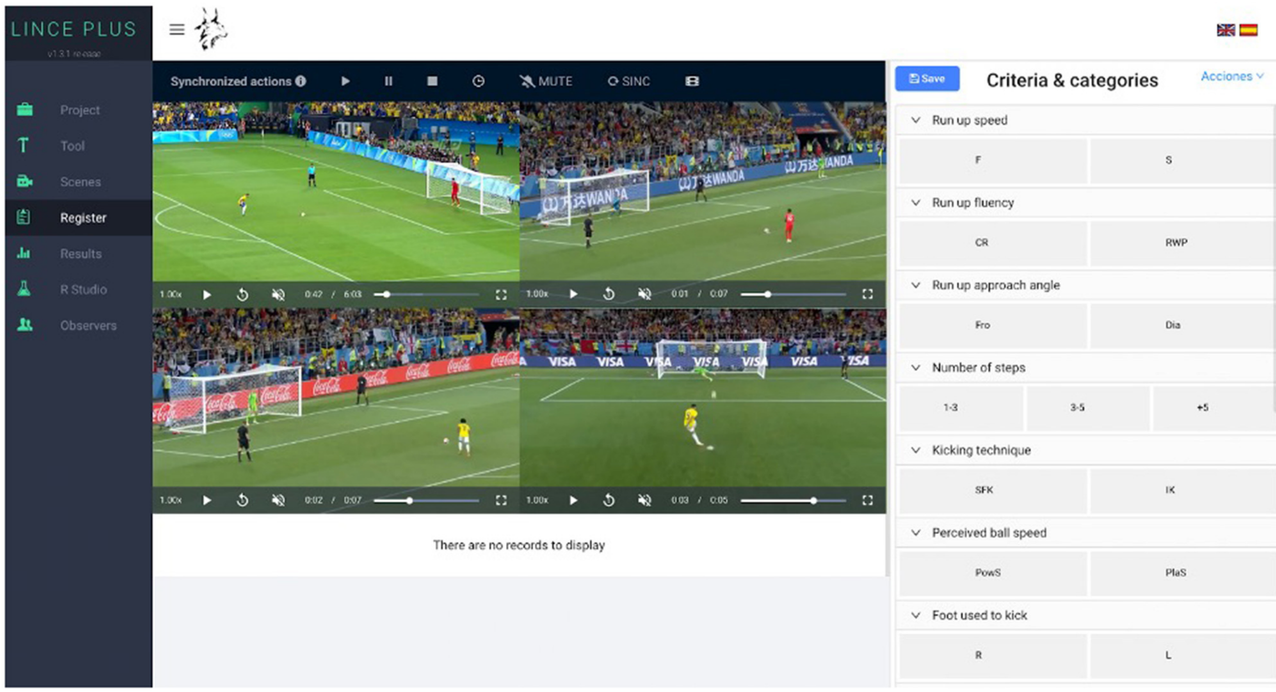


FIGURE 2 | Lince Plus interface. Example of criteria and categories coding.

significant, $\chi^2_{(36)} = 28.592$, $p < 0.001$. The model correctly classified 84.6% of cases. The goalkeeper's tactical action (awaiting) and run speed (slow) were related to a kicker-dependent strategy. While including the correlated OpenPose variable (goalkeeper anticipation) in the model [$\chi^2_{(35)} = 49.648$, $p < 0.001$], the accuracy is increased to 97.0%. Therefore, lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and run-up speed (slow) were associated with a kicker-dependent strategy.

Influence Variables on Penalty Taker Strategy

The association between all the OSPAF and OpenPose variables with the penalty taker's strategy was analyzed. **Table 5** presents only the variables that presented association and the respective values.

A logistic regression (*enter method*) was performed to investigate the relationship between the correlated OSPAF variables (run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed) on the likelihood of the goalkeeper-dependent strategy. The logistic regression model was statistically significant, $\chi^2_{(33)} = 24.819$, $p < 0.001$. The model correctly classified 97.1% of cases. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, the deception performed by the penalty taker, and lower ball speed were related to a goalkeeper-dependent strategy.

DISCUSSION

A unique method to calculate football players' orientation in in-match penalty kicks from a video has been tested. The mean confidence score of OpenPose variables was 0.80 and test-retest reliability showed an excellent reliability (Vogt, 2005; Lindstrom, 2010). The selected body orientation angle (goalkeeper anticipation) measured through OpenPose correlated significantly with the goalkeeper strategy. The prediction model of the goalkeeper's strategy had its accuracy increased when the variable goalkeeper anticipation was included. This finding corroborates the applicability of OpenPose to obtain the body orientation of professional football players during matches (Sangüesa et al., 2019).

Goalkeepers face a clear trade-off between moving early and moving in the correct direction (Hunter et al., 2018). The goalkeeper's chance of successfully saving a penalty kick is lower than that of the penalty taker to score and he must try to reverse this disadvantage by positioning himself to anticipate the direction of the kick that is about to come (Kuhn, 1988). In this study, the goalkeeper tactical action (awaiting) and run-up speed of the penalty taker (slow) were associated with a kicker-dependent strategy (84.6%). To further improve this model, the inclusion of the correlated OpenPose variable (i.e., goalkeeper anticipation) correctly classified 97.0% of cases. Corroborating previous studies (Nakai et al., 2019; Sangüesa et al., 2019, 2020a,b), the analysis of the body orientation through OpenPose has proved to be extremely useful on penalty

TABLE 2 | OpenPose confidence score per variable.

Player	Body orientation angle	Confidence score
Penalty taker	Non-kick foot orientation	0.51
	Shoulders	0.87
	Hips	0.85
Goalkeeper	Anticipation	0.87
	Left foot	0.84
	Right foot	0.83

TABLE 3 | Test-retest reliability per variable.

Player	Body orientation angle	*r*
Penalty taker	Non-kick foot orientation	0.924*
	Shoulders	0.998*
	Hips	0.991*
Goalkeeper	Anticipation	0.998*
	Left foot	0.953*
	Right foot	0.961*

*$p < 0.05$.

TABLE 4 | Association between OSPAF and OpenPose variables with the goalkeeper strategy.

	OSPAF variables	χ^2	*p*	Cramer's V
Goalkeeper strategy	Run up speed	4.875	<0.05	0.354
	GK tactical action	26.542	<0.05	0.825
	OpenPose variable	r_{pb}	*p*	
	Goalkeeper anticipation	0.959	<0.05	

TABLE 5 | Association between OSPAF and OpenPose variables with the penalty taker strategy.

	OSPAF variables	χ^2	*p*	Cramer's V
Penalty taker strategy	Run up speed	2.300	<0.05	0.243
	Run up fluency	5.512	<0.05	0.376
	Gaze behavior	22.224	<0.05	0.755
	Deception	8.770	<0.05	0.474
	OpenPose variable	r_{pb}	*p*	
	Ball speed	0.927	<0.05	

kick analytics. The improvement in the model related to the goalkeeper strategy shows the important practical application through the evaluation of the body orientation of football players by using OpenPose as a tool. These findings support previous study by Sangüesa et al. (2019, 2020a,b) and Nakai et al. (2019), which showed that skeletal data recognized by OpenPose are found to be highly applicable with sufficient accuracy. The acquisition of a set of biometric human body part orientations implies an improvement of the analysis of the penalty kick in elite football. Moreover, its integration with video allows this model to be used as a coaching resource to assess players' orientation and improve training strategies for game preparation.

