Division of Sport and Exercise
University of the West of Scotland

AN INTEGRATED APPROACH TO UNDERSTANDING THE DEMANDS IMPACTING SOCCER OFFICIALS' DECISION-MAKING: IMPLICATIONS FOR TRAINING AND PERFORMANCE

Gary Paul McEwan

A Doctoral Thesis

Submitted in partial fulfilment of the requirements of the University of the West of Scotland for the degree of Doctor of Philosophy

August 2021

ACKNOWLEDGEMENTS

Completing this programme of research has proven to be one of the biggest challenges of my life thus far and has pushed me to my absolute limits – both professionally and personally. However, through the many moments of frustration, uncertainty, and self-doubt, I am fortunate to have received the support and guidance of many people – without which, it simply wouldn't have been possible to reach the finish line.

First and foremost, I would like to thank my supervisory team of Dr Rosie Arthur, Professor Vish Unnithan, and Professor Chris Easton. Rosie, you have provided an unwavering amount of support throughout the entire process. From your initial efforts in setting up this opportunity to your support in finding solutions during more difficult times, you have had my best interests in mind and for this I am extremely grateful. I also appreciate the flexibility and freedom you have offered me to work in my own way. Vish, your guidance, enthusiasm, and energy has been instrumental to shaping the Ph.D. and ensuring that it remained on track. Although you were not my Director of Studies, you have been incredibly attentive and have always been available to offer your support and advice. I am also immensely thankful for the additional opportunities you have afforded me over the past few years. And Chris, thank you also for the support and guidance you have provided over the past years.

I also owe a great deal of gratitude to Dr Neil Gibson. Neil, since embarking on the MRes in 2015, you have provided me with a vast number of opportunities and have continued to push me out of my comfort zone. Though progress has sometimes been slower than we both would have liked, your guidance and patience has been pivotal to my development as both a researcher and practitioner.

I would also like to thank the Scottish Football Association who part-funded my studentship. Particular thanks must go to Mr Thomas Murphy who was invaluable in helping to arrange for the studies of this thesis to take place. And of course, to the referees and other participants who gave up much of their time, thank you.

On a more personal note, the completion of this thesis has coincided with what has been the toughest period of my life. I am privileged, however, to have a close network of both friends and family who have always been there for me. Firstly, to two of my closest friends, Marc and Tom. United in your ability to distract me with your distinctly average banter, you have both been invaluable behind the scenes in keeping me sane and providing somewhat of a work-life balance. There will be a few beverages on me once this is all done and dusted!

I also have a small but incredibly loving family whose support and encouragement has been endless. Marcus and Sula, Iain and Caroline, it is impossible to put into so few words how thankful I am for everything you have done over the years for my mum and I so with that in mind, I will save it for a separate thank you. But for the time being, what I can promise is that I will always be there to look out for Thomas, Maisie, Marcus, and Henry in the same way that you have for me. And of course, a special mention to my late gran – Helen. Gran, although you never got to see me go off to university, I know you would have been extremely proud.

And finally, Melissa. I couldn't ask for a more understanding or supportive partner and I honestly don't know what I would do without you. Your enduring positivity and bubbliness provide the much-needed balance to my reserved and sometimes grumpy nature. You've been by my side for every step of the way, putting up with my very frequent frustrations but also celebrating every little milestone. I have no doubt that we'll continue to overcome whatever is thrown at us and I can't wait to see what the future holds.

DEDICATION

For you, Mum.

The photograph above has been reproduced with permission from the copyright holder.

PUBLICATIONS AND CONFERENCE PROCEEDINGS

Publications

McEwan, G.P., Unnithan, V., Easton, C., & Arthur, R. (2020). Development and validation of the Referee Training Activity Questionnaire (RTAQ): Towards a better understanding of the training practices of soccer officials. *Journal of Sport Sciences,* 38(24), 2782-2793. https://doi.org/10.1080/02640414.2020.1800371 **(Chapter 3)**

McEwan, G. P., Unnithan, V. B, Easton, C., & Arthur, R. (2022). Training practices and perceptions of soccer officials: Insights from the Referee Training Activity Questionnaire. *International Journal of Sport Science & Coaching,* online ahead of print. https://doi.org/10.1177/17479541221110707 **(Chapter 4)**

Conference Proceedings

McEwan, G.P., Unnithan, V., Easton, C., & Arthur, R. (2019). Development and validation of the Referee Training Activity Questionnaire (RTAQ): Towards a better understanding of the training practices of soccer officials. *24th annual Congress of the European College of Sport Science*, Prague, Czech Republic, 3rd – 6th July 2019. **(Chapter 3)**

McEwan, G.P., Unnithan, V., Easton, C., Glover, A. J., & Arthur, R. (2021). Validity and reliability of physiological and perceptual responses during a treadmill-based Soccer Referee Simulation (SRS). *26th annual Congress of the European College of Sport Science*, 8th – 10th September 2021. **(Chapter 5)**

ABSTRACT

Soccer officiating represents a highly challenging activity during which complex decision-making processes are performed under pressurised and physiologically taxing conditions. Though the multidimensional nature of soccer officiating is well accepted, previous literature has largely been monodisciplinary in nature, with the physiological, psychological, and perceptual-cognitive elements explored in isolation. During competitive match play, however, such demands are not experienced independently but are instead intricately linked. It therefore stands to reason that each should be considered in unison when evaluating the training and performances of soccer officials. With this in mind, the aims of this thesis were to: 1) develop valid and reliable measurement tools that could facilitate the multidisciplinary testing and evaluation of soccer officials; and 2) use these newly developed tools to collect information from which to inform and guide the future training prescription of soccer officials. Accordingly, Chapter 3 describes the development and validation of the Referee Training Activity Questionnaire (RTAQ) - a novel self-report questionnaire designed to quantify the engagement of soccer officials in the broad range of training practices pertinent to their successful performance. In demonstrating good levels of content, face, and criterion validity, the RTAQ was deemed an effective and practical method of monitoring the training practices of soccer officials. Building upon this work in Chapter 4, the RTAQ was utilised to document the types and volumes of training engaged in by soccer officials of varying roles and levels of professional attainment. Findings broadly highlighted the training of match officials to be focused on physical conditioning with little attention directed towards other facets of performance such as decision-making training. The exposure of officials to combined physical and decision-making training was found to be particularly limited, therefore underlining the need for a protocol that developed the decision-making and physical abilities of soccer officials in unison. Thus, the purpose of Chapter 5 was to assess the validity and reliability of a novel