Previous study has shown that the penalty outcome depends, above all, on the emerging results of the "penalty taker—goalkeeper" dyadic interaction (Lopes et al., 2012; Almeida et al., 2016; Pinheiro et al., 2021b). In this study, lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and run-up speed of the penalty taker (slow) were associated with a kicker-dependent strategy. From a behavioral perspective, the present findings corroborate this dyadic interaction between the players in a penalty kick, as results showed that the goalkeeper strategy is influenced by the run-up speed of the penalty taker. Corroborating with this finding, Noël et al. (2021) indicated that goalkeepers must consider the penalty taker's run-up for deciding when to initiate their jump to the ball. It is presumed that more successful goalkeepers wait longer to decide for a goal side because this allows them to access more reliable information from the penalty taker's kicking actions to anticipate the penalty takers' intentions (Noël et al., 2021). Analytical procedures that integrate the study of criteria related to the interactions between opponents are highly recommended in game analysis in football (Sarmento et al., 2014). In real competitions, penalty kicks are an interaction process and the observable performance is rather the emergent result of this interaction process than the display of skills and abilities of the two parties (Lames, 2006). The new approach presented in this study, combining different methods, provides a deeper understanding of the player strategy in penalty kicks, through objective identification of the anticipation of the goalkeeper (i.e., angle: αGK measured via OpenPose). To further clarify the process of interaction in the penalty kick and the goalkeeper response time, future studies could introduce a time interval before the kick or an event (exact moment of the kick) as new variables with objective parameters to be analyzed by using OpenPose.

Regarding the penalty taker, the selected body angles measured through OpenPose did not associate significantly with the shooter strategy. A possible explanation could be that the biomechanical patterns of approaching the ball during the kick may vary from player to player, regardless of the strategy adopted. Previous study has shown that kicking from an approach angle of 45 and 60° may alter aspects of kick technique, such as enhancing pelvic rotation and thigh abduction of the kicking leg at impact (Scurr and Hall, 2009). Reinforcing this, Prassas et al. (1990) reported significant differences for a substantial number of variables, related to the kicking foot, leg, the non-kicking foot, trunk, and hip segments in football kicks.

A novelty of this study is the adoption of OpenPose measurements with notational analysis (i.e., OSPAF) to analyze penalty kicks. The OSPAF is an adequate and consistent instrument for analyzing successful and non-successful penalty kick patterns (Pinheiro et al., 2021b). The analysis of observational variables in penalty shooting may provide a general description of its technical execution, which allows for detecting the shooters and the goalkeeper's strategy based on the behavioral variables studied (Pinheiro et al., 2021b). Although the variables used to detect body angles possibly relevant to the analysis of strategy of the shooter in penalty kicks in football did not correlate significantly with the penalty taker strategy, the variables measured by OSPAF (i.e., run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed) were able to correctly classify 97.1% of the penalty taker strategy. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, the deception performed by the penalty taker, and lower ball speed were related to a goalkeeper-dependent strategy. Partially corroborating these findings, Noël et al. (2015) identified three variables (attention to the goalkeeper, run-up fluency, and kicking technique) that in combination could predict kick strategy in 92% of the penalties. Previous study had also shown that run-up and spatiotemporal patterns of gaze may differ between strategies (Noël and van der Kamp, 2012; Noël et al., 2015). The difference in fluency is probably a consequence of penalty takers who use a keeper-dependent strategy to increase time at the end of the run-up by waiting for the goalkeeper to commit to one side of the goal (van der Kamp, 2006). Studies in a realistic setup pointed those penalty takers by using the keeper-dependent strategy direct their gaze more toward the goalkeeper compared to the ball and the target location (Kurz et al., 2018). In contrast, penalty takers by using the keeper-independent strategy direct their gaze more toward the ball compared to the goalkeeper and the target location (Noël and van der Kamp, 2012).

Several studies have investigated the penalty kick strategies in football (van der Kamp, 2006; Noël et al., 2015, 2021; Pinheiro et al., 2021b). However, to the best of our knowledge, this is the first study to use OpenPose to detect relevant body orientation angles in penalty kicks in elite football from TV broadcast. This study is a preliminary study in penalty kick analysis and, thus, requires further examination. This study limitation was to not use a larger sample (e.g., full season), as it could bring practical applications and be more representative. Another limitation of this study was using only one viewing angle. It was included only one standard viewing angle and video quality was standardized, as recommended by (Sangüesa et al., 2020b). Nevertheless, for comparison of penalties from different viewing angles, a 3D transformation must be adopted when using OpenPose. Camera positioning (e.g., viewing angles) could affect the accuracy and, thus, the feasibility of the systems in practical settings (Zago et al., 2020). Nonetheless, this study presents an innovative approach to the analysis of penalty kicks in football, combining notational analysis with OpenPose. Its integration with video specification allows this model to be used as a coaching tool to assess players' orientation under different penalty kicks, improving sports preparation against upcoming opponents.

Multiple practical applications can be provided, from improving and refining player strategy in penalty kicks, to producing a precise assessment of player orientation in high-level competitive scenarios. Although it is not optimal to analyze only 34 penalty kicks, results from the present preliminary data indicate that it is possible to distinguish the goalkeeper's strategy (i.e., kicker dependent vs. kicker independent) based on the degree of goalkeeper anticipation, extracted through OpenPose. The body orientation analysis gives practitioners the potential to quickly evaluate the temporal decision-making of the goalkeeper (i.e., anticipation movement of the goalkeeper) with consideration to choosing when to initiate their jump to the ball. This could help to identify which goalkeepers move early or late in the penalty kick situation. Based on the pattern of anticipation of the goalkeeper in official competitions, specific training strategies can then be developed. Besides, having a time-based set of player orientations enhances analysts' ability to evaluate the relationship of on-ball and off-ball direction with the anatomical patterns. Posture analysis by using OpenPose has been verified to be practical with our model on the goalkeeper strategy identification. Future study could train a deep learning model to provide results about pose orientation automatically and faster.

CONCLUSION

This study tested an innovative approach in applying OpenPose measures integrated with notational analysis to investigate the factors influencing the players' strategy in penalty kicks. Results showed the applicability of OpenPose for in-match penalty kick analysis and an improvement in the prediction of the goalkeeper strategy by using a body orientation variable (anticipation) extracted via OpenPose. The goalkeeper degree of anticipation, tactical action, and run-up speed of the penalty taker can be associated with the goalkeeper strategy. Observable variables such as run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed may identify the shooter strategy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article. The video recordings used were public, thus authorization was not required from the observed players.

AUTHOR CONTRIBUTIONS

GP and ML contributed to the conception, design of this study, and wrote the first draft of the manuscript. GP and XJ organized the database. GP performed the statistical analysis. GP, VC, and ML contributed to the revision of the manuscript and read and approved the presented version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Funding as a doctoral scholarship was provided by the Katholische Akademische Ausländerdienst (KAAD); the publication costs were supported by the Technical University of Munich (TUM) Publishing Fund.

ACKNOWLEDGMENTS

The authors would like to thank the Chair for Performance Analysis and Sports Informatics of the TUM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fspor.2022.818556/full#supplementary-material

REFERENCES

Almeida, C. H., Volossovitch, A., and Duarte, R. (2016). Penalty kick outcomes in UEFA club competitions (2010–2015): the roles of situational, individual and performance factors. *Int. J. Perform. Anal. Sport*. 16, 508–522. doi: 10.1080/24748668.2016.11868905

Altman, D. G. (1991). *Practical Statistics for Medical Research*. Boca Raton, FL: CRC Press.

Aranda, R., González-Ródenas, J., López-Bondia, I., Aranda-Malavés, R., Tudela-Desantes, A., and Anguera, M. T. (2019). "REOFUT" as an observation tool for tactical analysis on offensive performance in soccer: mixed method perspective. *Front. Psychol.* 10:1476. doi: 10.3389/fpsyg.2019.01476

Cao, Z., Simon, T., Wei, S., and Sheikh, Y. (2017). "Realtime multi-person 2D pose estimation using part affinity fields," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Honolulu, HI: IEEE), 1302–1310. doi: 10.1109/CVPR.2017.143

Casal, C. A., Maneiro, R., Ardá, T., Marí, F. J., and Losada, J. L. (2017). Possession zone as a performance indicator in football. The game of the best teams. *Front. Psychol*. 8:1176. doi: 10.3389/fpsyg.2017.01176

Castellano, J., Alvarez, D., and Bradley, P. S. (2014). Evaluation of research using computerised tracking systems (Amisco® and Prozone®) to analyse physical performance in elite soccer: a systematic review. *Sports Med*. 44, 701–712. doi: 10.1007/s40279-014-0144-3

CMU-Perceptual-Computing-Lab (2017). *OpenPose: Real-Time Multi-Person Keypoint Detection Library for Body, Face, Hands, and Foot Estimation*. Available online at: https://github.com/CMU-Perceptual-Computing-Lab/openpose (accessed November 1, 2021).

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2). *Auflage. Hillsdale, NJ: Erlbaum.*

Fariña, A. R., Fábrica, G., Tambusso, P. S., and Alonso, R. (2013). Taking the goalkeeper's side in association football penalty kicks. *Int. J. Perform. Anal. Sport*. 13, 96–109. doi: 10.1080/24748668.2013.11868634

Fernandes, T., Camerino, O., Garganta, J., Pereira, R., and Barreira, D. (2019). Design and validation of an observational instrument for defence in soccer based on the Dynamical Systems Theory. *Int. J. Sports Sci. Coach*. 14, 138–152. doi: 10.1177/1747954119827283

Gabin, B., Camerino, O., Anguera, M. T., and Castañer, M. (2012). Lince: multiplatform sport analysis software. *Proc. Soc. Behav. Sci*. 46, 4692–4694. doi: 10.1016/j.sbspro.2012.06.320

Hartley, R., and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, *2nd Edn*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511811685

Hughes, M., and Bartlett, R. (2004). "The use of performance indicators in performance analysis," in *Notational Analysis of Sport, 2nd Edn*, eds M. Hughes and I. Franks (London: Routledge), 166–188.

Hunter, A. H., Murphy, S. C., Angilletta, M. J. Jr., and Wilson, R. S. (2018). Anticipating the direction of soccer penalty shots depends on the speed and technique of the kick. *Sports* 6:73. doi: 10.3390/sports6030073

Jamil, M., Littman, P., and Beato, M. (2020). Investigating inter-league and inter-nation variations of key determinants for penalty success across European football. *Int. J. Perform. Anal. Sport*. 20, 892–907. doi: 10.1080/24748668.2020.1794720

JASP Team (2021). *JASP (Version 0.16)[Computer software]*.

Kuhn, W. (1988). "Penalty-kick strategies for shooters and goalkeepers," in *Science and Football*, eds T. Reilly, K. Lees, and K. Davids (London: E and FN Spon), 489–492.

Kurz, J., Hegele, M., and Munzert, J. (2018). Gaze behaviour in a natural environment with a task-relevant distractor: how the presence of a goalkeeper distracts the penalty taker. *Front. Psychol*. 9:19. doi: 10.3389/fpsyg.2018.00019

Lames, M. (2006). Modelling the interaction in game sports - relative phase and moving correlations. *J. Sports Sci. Med*. 5, 556–560.

Lames, M., and Hansen, G. (2001). Designing observational systems to support top-level teams in game sports. *Int. J. Perform. Anal. Sport* 1, 85–91. doi: 10.1080/24748668.2001.11868251

Li, Y., Alexander, M., Glazebrook, C., and Leiter, J. (2015). Prediction of kick direction from kinematics during the soccer penalty kick. *Int. J. Kin. Sports Sci.* 3, 1–7. doi: 10.7575/aiac.ijkss.v.3n.4p.1

Lindstrom, D. (2010). *Schaum's Easy Outline of Statistics, 2nd Edn*. New York, NY: McGraw-Hill Education.

Lopes, J. E., Araújo, D., Duarte, R., Davids, K., and Fernandes, O. (2012). Instructional constraints on movement and performance of players in the penalty kick. *Int. J. Perform. Anal. Sport* 12, 311–345. doi: 10.1080/24748668.2012.11868602

Makaruk, H., Porter, J. M., Sadowski, J., Bodasińska, A., Zieliński, J., Niznikowski, T., et al. (2019). The effects of combining focus of attention and autonomy support on shot accuracy in the penalty kick. *PLoS ONE* 14:e0213487. doi: 10.1371/journal.pone.0213487

Memmert, D., and Nöel, B. (2020). *The Penalty Kick: The Psychology of Success*. Aachen: Meyer and Meyer Sport.

Nakai, M., Tsunodam, Y., Hayashi, H., and Murakoshi, H. (2019). "Prediction of basketball free throw shooting by OpenPose," in *New Frontiers in Artificial Intelligence. JSAI-isAI 2018. Lecture Notes in Computer Science*, *vol 11717*, eds K. Kojima, M. Sakamoto, K. Mineshima, and K. Satoh (Cham: Springer), 31. doi: 10.1007/978-3-030-31605-1_31

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., et al. (2020). Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras. *Front. Sports Act. Living*. 2:50. doi: 10.3389/fspor.2020.00050

Noël, B., Furley, P., van der Kamp, J., Dicks, M., and Memmert, D. (2015). The development of a method for identifying penalty kick strategies in association football. *J. Sports Sci*. 33, 1–10. doi: 10.1080/02640414.2014.926383

Noël, B., and van der Kamp, J. (2012). Gaze behaviour during the soccer penalty kick: an investigation of the effects of strategy and anxiety. *Int. J. Sport Psychol*. 43, 326–345.

Noël, B., van der Kamp, J., and Klatt, S. (2021). The interplay of goalkeepers and penalty takers affects their chances of success. *Front. Psychol*. 12:645312. doi: 10.3389/fpsyg.2021.645312

O'Donoghue, P. (2010). *Research Methods for Sports Performance Analysis*. London: Routledge.

Paterson, G., van der Kamp, J., and Savelsbergh, G. (2020). Moving advertisements systematically affect gaze behavior and performance in the soccer penalty kick. *Front. Sports Act. Living*. 1:69. doi: 10.3389/fspor.2019.00069

Pinheiro, G. S., Costa, V. T., and Lames, M. (2021a). "Penalty kicks in elite football: identifying factors related to the player strategy," in *International Society of Sport Psychology (ISSP) 15th World Congress Proceeding*, Taipei, 228–229, Pinheiro, G. S., Nascimento, V. B., Dicks, M., Costa, V. T., and Lames, M. (2021b). Design and validation of an observational system for penalty kick analysis in football (OSPAF). *Front. Psychol*. 12:661179. doi: 10.3389/fpsyg.2021.661179

Prassas, S. G., Terauds, J., and Nathan, T. (1990). "Three-dimensional kinematic analysis of high and low trajectory kicks in soccer," in *Proceedings of the VIIIth Symposium of the International Society of Biomechanics in Sports*, eds N. Nosek, D. Sojka, W. Morrison, and P. Susanka (Prague: Conex), 145–149.

Preciado, M., Anguera, M. T., Olarte, M., and Lapresa, D. (2019). Observational studies in male elite football: a systematic mixed study review. *Front. Psychol*. 10:2077. doi: 10.3389/fpsyg.2019.02077

Ramakrishna, V., Munoz, D., Hebert, M., Bagnell, J. A., and Sheikh, Y. (2014). "Pose machines: articulated pose estimation *via* inference machines," in *European Conference on Computer Vision* (Berlin: Springer), 33–47. doi: 10.1007/978-3-319-10605-2_3

Sangüesa, A. A., Haro, G., Ballester, C., and Martín, A. (2019). *Head*, *Shoulders, Hip and Ball... Hip and Ball! Using Pose Data to Leverage Football Player Orientation*. *Barça sports analytics summit*. Barcelona: Barça Innovation Hub.