treadmill-based Soccer Referee Simulation (SRS) that combined match-related decision-making with the physiological conditions synonymous of match play. No differences were detected between the SRS and actual match play for any of the selected heart rate (HR) measures, whilst the physiological and perceptual responses recorded were also aligned with those observed previously amongst elite field referees (FR). Physiological and perceptual responses elicited during the SRS were also found to be highly reproducible, with good levels of test-retest reliability observed for all outcome measures. The SRS was therefore found to be a valid and reliable protocol that closely replicates the physiological and perceptual loads imposed upon soccer FR during match play. Finally, the SRS was utilised in Chapter 6 to explore the impact that enhanced physiological loads have on the decision-making performances of soccer FR. Findings revealed for the first time that the accuracy of the decisions made by soccer FR may be compromised when decisions are made during (or shortly after) acute periods of extremely high physical exertion. Specifically, officials were found to be more likely to make a decisional error when: 1) HR ≥90% of maximal HR (HR_{max}); 2) respiratory rate (RR) ≥80% of peak RR (RR_{peak}); 3) ratings of perceived breathlessness (RPE-B) was rated as *very strong* to *maximal* (75–100 au); and 4) running speeds ≥250 $m \cdot min^{-1}$ were performed. Taken together, the current programme of research has made numerous original contributions to the literature concerning the science of soccer officiating and offers important implications of both a practical and theoretical nature.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

The following abbreviations are used frequently throughout this thesis and have been defined within the text in the first instance.

ANOVA ..... Analysis of Variance

AAR ..... Additional Assistant Referee

AR ..... Assistant Referee

CI ..... Confidence Intervals

CV ..... Coefficient of Variation

CVI ..... Content Validity Index

ES ..... Effect Size

FIFA ..... Fédération Internationale de Football Association

FR ..... Field Referee

GPS ..... Global Positioning System

HIIT ..... High Intensity Interval Training

HR ..... Heart Rate

HR_{max} ..... Maximal Heart Rate

ICC ..... Intraclass Correlation Coefficients

LOA ..... Limits of Agreement

OR ..... Odds Ratio

RER ..... Respiratory Exchange Ratio

RPE ..... Ratings of Perceived Exertion

RPE-B ..... Ratings of Perceived Exertion – Breathlessness

RPE-M ..... Ratings of Perceived Exertion – Muscular

RPE-T ..... Ratings of Perceived Exertion – Total

RR ..... Respiratory Rate

RR_{peak} ..... Peak Respiratory Rate

RTAQ ..... Referee Training Activity Questionnaire

SCORE ..... Scottish Centre of Refereeing Excellence

SD ..... Standard Deviation

SFA ..... Scottish Football Association

SR ..... Standardised Residuals

SRS ..... Soccer Referee Simulation

SWC ..... Smallest Worthwhile Change

TEM ..... Typical Error of Measurement

TRIMP ..... Training Impulse

UEFA ..... Union of European Football Associations

VAR ..... Video Assistant Referee

$\dot{V}E$ ..... Minute Ventilation

$\dot{V}E_{peak}$ ..... Peak Minute Ventilation

$\dot{V}O_2$ ..... Oxygen Uptake

$\dot{V}O_{2max}$ ..... Maximal Oxygen Uptake

WCS ..... Worst Case Scenario

CHAPTER 1

Introduction

1.1. BACKGROUND

Consequent to its widespread popularity, soccer (association football) has received considerable attention from the scientific community (Reilly & Gilbourne, 2003). Whilst much of this literature has focused on the players, growing attention has been directed towards those responsible for enforcing the laws of the game – the match officials (Castagna, Abt & D'Ottavio, 2007; Weston et al., 2012). Conventionally, competitive matches are governed by a team of officials comprising a field referee (FR), two assistants (AR), and a fourth official. At the highest levels of the game, this team may also include additional assistant referees (AAR), and more recently, video assistant referees (VAR). In their role as the final decision-making authority on the field of play, the FR is responsible for ensuring that match play is contested in a fair and safe manner (FIFA, 2016). A key aspect of this responsibility relates to the identification of incidents of foul play, with ~26 fouls awarded per match (Weston et al., 2011b; Mallo et al., 2012). Conversely, the main task of the AR is to identify incidents of offside, with ~30 occurring per match (Catteeuw et al., 2010c). As these decisions can impact a match's outcome, the correct interpretation and application of the laws of the game is essential (Červený et al., 2018). Accurate decision-making therefore represents the most pertinent metric of officiating performance (Helsen & Bultynck, 2004).

Given the dynamic and fast-paced environment in which they operate, the task of making consistently fair and accurate decisions is far from straightforward and necessitates a wide-ranging and multifaceted skill profile. First and foremost, soccer officials must possess comprehensive declarative and procedural knowledge of the laws of the game in order to interpret and apply them successfully (Helsen et al., 2019). Highly developed perceptual-cognitive skills are also integral to successful decision-making and facilitate the officials' ability to identify and extract the relevant information from which to base their judgements (Spitz et al., 2018b). The acquisition of such skills should therefore be at the forefront of the

officials' training and development. However, due consideration must also be given to preparing officials for the highly stressful and physically demanding conditions that characterise match play, as failure to cope with these demands may hinder their technical performance. Thus, a key challenge facing those responsible for training soccer officials is the requirement to address and develop the multiple elements of performance simultaneously.