Sangüesa, A. A., Martín, A., Fernández, J., Ballester, C., and Haro, G. (2020a). "Using player's body-orientation to model pass feasibility in soccer," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (Nashville, TN), 3875–3884.

Sangüesa, A. A., Martín, A., Fernandez, J., Rodríguez, C., Haro, G., and Ballester, C. (2020b). Always look on the bright side of the field: merging pose and contextual data to estimate orientation of soccer players. *arXiv e-prints* arXiv:2003.00943.

Sarmento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N., and Leitão, J. C. (2014). Match analysis in football: a systematic review. *J. Sports Sci*. 32, 1831–1843. doi: 10.1080/02640414.2014.898852

Scurr, J., and Hall, B. (2009). The effects of approach angle on penalty kicking accuracy and kick kinematics with recreational soccer players. *J. Sports Sci. Med*. 8, 230–234.

Soto, A., Camerino, O., Iglesias, X., Anguera, M. T., and Castañer, M. (2019). LINCE PLUS: research software for behavior video analysis. *Apunts Educ. Fís. Deportes* 137, 149–153. doi: 10.5672/apunts.2014-0983.es.(2019/3).137.11

van der Kamp, J. (2006). A field simulation study of the effectiveness of penalty kick strategies in soccer: late alterations of kick direction increase errors and reduce accuracy. *J. Sports Sci*. 24, 467–477. doi: 10.1080/02640410500190841

Vilagut, G. (2014). "Test-retest reliability," in *Encyclopedia of Quality of Life and Well-Being Research*, eds A. C. Michalos (Dordrecht: Springer), 3001. doi: 10.1007/978-94-007-0753-5_3001

Vogt, W. P. (2005). *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*. Newcastle upon Tyne: Sage.

Wei, S., Ramakrishna, V., Kanade, T., and Sheikh, Y. (2016). "Convolutional pose machines," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (San Juan, PR: IEEE), 4724–4732. doi: 10.1109/CVPR.2016.511

Zago, M., Luzzago, M., Marangoni, T., De Cecco, M., Tarabini, M., and Galli, M. (2020). 3D tracking of human motion using visual skeletonization and stereoscopic vision. *Front. Bioeng. Biotechnol*. 8:181. doi: 10.3389/fbioe.2020.00181

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Original research article

International Journal of Sports Science & Coaching

International Journal of Sports Science & Coaching
1–8

DOI: 10.1177/17479541211069935
journals.sagepub.com/home/spo
SAGE

Small-sided games do not replicate all external and internal loads of a football match-play during pre-season: A case study

Guilherme de Sousa Pinheiro[1], Roberto Chiari Quintão[2], Vitor Bertoli Nascimento[3], João Gustavo Claudino[4], Adriano Lima Alves[5], Israrel Teoldo da Costa[6] and Varley Teoldo da Costa[7]

Abstract

This study investigated the differences in external and internal load during pre-season training sessions carried out with different SSGs and a friendly match in top-class professional football players. The study was conducted over a full pre-season. Participants were 9 male top-class professional football players (25 ± 5 years; 74 ± 8 kg; 177 ± 8 cm). The following variables were measured: training session duration (min), average heart rate (bpm), total distance (m), distance covered per minute (m/min), the total number of accelerations > 2.5 m/s^2, number of accelerations > 2.5 m/s^2 per minute, average distance of accelerations (m), the average value of acceleration (m/s^2). One-way ANOVA was performed to analyze the variance of all evaluated variables. No differences were found in the average accelerations (m/s^2) (0.128) among all the training formats. Moderate differences were found in number of accelerations > 2.5 m/s^2 per minute ($\eta2 = 0.396$, moderate effect) and average distance of accelerations ($\eta2 = 0.545$). Strong differences were found in HR ($\eta2 = 0.788$, large effect), total distance ($\eta2 = 0.797$, strong effect), distance per minute ($\eta2 = 0.775$ strong effect), total number of accelerations > 2.5 m/s^2 ($\eta2 = 0.699$ strong effect). Significant correlations were found just for the number of accelerations > 2.5 m/s^2 and the number of accelerations > 2.5 m/s^2 per minute with the 4v4, 8v8 and the FM ($r = 0.828$–0.890, $r^2 = 69\% – 79\%$; $p < 0.01$). External and internal loads differ across different SSGs and a FM during the pre-season training sessions.

Keywords

Acceleration, global positioning system, heart rate, soccer, training

Reviewers: Bruno Couto (Federal University of Minas Gerais, Brazil)
José Ponce (University of Extremadura, Spain)

[1]Chair of Performance Analysis and Sports Informatics, Technical University of Munich, Munich, Bayern, Germany
[2]Clube Atlético Mineiro, Belo Horizonte, Brazil
[3]Research and Development Department, Load Control, Contagem, Brazil
[4]Universidade de São Paulo, Sao Paulo, Brazil
[5]Universidade Federal de Minas Gerais, MG, Brazil
[6]Center of Research and Studies in Soccer, Universidade Federal de Vicosa, Vicosa, MG, Brazil
[7]UFMG Soccer Science Center, Universidade Federal de Minas Gerais, MG, Brazil

Corresponding author:
Guilherme de Sousa Pinheiro, Chair of Performance Analysis and Sports Informatics, Technical University of Munich, Georg-Brauchle-Ring 60/62 D, Munich, Bayern, Germany.
Email: guilherme.pinheiro@tum.de

Introduction

Training methods in football have evolved over the years.[1] Different training formats that claim to simultaneously improve physical capacities along with technical and tactical skills, under the demands of professional football settings have been used.[2] The different small-sided games are an efficient option for football players to simulate real match play situations and a proper tool to improve the physical fitness of players.[3–5]

Small-sided games (SSG´s) are one of the most used training tasks in football training methodology by coaches and are widely studied by sports scientists.[4,6–8] Modified games played on reduced field areas and often using adapted rules involving a smaller number of players compared with traditional games or high-intensity interval training are attractive exercise modalities to simultaneously develop endurance capacity, tactical and technical football skills.[9–11] Internal load (e.g. perceived exertion, heart rate, and blood lactate concentration) and skill requirements can be modified during SSG by altering certain factors, such as the number of players (e.g. 1vs1 to 3vs3, 4vs4 to 5vs5 or even 5vs5 to 5vs3),[4] the pitch size (e.g. small to large),[12] the rules of the game (e.g. offside)[13] and coach encouragement.[4,12]

Sided games are not a one size fits all training methodology when it comes to player loading.[14] Although the external and internal loads of different SSG conditions are one of the most studied performance parameters, some limitations remain in the international literature.[6,15] The heterogeneity that exists among designs causes a lack of consistency about the design of SSG,[3] the level of ability of the players, the coach encouragement, and the rules (e.g. number and type of goals, number of touches in the ball) used by researchers.[9,13] Such inconsistencies make it difficult to compare studies. Additionally, the main body of research in this area is conducted with young players[3,6] while studies conducted with senior players and especially with top-level professional players are less known.[15]

Despite the availability of evidence about training load within SSG-based programmes is large,[3,14,16] few studies compared different SSG formats and competitive friendly matches.[17–20] Giménez et al.[17] investigated, in professional football, the relationships among external loads (e.g. running, acceleration) and perception of exertion of friendly matches. The authors used three task training sessions (SSG/LSG: small or large-sided game, mini-goals - MG, ball circuit training- CT), in different design combinations (Design 1: SSG + MG + LSG; Design 2: SSG + CT + LSG, and Design 3: MG + CT + LSG). Those authors found that the training tasks did not replicate the main set of high-intensity efforts experienced in competitive conditions. Casamichana et al.[18] examined the impact of developing SSG´s training sessions compared to conducting friendly matches in semiprofessional players, and the results indicated that during friendly matches more sprints per hour of play were performed, with greater mean durations and distances, greater maximum durations and distances, and a greater frequency per hour of play for sprints, compared to the SSG´s. Dellal et al.[19] suggested that 4 versus 4 SSGs with specific conditions imposed (1 or 2 ball touch rules) induced a high proportion of high-intensity running, significant loading of the aerobic system (HR response). Castellano & Casamichana[20] analyzed the differences in the number of accelerations between small-sided games and friendly matches also in semiprofessional players. Those authors showed that the number of accelerations was higher during SSG used as part of training than it was during friendly matches. This finding might be related to greater neuromuscular fatigue and increased metabolic cost during matches, although in that study the players' heart rate was not monitored.[20]