Improvements in officiating performance can be effectively driven by the training process. That is, through engagement in structured and purposeful training, soccer officials may acquire the multifaceted skills and expertise required to officiate at the elite level. On the other hand, insufficient or inappropriate training may stall an official's development and impede their ability to carry out their role on the field of play. It is therefore crucial that practitioners possess a detailed understanding of current practices and monitor the training process closely, as in doing so, training may be refined and adjusted accordingly (Saw et al., 2015). Nonetheless, the training activities engaged in by soccer officials are not currently well documented. It is generally reported that their preparation is largely focused on the development of physical fitness, with perceptual-cognitive skills acquired mainly through observational learning or the officiating of competitive matches (MacMahon et al., 2007). What remains unclear, however, is whether such reports simply reflect the difficulties of capturing information on the wide range of training activities pertinent to this unique cohort. For instance, whereas heart rate (HR) and Global Positioning System (GPS) devices facilitate the objective monitoring of physical training sessions, no standardised method currently exists for monitoring training that is invariably self-led and performed off-field (i.e., decision-making, psychological, and technical skills training). A key objective of the present thesis was therefore to develop a valid and reliable measurement tool with the potential to assess these crucial but often overlooked forms of training, and to use this tool to reveal novel insights into the training activities currently engaged in by soccer officials. Previous attempts to document the training practices of soccer

officials were undertaken over a decade ago, and focused exclusively on elite FR (MacMahon et al., 2007; Weston et al., 2011). However, given the differing demands associated with their roles (Castillo et al., 2016b), we could expect the training practices of FR and AR to differ. It is also possible that differences in levels of experience and knowledge may result in different training practices being engaged in by officials at different levels of professional attainment. Such differences could potentially provide an insight into predictive relationships, whereby particular training profiles appear critical to the development of soccer officials and their ability to reach the elite levels of their profession. An up-to-date account of the training practices engaged in by soccer officials of varying roles and levels of experience is therefore warranted.

To aid their development and devise training interventions that prepare officials for the challenges of match play, a comprehensive understanding of each facet of soccer officiating is warranted. Considerable efforts have therefore been made to document the physiological (Castagna et al., 2007; Weston et al., 2012), psychological (Samuel, 2015), and perceptual-cognitive (Helsen et al., 2019) challenges encountered during match play. Though offering valuable insights into the competing demands placed upon soccer officials, previous literature has largely adopted a reductionist approach whereby each element has been explored in isolation. The demands of match officiating should not be considered as mutually exclusive, however, as they are instead intricately linked and imposed concurrently (Weston, 2015). A more integrated and multidisciplinary approach is therefore necessary if the challenges and complexities of soccer officiating are to be truly understood (Piggott et al., 2019). In particular, it is important that the physiological and psychological demands of match play are considered in unison with respect to their impact on the most important aspect of match officiating – accurate decision-making. By understanding how these demands influence the judgements made by soccer officials, training interventions that replicate the types of situations and contexts that prove most problematic to their decision-making may be developed.

Evidence exists to suggest that the decision-making of soccer officials may be impaired by excessive increases in psychological stress and anxiety incurred as a result of situational and social pressures (Johansen & Haugen, 2013). Amongst the stressors reported to influence the officials' judgements are the size (Picazo-Tadeo et al., 2017), noise (Unkelbach & Memmert, 2010), and density (Dawson & Dobson, 2010) of the crowd, as well as contextual factors relating to the elapsed time of the match (Unkelbach & Memmert, 2008) and the score differential (Voight, 2009). In contrast, a clear association is yet to be established between the physical and decision-making performances of soccer officials. The traditional approach to explore this association has been to ascertain the accuracy of the in-game official's decisions across distinct 10–20-minute epochs and contextualise these against the physiological loads recorded during that same period (Mascarenhas et al., 2009; Elsworthy et al., 2014b; Emmonds et al., 2015). There does however exist a number of limitations with this approach. Firstly, by aggregating data over prolonged epochs, acute periods of high physiological stress are likely to be concealed. Although lasting only seconds (Barbero-Álvarez et al., 2012), it is these high-intensity exertions that are perhaps most likely to impair an official's decision-making. Meanwhile, the aggregation of data over fixed 10–20-minute periods may also underrepresent the actual physical demands of match play (Fereday et al., 2020). If a meaningful association is to be identified between the physical and decision-making performances of soccer officials, a more sensitive analysis appears necessary whereby the accuracy of the officials' decisions is considered with respect to the physiological loads incurred at the time of (or in the moments preceding) the decision.

The inability of previous literature to detect a clear association between the physical and decision-making performances of soccer officials during match play may also reflect the challenges of exploring this association in-situ. For instance, as match play is susceptible to an array of environmental and contextual variables (Weston et al., 2006; Castillo et al., 2018),

delineating the influence of physical load on the in-match decisions of soccer officials is difficult. Alternatively, simulation protocols that mimic match play are often utilised to circumvent the challenges of conducting research in the actual performance setting (Russell et al., 2011). Whilst negating the contextual factors that confound match play data, simulation protocols enable physiological responses to be monitored in a more controlled environment. A number of simulations have therefore been developed for soccer players (Russell et al., 2011; Aldous et al., 2014; Page et al., 2015). In comparison, a valid and reliable simulation protocol does not currently exist for soccer officials. Though the activity profiles of the officials and players are interrelated, distinct differences exist with officials generally covering greater distances comprising more low intensity running (but less sprinting) than their playing counterparts (Weston et al., 2011c). Furthermore, as the physical and decision-making demands of officiating are superimposed upon one another, a compelling case exists for the inclusion of match-related decision-making into the physical training and testing protocols of sport officials (Kittle et al., 2019c). A valid and reliable protocol that replicates both the physical and decision-making elements of soccer officiating is therefore warranted.

1.2. AIMS AND OBJECTIVES OF THE THESIS

The overarching aims of this thesis were therefore to: 1) develop valid and reliable measurement tools (i.e., a training activity questionnaire and a match simulation protocol) that could facilitate the multidisciplinary testing and evaluation of soccer officials; and 2) use these newly developed tools to glean information from which to inform and guide the training processes of soccer officials. These aims were addressed over the course of four experimental chapters, which set out to achieve the following objectives:

1. To develop and validate a novel measurement tool for assessing the training practices of soccer officials (Chapter 3).
2. To document the current training practices of soccer officials in relation to their roles (FR and AR) and level of professional attainment. An additional objective was to explore the association between the perceptions and training practices of soccer officials (Chapter 4).
3. To develop a valid and reliable treadmill-based simulation that replicates the physiological and decision-making demands of soccer officiating (Chapter 5).
4. To assess the decision-making performances of soccer officials in relation to the physiological loads recorded immediately prior to the decision (Chapter 6).

CHAPTER 2

Literature Review

2.1. INTRODUCTION

To ensure that the laws of the game are upheld, match play is governed by a team of officials comprising a FR, two side-line AR, and a fourth official. At the highest levels of the game, this team may also include additional officials such as AAR, and more recently, VAR. As their decisions can influence a match's outcome, accurate decision-making represents the most pertinent metric of an official's performance (Helsen & Bultynck, 2004; Mallo et al., 2012). To facilitate optimal decision-making within dynamic and fast-paced performance environments, a wide-ranging and multifaceted skill profile is required. In particular, perceptual-cognitive skills are integral to successful decision-making, and facilitate an official's ability to identify and extract relevant information from the environment (Spitz et al., 2018b). Such information is then integrated with an official's technical expertise and knowledge of the laws of the game in order to select and execute an appropriate decision (Plessner & Haar, 2006). To compound matters further, these complex decision-making processes are performed under physiologically and psychologically taxing conditions (Castagna et al., 2007; Weston et al., 2012; Webb et al., 2020). The decision-making performances of soccer officials therefore appear dependent upon the multifaceted integration of several distinct yet interacting processes.

A common feature within much of the literature to date has been the adoption of a reductionist approach whereby the various elements of soccer officiating have been discussed independently from one another. That is, previous reviews have largely been monodisciplinary in nature, whereby the physiological (Castagna et al., 2007; Weston et al., 2012), psychological (Guillén & Feltz, 2011; Samuel, 2015), and perceptual-cognitive (Gaoua et al., 2017; Helsen et al., 2019) elements of match officiating have been discussed in isolation. Within the context of competitive match play, however, these demands are superimposed upon one another in a dynamic manner as opposed to being mutually exclusive (Weston, 2015). A more holistic and

integrated approach therefore appears necessary if the demands and complexity of soccer officiating is to be better understood (Piggott et al., 2019). Additionally, a review of the literature is yet to be completed whereby the multidimensional demands imposed during match play are discussed in relation to their impact on the most important aspect of soccer officiating – decision-making. By identifying the attributes that underpin successful decision-making, as well as the physiological and psychological conditions that prove most problematic, valuable information may be obtained that can inform the development of future training interventions. Thus, scope exists for a review to synthesise the available literature to provide a global overview of the factors that facilitate (or inhibit) the successful decision-making performances of soccer officials. The objectives of this review were therefore to: 1) explore the multifaceted demands imposed upon soccer officials and discuss the extent to which each impact upon decision-making performance; and 2) use this body of literature to enhance our understanding of the training requirements for soccer officials.

To obtain the necessary articles for this review, two electronic databases (PubMed and Web of Science) were searched from their inception until December 2019. Search terms were as follows: (*"referee"* OR *"umpire"* OR *"match official"*) AND (*"decision-making"* OR *"decisions"*). Secondary searches included manual examinations of the reference lists of the recovered articles as well as previously known studies. Only peer-reviewed articles published in English were considered with studies deemed eligible if they were concerned with the factors influencing the decision-making performances of team sport officials. Although the scope of this review is primarily concerned with soccer officials (both FR and AR), studies concerning interactor officials in other team-based sports (Australian football, rugby league, rugby union etc.) were also considered to account for the paucity of data with soccer officials in some areas of research. Though subtle differences exist in relation to the roles of match officials from different sports, it remains that the demands imposed upon team-sport officials are largely

similar. Specifically, team-sport officials have a shared responsibility to process perceptual cues within dynamic performance environments, maintain constant interaction with the players, and undertake complex decision-making processes within challenging psycho-physiological conditions (MacMahon & Plessner, 2008). A total of 877 articles were identified following these initial searches. Each article was screened by examining the title, abstract, and keywords, with 796 articles being excluded on the basis of ineligibility or title duplication. The full texts of the remaining records (n=81) were fully examined in respect to the scope of this review – this resulted in the exclusion of an additional 13 records, leaving a total of 68 articles for discussion (soccer, n=53; Australian football, n=6; basketball, n=5; rugby league, n=2; rugby union, n=1; ice hockey, n=1).

2.2. DECISION-MAKING PROCESSES OF SOCCER OFFICIALS – THE ROLE OF PERCEPTUAL-COGNITIVE AND TECHNICAL EXPERTISE

2.2.1. Role of Perceptual-Cognitive Expertise During the Decision-Making Processes of Team Sport Officials

Whilst a paucity of literature exists concerning the decision-making processes of soccer officials, highly developed perceptual-cognitive skills are considered integral to successful decision-making within dynamic, time-constrained domains (Williams et al., 2011). Generally, perceptual-cognitive expertise refers to the ability to identify relevant cues from the environment, process said information, and integrate it with existing knowledge and experience in order to formulate an appropriate response (Marteniuk, 1976). It therefore follows that such skills are essential to soccer officials whereby the ability to quickly identify and extract relevant information within a fast-paced performance environment is critical (Spitz et al., 2018b). Recently, perceptual-cognitive expertise was shown to be positively related to the accurate decision-making performances of soccer officials during off-field video-based assessments,

and successfully differentiates between elite and sub-elite FR (Spitz et al., 2018b) and AR (Gilis et al., 2008). Although the transfer of such findings to competitive settings remains to be elucidated, the ability to process game-related perceptual information represents a crucial component of expertise that appears to contribute to the more accurate decision-making of soccer officials.