Previous studies have compared the training load in various training tasks with friendly matches.[17,18,20–22] However, to the best of our knowledge, no studies compared the internal and external training load of different pre-season activities (e.g. small-sided games and a friendly match) in top elite football players. Coaches commonly plan friendly matches during the preseason to prepare the players for the dynamic effort usually demanded by official matches.[21] Monitoring post-match fatigue-related markers and planning effective training loads are among the key issues in sports preparation.[23] The preparatory period, commonly referred to as pre-season, is designed to develop players' physical capacities and prepare them for the various demands of match-play in the whole season.[24] Small-sided games are widely used as a training tool in football preparation, and the physical demands of different SSG conditions are the among most studied performance parameters.[6,25] Given its importance in planning effective training loads, especially during the pre-season, more research is needed to clarify whether the manipulation of different SSG formats causes similar responses to competitive scenarios in elite football. Understanding the possible differences between SSGs and friendly matches would be useful for coaches and practitioners dealing with training prescription in elite football, assisting with preseason programming, to better cope with training load management and prevent muscular injuries.[23] Therefore, the aim of this pilot study is to investigate the differences in external and internal load during pre-season training sessions carried out with different SSGs and a friendly match in top-class professional football players.

Materials and Methods

Sample

An initial sample of 23 professional football players (who were already champions of the main competitions in

Brazil, South America, Europe, as well as the FIFA World Cup) was monitored in a competitive season across the national championship (38 matches). In an attempt to provide a representative profile for this study, we elected to only include players who were regular starters with ≥ 60% participation in the total matches of the season and who not absent for more than 21 days due to injuries. The final sample, therefore, was composed of 9 male professional football players (25.11 ± 4.59 years, body mass 74.33 ± 8.3 kg, height 176.56 ± 7.94 cm). The group was composed of defenders, midfielders, and forwards. Goalkeepers (GK) were excluded from the analysis. These individuals were part of a first division team of Brazilian football, with professional experience in training and competitions of national and international level, recognized by the Brazilian Football Confederation (CBF) and South American Football Confederation (CONMEBOL). All players were submitted to medical evaluations by the club's medical staff and presented adequate health status for the practice of professional football.

All participants signed the Free and Informed Consent Term. The anonymity of the participants was preserved throughout the process. The data were only involved in this study after the agreement of the participants. It was obtained a consent letter from the club agreeing with the procedures. All research procedures were conducted according to the norms established by the National Health Council Resolution (466/2012) and the Declaration of Helsinki for research with humans. The project was approved by the Human Research Ethics Committee (485/10).

Experimental design: training games formats

A descriptive comparative design was used to investigate possible differences between several pre-season sessions, including SSG formats and friendly match. The study was conducted over a full pre-season and lasted three weeks. Various studies had used similar weekly training frequency.[8] The following game formats: 4v4, 6v6, 7v7, 8v8, 10v10, and 14v14 were compared with 1 friendly match (FM) in terms of the activity profiles of the players. The different SSGs used in this study were similar to previous studies.[3–5,7,15] The pre-season training session's formats carried out with different SSGs are described in table 1.

The pre-season sessions were performed in three consecutive weeks and were part of a regular training session. The 4v4 and 6v6 SSGs were performed on Tuesday and Thursday of the first week, respectively. The 7v7, 8v8 and 10v10 SSGs were performed on Tuesday, Thursday, and Saturday of the second week, respectively. The 10v10, 14v14 SSGs and FM were performed on Tuesday, Thursday, and Saturday of the third week, respectively. We chose the same weekdays to minimize the influence of the distribution of training loads over the weeks on player's physical responses; therefore, the recovery time was standardized. In the FM, the opponent was a professional football team, which competes in championships at a national level.

The players were familiarized with the use of these devices and with the SSG formats used. Coaches gave verbal encouragement to players during the training formats. Each team was composed of at least one defender, midfielder, and forward, to allow teams to explore the physical and technical-tactical specificities of each playing position during the different SSGs.[13] Considering that all players are part of a top-class professional team, we assumed that the homogeneity of the sample would not require an intentional team creation and opponent composition by the researchers. The decision, therefore, creating SSG´s teams and defining opponents was coach-driven.

All SSGs were implemented immediately after a warm-up (15–30 min) containing preparatory activities such as moderate running, dynamic stretching, balance and agility exercises, and accelerations. This process helped to ensure similar conditions across all SSGs. The ball was always available by prompt replacement when out to maximize effective playing time, except in the 4v4 format, which was played in a special arena surrounded by walls that kept the ball in play continuously. The number of touches to the ball was free for each player. Official FIFA-approved goals (7.32 × 2.44 m) were used. The offside rule was applied in all training formats, as well as in the match. On the rest periods, players were allowed to drink liquids *ad libitum*.

Data collection

The external load variables were obtained from portable GPS devices (GPSports SPI Pro X). According to the manufacturer, the GPS device has a sampling frequency of 15 Hz and includes a 100-Hz triaxial accelerometer. The manufacturer supplemented the GPS frequency to provide a sampling rate of 15 Hz.[25] Each player used a special vest which enabled the device to be fitted to the upper part of his back. The use of the special GPS vest and the heart rate (HR) monitor has not influenced the player's performance, as the club uses these devices in training sessions and official matches. The GPS devices were activated 15 min before the beginning of each training session, following the manufacturer's instructions. The data were transferred to a computer and analyzed in the software Team AMS (R1 2016, Australia). These devices have been previously used in elite football.[26] The registered variables were training session duration (min), average heart rate (bpm), total distance (m), distance covered per minute (m/min), the total number of accelerations >2.5 m/s^2, number of accelerations >2.5 m/s^2 per minute, average distance of accelerations (m), the average value of acceleration (m/s^2).

Table 1. Description of the SSG´s.

SSG	Pitch size	Area per player	SSG planning	Game description	Game purpose
4v4	32 × 20 m	80 m^2	4 sets × 7.5 min, 5 min of passive rest	Small SSG + GK	Free play
6v6	40 × 30 m	100 m^2	4 sets × 12.5 min, 5 min of passive rest	Small SSG + GK	Free play
7v7	52,5 × 68 m	255 m^2	4 sets × 10 min, 5 min of passive rest	Medium SSG + GK	Free play
8v8	35 × 40 m	87.5 m^2	4 sets × 7.5 min, 5 min of passive rest	Small SSG + GK	Ball possession
10v10	105 × 68 m	357 m^2	4 sets × 7.5 min, 5 min of passive rest	Large SSG + GK	Set piece training
10v10	105 × 68 m	357 m^2	4 sets × 7.5 min, 5 min of passive rest	Large SSG + GK	Free play
14v14	52,5 × 68 m	127.5 m^2	4 sets × 7.5 min, 5 min of passive rest	Small SSG + GK	Recreational training

Legend: SSG = small-sided game; GK = goalkeeper. The friendly match was played on a 105 × 68 m pitch (area per player: 357m^2).

The monitored training sessions occurred in the morning, between 9:00 am and 11:00 am, with sunny weather conditions and similar temperatures (24.3 ± 3.3°C), separated by an interval of 24 h between them. The majority of the training formats were performed on the same field with natural grass. Only the 4v4 sided game was played on artificial grass. The official match was played in home condition, between 4:00 and 6:00 pm on natural grass with sunny conditions and similar temperatures (26.3 ± 2.7°C). Pauses in the training game formats were excluded from the analysis. In order to ensure the ecological validity of the data collected, the planning and execution of the training did not suffer interference from the researchers.[26,27] During the study period, the athletes performed 2–3 strength training sessions per week. These sessions took place in the gym under the supervision of the club.

Statistical analysis

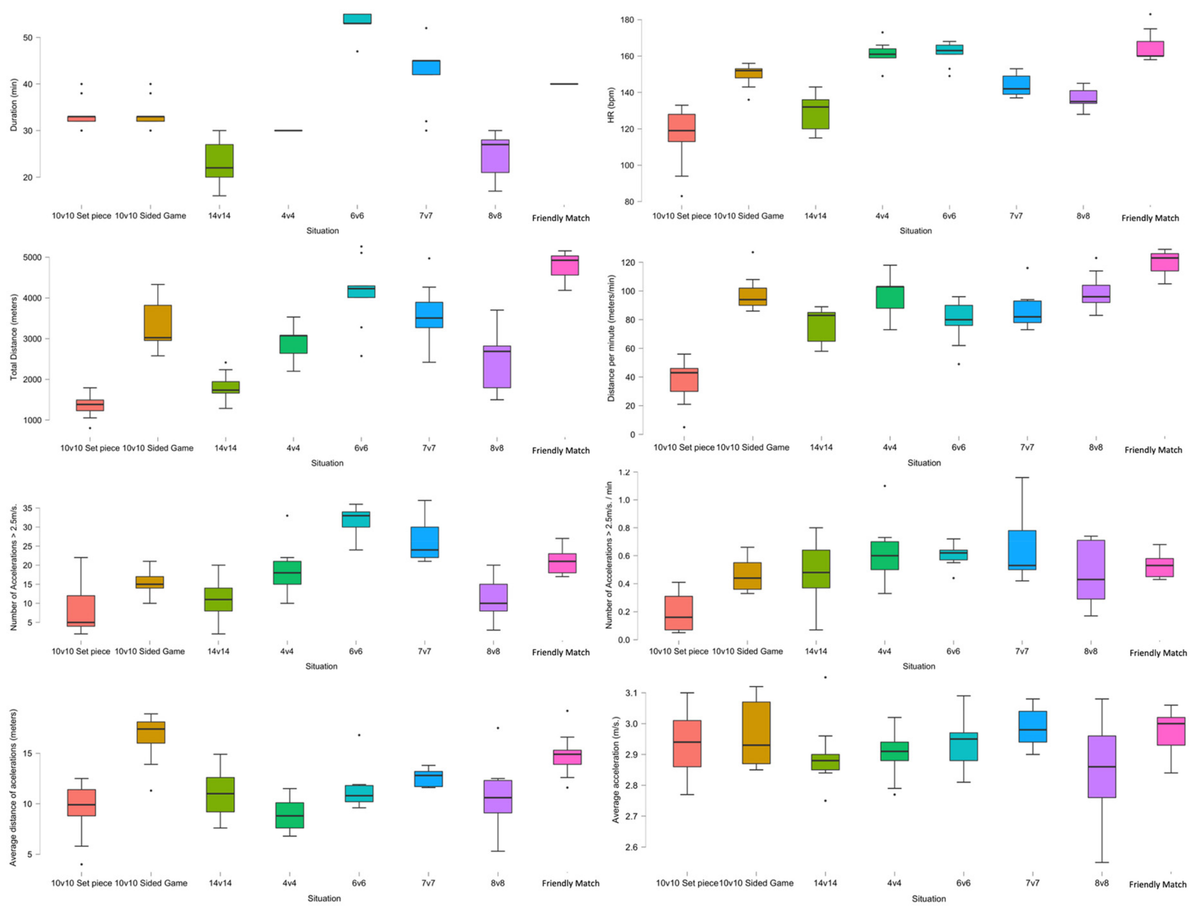
G*Power 3.1 software was used for the sample calculation. The hypothesized effect size assumed was 50%. A sample generalization power of up to 44% was achieved. A descriptive analysis of the data was performed, presenting the results in mean and standard deviation. Shapiro Wilk test was performed to verify data normality. As data presented normal distribution, parametric tests were performed. One-way ANOVA was performed to analyze the variance of all evaluated variables. The partial eta squared ($\eta 2$) has tested the effect size (ES) of ANOVA. The Ferguson's classification for the ES was used[27]: no effect (ES < 0.04); minimum effect (0.04 < ES < 0.25); moderate effect (0.25 < ES < 0.64); and strong effect (ES > 0.64). The pairwise comparisons were tested with Cohen's d to analyze the effect size. The following classification to measure the magnitude of ES was used[28]: no effect (d < 0.41), minimum effect (0.41 < d < 1.15), moderate effect (1.15 < d < 2.70) and strong effect (d > 2.70). The level of statistical significance adopted was $\alpha = 0.05$. All data were analyzed using JASP software (Team, 2020; JASP Version 0.14; Computer software).

Results

Descriptive statistics of duration, heart rate, total distance (m), distance per minute (m/min), the total number of accelerations >2.5 m/s^2, number of accelerations >2.5 m/s^2 per minute, average distance of accelerations (m) and average accelerations (m/s^2) are shown in Graph 1.

There was no difference in the training session duration (p = 0.995) across all training formats. Comparing all the pre-season training tasks, no differences were found just in the average accelerations (m/s^2) (p = 0.128). Moderate differences were found in number of accelerations >2.5 m/s^2 per minute ($\eta 2 = 0.396$, moderate effect; p < 0.001) and average distance of accelerations ($\eta 2 = 0.545$ moderate effect; p < 0.001). Strong differences were found in HR ($\eta 2 = 0.788$, large effect; p < 0.001), total distance ($\eta 2 = 0.797$, strong effect; p < 0.001), distance per minute ($\eta 2 = 0.775$ strong effect; p < 0.001), total number of accelerations >2.5 m/s^2 ($\eta 2 = 0.699$, strong effect; p < 0.001).

Tukey post hoc analysis revealed that the total number of accelerations >2.5 m/s^2 in the FM was higher than the 10v10 set piece (d = 2.409), 14v14 (d = 2.229) and 8v8 (d = 2.049). No difference was found comparing the match with the 10v10 sided game (p = 0.265), 4v4 (p = 0.997) and 7v7 (p = 0.440). Lower values were found comparing the FM with 6v6 (d = 2.979). The number of accelerations >2.5 m/s^2 per minute in the FM was higher than the 10v10 set piece (d = 2.926, strong effect). No differences were found comparing the FM with the 10v10 sided game (p = 0.988), the 14v14 (p = 0.975), the 4v4 (p = 0.955), the 6v6 (p = 0.993), the 7v7 (p = 0.828) and the 8v8 (p = 0.998). Average distance of accelerations in the FM was higher than the 10v10 set piece (d = 2.188), 14v14 (d = 1.670), 4v4 (d = 3.046) and 8v8 (d = 1.443). No difference was found comparing the FM with the 10v10 sided game (p = 0.842), the 6v6 (p = 0.063) and 7v7 (p = 0.472). HR in the FM was higher than the 10v10 set piece (d = 3.734), 10v10 sided game (d = 2.071), 14v14 (d = 3.773), 7v7 (d = 2.853) and the 8v8 (d = 4.070). No difference was found comparing the FM with the 4v4 (p = 0.984) and the 6v6 (p = 0.985). Total distance



Graph 1. Descriptive boxplots.

in the FM was higher than the 10v10 set piece (d = 10.967), 10v10 sided game (d = 3.255), 14v14 (d = 8.839), 4v4 (d = 5.385), 7v7 (d = 2.116) and the 8v8 (d = 3.889). No difference was found comparing the FM with the 6v6 (p = 0.233). Distance per minute in the FM was higher than the 10v10 set piece (d = 8.680), 10v10 sided game (d = 2.090),14v14 (d = 4.390), 4v4 (d = 2.162), 6v6 (d = 3.521), 7v7 (d = 3.104) and the 8v8 (d = 1.998). Further tables providing comparisons between the FM and the SSG´s are available as supplementary material.