The ability to quickly and accurately predict the outcome of a situation before it has arisen, termed anticipation, is frequently cited as a key perceptual-cognitive characteristic underpinning successful sporting performance (Mann et al., 2007). The significance of anticipation is particularly profound within temporally constrained environments where viewing time is limited (North et al., 2011). In relation to soccer match play, the visual cues relied upon are typically in motion and appear rapidly within continuously changing visual displays. Under such conditions, the time to identify and process relevant information is limited. The official's ability to anticipate the movement of players and sequences of play before they unfold may therefore create additional time for the selection and execution of an appropriate response, thus facilitating the decision-making process. Using the temporal occlusion paradigm during a video-based task, elite soccer FR were recently found to possess a greater anticipatory ability than their sub-elite counterparts, with differences paralleling their ability to correctly assess incidents of foul play (Spitz et al., 2018b). Although the relative contribution of individual perceptual-cognitive skills during the decision-making process remains unclear, anticipation appears vital (Larkin et al., 2018a). Regarding the factors supporting the superior anticipatory judgments of experts, pattern recognition and the capacity to utilise advance cues effectively appear important - both of which may develop with enhanced experience (North et al., 2016). The interpretation being that with greater experience, elite performers possess more comprehensive memory representations that allow them to recognise situations, and feed-forward information to predict likely future outcomes (Farrow et al., 2010).

Consistent with this literature, elite soccer FR (Spitz et al., 2018) and AR (Gilis et al., 2008) exhibit a greater capacity to recall foul play situations than their sub-elite counterparts. Essentially, the official's ability to recall patterns of play and utilise visual cues in the build-up to notable events, such as a player-to-player contact, may facilitate successful anticipation, thus enhancing the likelihood of a correct decision being reached.

The visual search strategies employed by match officials, characterised by the number and duration of fixations, also warrant consideration. Across a range of sporting disciplines, evidence suggests elite performers exhibit fewer fixations of longer duration than their non-elite counterparts (Mann et al., 2007). Interestingly, such findings are not replicated within the sports officiating literature as a comparable number and duration of fixations are performed during decision-making scenarios by elite and sub-elite soccer FR (Spitz et al., 2016) and AR (Catteeuw et al., 2009a; Catteeuw et al., 2010d; Schnyder et al., 2017). Similar observations have been reported amongst elite and sub-elite ice hockey (Hancock & Ste-Marie, 2013) and rugby union (Moore et al., 2019) officials. Alternatively, it appears likely that higher-level officials extract more relevant information from the same visual field as compared to lower-level officials (Catteeuw et al., 2009a; Hancock & Ste-Marie, 2013). A more pertinent feature of an official's visual search behaviour may therefore relate to the location of their visual fixations. When the visual scan patterns of elite and sub-elite soccer officials during a series of foul play situations were assessed, elite FR spent significantly more time fixating the most informative areas of their visual display (Spitz et al., 2016). Specifically, elite FR spent more time fixating on the contact zone of the attacking player and less time fixating the body parts detached from the infringement. Such findings are sustained by those of previous literature whereby goal-driven processes are proposed to dominate the visual searches of experts, with non-experts typically applying a more random or stimulus-driven control of fixation location

(Del Campo et al., 2018). Generally, experts demonstrate a tendency of knowing where and when to look, with non-experts drawn to more salient or superficial areas of the display.

The ability of officials to sustain high levels of visual attention is also critical, as even a temporary lapse of attention may compromise the correctness of a decision (Morris & O'Connor, 2017; Spitz et al., 2018b). Accordingly, highly developed visual and attentional capabilities represent important characteristics of successful match officials, and reliably discriminate between elite and sub-elite soccer FR (Ghasemi et al., 2011) and AR (Pietraszewski et al., 2014). Visual attention may be categorised into various sub-processes such as selective, sustained, and divided attention, with the extent to which each is exercised being contingent upon factors such as the type of decision (Memmert, 2009). For instance, FR assessing foul play situations must focus their attention on the immediate situation (selective and sustained attention), as diverting attention to external stimuli not directly involved may increase the possibility of missing crucial information. Conversely, AR assessing offside scenarios must simultaneously focus their attention on both the offside line and the player in possession of the ball, thus exercising their divided attention. Recent research also emphasises the importance of match officials being able to spread their attentional resources along its different meridians (Hüttermann et al., 2018). As elite soccer FR typically fixate their gaze on specific body parts (i.e., upper or lower body) when assessing incidents of foul play (Spitz et al., 2016), such instances may necessitate the concentration of attentional resources along the vertical and diagonal meridians. A wide horizontal breadth of attention, however, is likely required for the correct assessment of offside situations by AR. Such assertions are supported by recent observations where AR demonstrating a greater capacity to identify stimuli alongside the horizontal meridian of their attentional focus, demonstrated a greater ability to correctly assess offside incidents during a video-based task (Hüttermann et al., 2018).

Notwithstanding the significant efforts that have been made to identify the component perceptual-cognitive skills that contribute to the accurate decision-making of soccer officials, the actual cognitive processes involved during the decision-making moment have gone largely unexplored. That is, prevailing research approaches have invariably examined the association between isolated perceptual-cognitive skills and decision-making during simulated off-field tasks, with little attention paid to the subjective or lived experience of the officials themselves (Raab et al., 2018). There also exists a paucity of research that has been conducted within naturalistic environments, therefore raising questions over the transferability of previous findings to the competitive setting (Piggott et al., 2019). To address these shortcomings and gain insights into how perceptual and environmental information is processed to guide the soccer officials' decisions, researchers may wish to explore the concurrent use of qualitative methods of inquiry. In the context of soccer players, for example, semi-structured interviews have emerged as a key method of exploring the cognitions involved during the decision-making process (Levi & Jackson, 2018). The recall of cognitions and thought processes during post-performance interviews may however be vulnerable to distorted accounts caused by memory decay (Nicholls & Polman, 2008; Allen-Collinson, 2009). Traditional interview methods are also based upon the premise that the interviewee is conscious or aware of their thought processes; however, many of these cognitions are likely intuitive and occur at a sub-conscious level (Schweizer et al., 2011). Alternatively, "think aloud" protocols that involve the verbalisation of thought processes during the execution of a task may provide unique insights into the cognitions present at the time of the decision (Eccles & Arsal, 2017). In contrast to the growing body of work that has employed think aloud protocols to capture the in-event cognitions of athletes (Elliott et al., 2020), this method has not yet been utilised within the context of soccer officials. A mixed method approach to future research is therefore

recommended whereby both quantitative and qualitative methods are used to explore the decision-making processes of soccer officials in greater detail.

2.2.2. Integration of Perceptual-Cognitive and Technical Expertise During the Decision-Making Processes of Team Sport Officials

Soccer officiating represents a highly domain-specific activity whereby expertise is reserved to the specific role that one occupies (MacMahon et al., 2007; Catteeuw et al., 2009b). Consequently, highly developed perceptual-cognitive skills alone are likely insufficient to facilitate the accurate decision-making of soccer officials. For example, although possessing comparable perceptual-cognitive capacities, elite soccer officials unsurprisingly demonstrate a superior ability to correctly assess infringements compared to players (MacMahon et al., 2007). Likewise, the performance of soccer FR and AR during foul play and offside decision-making tasks reflects high levels of role specificity (Catteeuw et al., 2009). That is, FR make more accurate foul play decisions than AR, with AR demonstrating a superior ability to correctly adjudicate scenarios of offside. It therefore appears prudent to consider the role that technical expertise plays within the decision-making performances of soccer officials, as we may speculate that it is the integration of an official's technical and perceptual-cognitive expertise that underpins their decision-making ability.

As aforementioned, superior anticipatory skill is likely to contribute to successful decision-making, and partially reflects an official's domain-specific knowledge and capacity to recognise patterns (Spitz et al., 2018b). One source of information that appears particularly influential towards successful anticipation is tactical knowledge and awareness (Slack et al., 2013). A greater tactical awareness may also benefit other aspects of a match official's technical and decision-making performance such as their on-field positioning (Morris & O'Connor, 2017; Slack et al., 2013). Indeed, by understanding the tactical behaviours of the

players, officials may be better equipped to predict sequences of play, which in turn may enable the adoption of more suitable viewing positions. In relation to on-field positioning, being near infringements and adjudicating from optimal viewing angles are commonly perceived as vital if an official is to obtain an unobstructed view and enhance their perception of each infringement. In partial support of these suggestions are the findings of Hüttermann and colleagues (2017), who found narrower visual angles to be advantageous for the successful offside decision-making of AR. It has also been reported that the error rate of FR is lowest when adjudicating on infringements in the central area of the field from distances of 11-15 m (Mallo et al., 2012). A considerable body of evidence, however, suggests that on-field positioning has little effect upon the decision-making accuracy of soccer FR (Mallo et al., 2012; De Oliveira et al., 2011; Hossner et al., 2019; Riiser et al., 2019) and AR (Oudejans et al., 2005; Helsen et al., 2006; Catteeuw et al., 2010c; Barte & Oudejans, 2012). Nonetheless, it appears rash to conclude that on-field positioning has no effect upon decision-making accuracy, as its impact should be explored in relation to other factors such as teamwork. During high-standard competitive matches, soccer officials operate within a team consisting of a FR, two AR, and a fourth official. This team of match officials may also receive additional support from AAR and VAR. Where the FR's view of a situation is obstructed, information obtained from colleagues may allow them to uphold the correctness of their decisions (Mallo et al., 2012). These suggestions are substantiated by recent research whereby the on-field decision-making performances of a cohort of FR was shown to be compromised most within the lateral areas of the pitch farthest away from the AR (right defensive and left attacking zones) (Joo & Jee, 2019). Exploring the impact of on-field positioning upon decision-making accuracy, in relation to contextual factors such as teamwork therefore appears important and warrants further investigation.

In the final stages of the decision-making process, the visual information obtained following previous perceptual-cognitive processing is then integrated with the officials' existing knowledge and technical expertise (Plessner & Haar, 2006). In particular, the nature (i.e., foul/no foul or onside/offside) and severity (i.e., a careless/reckless challenge) of the perceived infringement must be categorised according to the laws of the game, and then acted upon in the form of an appropriate decision (i.e., free kick and yellow card). As the primary responsibility of soccer officials is to ensure match play is contested in accordance with the laws of the game, their understanding of these laws represents the most fundamental aspect underpinning successful decision-making. It is also noteworthy that soccer officials often endeavour to avoid unnecessary disruptions to match flow and momentum by applying a practical (rather than literal) interpretation of the laws of the game (Unkelbach & Memmert, 2008). Indeed, empirical evidence suggests that when infringements are viewed in the context of the game, fewer sanctions are awarded than when the same infringements are viewed as random successions of individual scenes (Unkelbach & Memmert, 2008; Brand et al., 2006). Soccer officials' decisions therefore appear contingent upon contextual judgement whereby decisions are not viewed as *"black-or-white"* but are instead considered in relation to various contextual factors (Larkin et al., 2018a). For instance, foul calls have been reported to occur most against the team with the fewest fouls (Anderson & Pierce, 2009). Soccer officials have also been reported to tacitly bias subsequent *50-50* decisions in favour of the opposite team in order to *"put things right"* following an erroneous decision (Lane et al., 2006; Schwarz, 2011). Such tactics may be better understood from a game management perspective and as a strategy used to return play to a state of equilibrium (Lane et al., 2006). Whilst highlighting the complex dynamics that confound the in-match decisions made by soccer officials, it is also important to acknowledge the implications that such findings have in relation to how the accuracy of these decisions are assessed within the literature. Here, expert panels comprising former top-level

officials typically use video footage to determine the accuracy of each decision in isolation. In contrast to the in-game officials, however, these assessments are made in the absence of contextual information such as the elapsed time of the match, the score line, and the outcome of previous decisions (Mallo et al., 2012). Moreover, expert panels enjoy the utility of video technology whereby incidents can be replayed in slow motion. Although slow motion replays may facilitate the correct assessment of objective decisions that require both spatial and temporal precision (i.e., offside decisions), their use for traditionally subjective decisions that involve rule interpretation and application (i.e., foul play decisions) may result in the more severe penalisation of infringements as compared to real time (Spitz et al., 2018a). Thus, the use of agreement scores with expert panels to assess the decision-making performances of soccer officials has been criticised, as the different conditions under which incidents are assessed are likely to result in different decisions being reached (Weston, 2015). In considering the findings of previous literature concerning the decision-making performances of soccer officials, it is therefore important that they be interpreted in relation to the presence (or absence) of contextual information.