Discussion

The purpose of the present pilot study was to investigate the differences in external and internal load during pre-season training sessions carried out with different SSGs and a friendly match in top-class professional football players. Our main findings showed no differences only in the average value of accelerations (m/s^2), across different SSGs and the FM. However, there were differences in the total number of accelerations > 2.5 m/s^2 and number of accelerations > 2.5 m/s^2 per minute, average distance of accelerations (m), HR (bpm), total distance (m) and distance per minute (m/min). The present findings are in accordance with an extensive literature supporting the hypothesis that different game formats demand particular external and internal loads, provoking a specific response in players and having acute effects on physical condition.[4,7] In addition, current results shows that the external and internal loads differ across different SSGs in relation to the FM during the pre-season training sessions, corroborating with previous research.[17] Therefore, further studies with a representative sample are encouraged to investigate if the SSGs can fully simulate the demands of a competitive match-play.

In elite football, the number and value of accelerations performed by players are important parameters due to their relevance in the competitive performance,[20] as well as the impact on the recovery time of this type of action. Although it has been shown that greater accelerations are required in game formats with larger pitch areas than in small ones,[3] our study showed an absence of differences in the mean value of accelerations. These findings indicate that the intensity of the actions was statistically similar in all training formats and the FM. Our findings are in accordance with Rago et al.,[2] showing that SSG seems to replicate well the acceleration demands observed during full-sized games.

One explanation in the present study could be because only a range of intensity was analyzed, as the analysis of different intensity ranges previously showed differences.[20]

Nevertheless, the absolute frequency of the accelerations >2.5 m/s^2 showed differences, indicating that the mechanical load may vary among the investigated training tasks. The FM elicited a greater number of accelerations >2.5 m/s^2 than the 10v10 (set-piece game purpose), the 14v14 and the 8v8, which can be partially explained by the game purpose, such as recreational or ball possession. When comparing the FM with the 10v10 sided game, 4v4 and 7v7, no difference was found. The 6v6 game elicited a greater number of accelerations >2.5 m/s^2 than the FM. Conversely with the literature, when both variables (dimensions and players) are lower, more demand is placed on acceleration and deceleration variables.[20] However, normalizing this analysis per minute, the relative frequency of the accelerations >2.5 m/s^2 was greater only than the 10v10 (set-piece game purpose). No differences were found comparing the FM with all other training formats. The very large correlation found among the number of accelerations > 2.5 m/s^2 in the 4v4, 8v8 and the FM corroborates previous findings from Rago et al..[2] Those authors have found a moderate correlation between acceleration and full-sized games. However, in that study, only 7v7 format was used, which could compromise the full analysis of the SSG´s impacts on acceleration.

Short-sprint ($\leq$ 20 m) performance is an important quality for success in football.[29] In the present study, the FM elicited a higher average distance of accelerations than the 10v10 (set-piece game purpose), 14v14, 8v8 (all with moderate effect size), and 4v4 (strong effect size) formats. Even though the 8v8, 10v10, and 14v14 formats were played in medium, large and full-size pitches, it was expected that the distance from the accelerations would be shorter, due to the specific tactical objective of this games (i.e. ball possession, recreational). However, the very strong effect size (d = 3.046) found in relation to 4v4 indicates that the average distance of the accelerations in the FM was substantially larger. This finding leads to the hypothesis that the exponential increase in the use of SSG´s in football training,[4,6] which does not entirely fulfil the complete match demand itself,[17] might be related to increased hamstring injuries in recent years.

No difference was found comparing the match with the 10v10 (free play), 6v6 and 7v7. The absence of difference with 10v10 is somehow obvious given the same area, number of players used and specific tactical purposes of the training format. Whether the objective of the training is to perform sprints similar to those made in a match,[3,4] playing situations with larger spatial dimensions or finishing situations involving few players and a large space, may be effective options in this process of optimizing the players' state of physical conditioning.

Heart rate is commonly used to monitor training intensity in elite football.[26] In the present study, the FM elicited greater HR than the 10v10 (set-piece game purpose), 10v10 (free play), 14v14, 7v7 and the 8v8. A systematic review showed that players obtained higher %HR when playing in a smaller format compared to other higher format of SSG´s,[3] however our present findings indicated that no difference comparing the FM with the 4v4 and the 6v6 preseason training tasks. In this way, 4v4 and 6v6 formats could promote similar levels of HR as in the FM. Our results could be explained by the dynamic movement pattern in such game formats, especially in the 4v4 format, which was played in a special pitch surrounded by walls that kept the ball in play continuously. These findings are in accordance with the literature, showing that SSG ´s may be effective in maintaining aerobic fitness.[4]

Our results showed that the FM elicited a higher total distance than the 10v10 (set-piece game purpose), 10v10 (free play), 14v14, 4v4, 7v7 and 8v8. This data corroborates past findings that indicate higher external load (i.e. distance covered) in SSGs played on medium and large pitches than on small pitches, for both amateur and professional level players.[3] Still, no difference was found comparing the FM with the 6v6. Tactical rules applied in SSG protocols could lead to a significant increase in total distance,[3,4] which might explain the statistically similar behaviour in the total distance in the 6v6 format. However, normalizing the analysis per time was possible to observe that the FM elicited greater distance per minute than all evaluated training formats. As has been indicated previously, the lack of similarity between the demands various groups of training formats could suggest the need to use the whole range of training game formats (e.g. from 1 vs. 1 to 10 vs. 10) when coaches want to overstimulate or replicate the demands of the game.

The limitation of this study was to have a small sample size. However, even with such a sample size, a sample generalization of the data of up to 44% was achievable. Therefore, future studies are encouraged to analyze a larger number of players. Another limitation was to not evaluate other constructs, such as perceived exertion, muscle soreness and tactical demands. These measures can complement the analysis of the SSG and the FM to understand better the differences. Despite this, the present study brings reflections and practical findings of external and internal load of various pre-season training tasks carried out with different SSG formats and FM, to further develop our understanding of the training stimulus provided by football specific training sessions during pre-season in professional football.

Before finalizing, we would like to do some reflections regarding the widespread use of the SSGs in elite football practices and its practical implications. Despite technological developments in sports training settings,[30,31] a longitudinal analysis between 2001 and 2013 found that hamstring injuries had annually increased by 4% in professional football.[24] Football matches' demands have also increased over the years. Yet, notwithstanding the popularity of SSG's and their tactical benefits,[11] it is unclear whether the common widely used SSG´s actually reproduce the demands of competitive matches, regarding physical performance[18] or if it may be effective to reduce the likelihood of muscular injuries.[23] Evidence has emerged in recent years supporting the need to include sprint stimuli in player's preparation with the objective of minimizing injury risk, in addition to the SSG practices. Furthermore, if players do not accumulate high chronic sprinting loads, they are more prone to "peaks" in acute sprinting loads during matches, which will ultimately lead them to increased likelihood of muscular injury, especially hamstrings injuries.[23] Although hamstrings injuries have been related to strength deficits, the current evidence is insufficient to recommend it.[30] Because there is inconsistency regarding the association between eccentric hamstrings strength and injury risk using different field devices.[30] Therefore, playeŕs physical preparation with similar demands to those of the match is necessary (i.e. similar acceleration distances), as an inadequate exposure to specific training loads could lead to muscular injury.