2.3. THE IMPACT OF PSYCHOLOGICAL STRESS UPON THE DECISION-MAKING PERFORMANCES OF TEAM SPORT OFFICIALS

2.3.1. Situation Criticality and the Decision-Making Performances of Team Sport Officials

In considering the economic importance of winning within the modern game, and the significant implications that decisional errors may have on a match's outcome, soccer officials face considerable pressure to make consistently fair and accurate decisions (Boyko et al., 2007; Riedl et al., 2015; Lago-Peñas C & Gómez-López, 2016; Rohde & Breuer, 2016). The psychological demands imposed upon match officials are therefore considerable (Mascarenhas

et al., 2006; Samuel, 2015; Aragão-Pina et al., 2018). Amongst the sources of psychological stress frequently cited by soccer officials is the level of competition and importance of the match, with higher competitive levels typically being congruent with greater psychological loads (Johansen & Haugen, 2013). The manifestation of potentially harmful emotional states such as stress and anxiety are particularly prevalent amongst top-level officials and are likely reflective of the increased expectations and scrutiny that accompanies elite-level soccer (Johansen & Haugen, 2013; Balmer et al., 2007; Voight, 2009). Although contingent upon the appraisal mechanisms employed by the official, high levels of anxiety likely exert a deleterious effect upon various components of the decision-making process such as anticipation, visual search strategies, and attentional control (Wilson et al., 2009; Vater et al., 2016; Alder et al., 2018). A number of theories exist to perhaps explain such findings (Ford et al., 2017). One line of thought suggests that high levels of anxiety may shift attention towards task-irrelevant cues such as worry and apprehension, therefore placing a strain on an individual's limited attentional resources (Eysenck et al., 2007). Given the pivotal role that high levels of attention play within the decision-making process, it is perhaps unsurprising that excessive levels of anxiety are purported to impair the decision-making performances of soccer FR (Balmer et al., 2007; Hoseini et al., 2011). Currently though, no data exists in relation to the association between anxiety and the decision-making performances of soccer AR. Considering the proximity of the AR to the crowd, and the important role that high levels of concentration and visual attention play during the adjudication of offside (Hüttermann et al., 2018), the anxiety-performance relationship requires further investigation amongst this cohort.

Contextual variables such as the stage of the match and the score differential also represent key sources of psychological stress amongst team sport officials and may exacerbate levels of anxiety (Unkelbach & Memmert, 2008; Voight, 2009; Hill et al., 2016; Ritchie et al., 2017). For instance, in the final stages of an important match with a low score differential, a single

mistake by the officials may have significant consequences on the match outcome (Lago-Peñas & Gómez-López, 2016). Whilst the stage of the match does not appear to have a clear effect upon decision-making performance (Castillo et al., 2019), small score differentials at the time of the decision were recently associated with a greater incidence of incorrect decisions than were large margins (Corrigan et al., 2019). Whether these findings reflect heightened levels of anxiety however remains speculative, as no measures of anxiety were obtained in this investigation. To examine the anxiety-performance relationship amongst soccer officials, the accuracy of the officials' decisions has previously been assessed in relation to pre- or post-performance measures of state and trait anxiety (Balmer et al., 2007; Johansen & Haugen, 2013). Although such measures may provide a useful indication of an official's psychological state at a particular time point, anxiety levels are likely to fluctuate throughout the course of a match in response to acute stressors such as crowd noise and previous errors (Plessner & Betsch, 2001). A key criticism of previous approaches therefore concerns their inability to assess an official's anxiety at the time of the decision. Moreover, a reliance on correlational designs means that the factors which influence how stress impacts on officiating are yet to be revealed. Future investigations using anxiety manipulations in experimental conditions could expand understanding regarding stress and the factors which may moderate the impact of stress on the decision-making of officials. However, it is important to note that replicating the stress of officiating in laboratory-based conditions is not without its challenges.

2.3.2. Social Pressures and the Decision-Making Performances of Team Sport Officials

During match play, soccer officials are in constant interaction with the players and coaches, with their decisions routinely coming under intense criticism (Webb et al., 2020). In particular, players often contest the officials' decisions with the intention of swaying the decision in their favour (Hacicaferoğlu & Gündoğdu, 2014; Gencay & Aydin, 2017). Interestingly, player