SSGs have been shown to be extremely relevant to enable football players to enhance their tactical and technical skills.[11] Nevertheless, it is important to mention that, although modulating SSG rules and pitch area could enable greater acceleration distances, the stimulus of SSG alone might not be enough to simulate mechanical demands of match-play regarding the distance of accelerations. Perhaps the most important element of preventing injuries in elite football lies in optimally managing player load.[23] SSGs typically mimics the physical intensity and movement patterns of match play, but can often result in limited exposure to sprinting distances due to the use of smaller areas. Match-play represents the largest stimulus in terms of high-speed running (HSR) and thus balancing training to reflect match play HSR load is important.[23] SSG´s should be used preferably for the development of greater technical-tactical skills and for aerobic-fitness development.[18] The inclusion of additional specific speed drills to SGG to prepare players for competition demands, with a suitable stimulus of acceleration distance, is necessary.

Conclusion

Pre-season training tasks carried out with different SSG formats elicit different demands on elite professional football players, regarding running, acceleration, and cardiovascular responses. Caution should be taken when selecting different SSGs, to improve the players' performance during the pre-season. This pilot study may help coaches to learn whether proposed tasks underload, replicate or overload the requirements of friendly match-play, something which might be considered when scheduling training sessions.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iDs

Guilherme de Sousa Pinheiro https://orcid.org/0000-0002-5572-1505
Adriano Lima Alves https://orcid.org/0000-0001-5315-8635
Varley Teoldo da Costa https://orcid.org/0000-0001-5786-633X

Supplemental material

Supplemental material for this article is available online.

References

1. Selmi O, Gonçalves B, Ouergui I, et al. Influence of well-being variables and recovery state in physical enjoyment of professional soccer players during small- sided games. *Res Sports Med* 2018; 26: 199–210.
2. Rago V, Silva JR, Mohr M, et al. The inter-individual relationship between training status and activity pattern during small-sided and full-sized games in professional male football players. *Sci Med Football* 2018; 2: 115–122.
3. Bujalance-Moreno P, Latorre-Román PÁ and García-Pinillos F. A systematic review on small-sided games in football players: acute and chronic adaptations. *J Sports Sci* 2019; 37: 921.
4. Hill-Haas SV, Dawson B, Impellizzeri FM, et al. Physiology of small-sided games training in football: a systematic review. *Sports Med* 2011; 41: 199–220.
5. Owen AL, Newton M, Shovlin A, et al. The Use of small-sided games as an aerobic fitness assessment supplement within elite level professional soccer. *J Hum Kinet* 2020; 71: 243–253
6. Folgado H, Bravo J, Pereira P, et al. Towards the use of multidimensional performance indicators in football small-sided games: the effects of pitch orientation. *J Sports Sci 2019*; 37: 1064–1071
7. Sgrò F, Bracco S, Pignato S, et al. Small-sided games and technical skills in soccer training: systematic review and implications for sport and physical education practitioners. *J Sports Sci* 2018; 6: 9–19.
8. Halouani J, Chtourou H, Gabbett T, et al. Small-sided games in team sports training. *J Strength Cond Res* 2014; 28: 3594–3618.
9. Praça GM, Andrade AGP, Bredt SGT, et al. Progression to the target vs. Regular rules in soccer small-sided games. *Sci Med Football* 2021; 5: 2–4. DOI: 10.1080/24733938.2020.1869811
10. Selmi O, Ouergui I, Levitt DE, et al. Small-Sided games are more enjoyable than high-intensity interval training of

similar exercise intensity in soccer. *Open Access J Sports Med 2020*; 11: 77–84

11. Teoldo I, Guilherme J and Garganta J. Training football for smart playing: on tactical performance of teams (1st edition). *Appris* 2017.
12. Rampinini E, Impellizzeri FM, Castagna C, et al. Factors influencing physiological responses to small-sided soccer games. *J Sports Sci* 2007; 25: 659–666.
13. Praça GM, Chagas MH, Bredt SGT, et al. The influence of the offside rule on players' positional dynamics in soccer small-sided games. *Sci Med Football* 2020; 4. DOI: 10.1080/24733938.2020.1819559
14. Lacome M, Simpson BM, Cholley Y, et al. Small-Sided games in elite soccer: does One size Fit All? *Int J Sports Physiol Perform* 2017: 1–24.
15. Clemente FM, Sarmento H, Rabbani A, et al. Variations of external load variables between medium- and large-sided soccer games in professional players. *Res Sports Med 2019*; 27: 50–59
16. Sarmento H, Clemente FM, Harper LD, et al. Small sided games in soccer – a systematic review. *Int J Perf Anal Sport* 2018; 18: https://doi.org/10.1080/24748668.2018.1517288
17. Giménez JV, Castellano J, Lipinska P, et al. Comparison of the physical demands of friendly matches and different types On-field integrated training sessions in professional soccer players. *Int J Env Res Pub HE 2020*; 17: 2904
18. Casamichana D, Castellano J and Castagna C. Comparing the physical demands of friendly matches and small-sided games in semiprofessional soccer players. *J Strength Cond Res 2012*; 26: 837–843
19. Dellal A, Owen A, Wong DP, et al. Technical and physical demands of small vs. Large sided games in relation to playing position in elite soccer. *Hum Mov Sci* 2012; 31: 957–969.
20. Castellano J and Casamichana D. Differences in the number of accelerations between small-sided games and friendly matches in soccer. *J Sport Sci Md 2013*; 12: 209–210.
21. Campos-Vázquez M, Castellano J, Toscano-Bendala FJ, et al. Comparison of the physical and physiological demands of friendly matches and different types of preseason training sessions in professional soccer players. *RICYDE* 2019; 58: 339–352.
22. Campos-Vázquez M, González-Jurado J, León-Prados J, et al. A comparison of internal load between friendly matches and a conditioned game in professional football players. *Cult Cien y Deporte* 2016; 11: 67–73. https://dx.doi.org/10.12800/ccd.v11i31.644.
23. Buckthorpe M, Wright S, Bruce-Low S, et al. Recommendations for hamstring injury prevention in elite football: translating research into practice. *Br J Sports Med* 2019; 53: 449–456.
24. Ekstrand J, Spreco A, Windt J, et al. Are elite soccer Teams' preseason training sessions associated With fewer In-season injuries? A 15–year analysis From the union of european football associations (UEFA) elite club injury study. *Am J Sports Med* 2020; 48: 723–729.
25. Johnston RJ, Watsford ML, Kelly SJ, et al. Validity and inter-unit reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. *J Strength Cond Res* 2014; 28: 1649–1655.
26. Pinheiro GS, Quintão RC, Custódio I, et al. Analysis of relationships between different training load monitoring tools in elite U-20 soccer. *Human Mov 2018* 2018; 2018: 52–59.
27. Quintão RC, Custódio IJO, Alves AL, et al. Quantification and comparison of external load of different soccer-specific training contents in relation to the game using a GPS with accelerometer. *Revista Brasileira de Futebol* 2013; 06: 3–12. In Portuguese.
28. Ferguson CJ. An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice* 2009; 40: 532–538.
29. Nicholson B, Dinsdale A, Jones B, et al. The training of short distance sprint performance in football code athletes: a systematic review and meta-analysis. *Sports Med* 2021; 51: 1179–1207
30. Claudino JG, Cardoso Filho CA, Bittencourt NFN., et al. Eccentric strength assessment of hamstring muscles with New technologies: a systematic review of current methods and clinical implications. *Sports Med - Open* 2021; 7: 10
31. Claudino JG, Capanema DO, de Souza TV, et al. Current approaches to the Use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review. *Sports Med Open* 2019 Jul 3; 5: 28.. PMID: 31270636; PMCID: PMC6609928.