vocalisations do not appear to influence the foul play decisions made by soccer FR (Lex et al., 2015). Once a decision has been made, however, player vocalisations do appear to increase the likelihood of a sanction being issued for the offending individual (Lex et al., 2015). Additionally, as matches are often contested in the presence of large partisan crowds, the influence of spectators on the decision-making performances of soccer officials has been extensively researched. Factors reported to influence the decisions made by soccer FR and AR include: crowd noise (Nevill et al., 2002; Unkelbach & Memmert, 2010); the number of spectators (Downward & Jones, 2007; Page & Page, 2010; Picazo-Tadeo et al., 2017); crowd density (Dawson & Dobson, 2010; Goumas, 2013); the composition of the crowd (Nevill et al., 2017); and the distance between the field and the stand (Dawson & Dobson, 2010; Scoopa, 2008; Buraimo et al., 2011). Together, these studies provide evidence substantiating the home-advantage phenomenon whereby sporting teams are observed to win a disproportionate number of matches at home, with the tendency of the match officials to bias their decisions in favour of the home side playing a pivotal role (Goumas, 2014). Notwithstanding the array of factors that likely contribute to home-advantage (Johnstone, 2008; Constantinou et al., 2014), work by Balmer and colleagues (2007) suggests that the inclination of soccer officials to bias their decisions in favour of the home team relates to heightened levels of cognitive anxiety, induced in response to amplified crowd noise. Although decision-making accuracy was not assessed *per se*, such data highlight the profound effect that heightened levels of anxiety have upon an official's decisions. Similarly, the decision-making performances of Iranian FR were negatively affected by increased levels of anxiety (Hoseini et al., 2011), whilst Sors and colleagues (2019) demonstrated the decision-making performances of highly anxious basketball officials to deteriorate under pressing crowd conditions. In partial explanation of these observations, it may be proposed that when faced with acute stressors such as crowd noise, more anxious individuals are likely to adopt avoidance coping strategies (Nevill et al.,

2017). For example, when presented with a contentious decision in front of a vociferous home support, officials may favour the home side in order to appease the crowd. This suggestion gains support from previous research whereby rugby union officials were documented to adopt avoidance-coping mechanisms in order to disengage from situations of interpersonal conflict (Hill et al., 2016). As match officials must maintain constant situational control, such strategies were reported to be ineffective and associated with under-performance and choking under pressure (Hill et al., 2016).

As discussed, stressors relating to the criticality of the decision and the social pressures experienced from players and spectators appear to have a particularly deleterious effect on the decision-making performances of soccer officials, with these decrements likely resulting from heightened levels of anxiety (Balmer et al., 2007; Johansen & Haugen, 2013). It is worth noting, however, that such stressors are ultimately subjective and may be interpreted differently between individuals. That is, whereas some individuals may view a given stressor as threatening, others may appraise the same stressor more positively. It also appears that some individuals are able to tolerate higher levels of anxiety before decrements in performance are witnessed, whilst others might only function optimally within a narrow range of anxiety (Raglin & Hanin, 2000). The magnitude and direction of the anxiety-performance relationship is therefore not unequivocal and is moderated by complex individual differences relating to factors such as the experience and psychological profile of the match official (Mellalieu et al., 2009). For example, athletes high in levels of self-efficacy are generally predisposed to interpreting anxiety as facilitative rather than debilitative, and often exhibit lower levels of both cognitive and somatic anxiety (Nicholls et al., 2010). Other factors that appear to moderate the anxiety-performance relationship include personality traits (Byrne et al., 2015) and coping strategies (Dias et al., 2012). In considering the aforementioned studies examining the anxiety-performance relationship amongst soccer officials, little attention has been paid to these

moderating variables. Future research may therefore wish to explore how different soccer officials respond to heightened levels of anxiety and the impact this has on their decision-making performance.

2.4. THE IMPACT OF PHYSICAL AND PHYSIOLOGICAL LOAD ON THE DECISION-MAKING PERFORMANCES OF TEAM SPORT OFFICIALS

The physical and physiological demands of soccer officiating are vast. Briefly, elite FR cover distances of 9-11 km during competitive matches (Krustrup et al., 2009; Weston et al., 2011b; Weston et al., 2011c; Castillo et al., 2017); considerable proportions of which incorporate high-intensity actions such as high-speed running and sprinting (Weston et al., 2011c; Barbero-Álvarez et al., 2012). Examinations of the cardiovascular responses elicited during match play suggest FR attain a mean HR of 85-95% of their maximal (Barbero-Álvarez et al., 2012; Weston et al., 2006; Castillo et al., 2016b). Conversely, AR cover lower distances of 5-7 km, and exhibit mean match HRs of 79% (Castillo et al., 2016b). As the physical and decision-making demands of match officiating are performed concurrently, the relationship between these facets of performance warrants consideration (Weston et al., 2012). Intuitively, the necessity to perform complex perceptual-cognitive processes under high levels of physical exertion would appear counterproductive to optimal decision-making (Hüttermann & Memmert, 2014; Smith et al., 2016). Nonetheless, our understanding of the impact that various markers of internal and external load have upon the decision-making performances of soccer officials remains limited.

2.4.1. Decision-Making Accuracy of Team Sport Officials in Relation to Markers of External Load

To examine the impact that high external loads have upon the decision-making performances of team sport officials, researchers have typically ascertained decision-making accuracy across 10–20-minute periods and assessed against the physical outputs achieved during that same period. Currently though, the association between decision-making accuracy and physical match demands remains unclear (Mascarenhas et al., 2009; Elsworthy et al., 2014b; Emmonds et al., 2015). In a small sample of seven soccer matches, Mascarenhas and colleagues (2009) revealed decision-making accuracy to be unrelated to the average running speed and total distance accumulated during each 15-minute period of the match. Such findings should however be interpreted with caution. Firstly, as the physical demands imposed upon soccer officials are best understood when total distance is separated into individualised speed-related categories, measures of total distance and mean running speed may not provide the most accurate reflection of physical match demands (Castagna et al., 2007). Alternatively, examining the impact of high-intensity activities such as high-speed running, sprinting, and changes of direction may prove more fruitful when examining the potential association between physical exertion and decision-making. It is also noteworthy that the activity profile of each official was ascertained using GPS technology that collected positional data at a sampling frequency of 1-Hz. Whilst 1-Hz GPS may accurately quantify total distance, such frequencies may not provide accurate measures of the distance covered by means of high-intensity activities (Coutts & Duffield, 2010). GPS systems possessing greater sampling rates are therefore necessary to accurately identify high-intensity activities such as high-speed running, sprinting, and changes of direction in relation to decision-making (Jennings et al., 2010).

Using 10-H_z GPS analysis, the decision-making accuracy of 8 rugby league officials was assessed against the volume of high-intensity running performed across each 10-minute period of match play, with trivial correlations found (Emmonds et al., 2015). In a similar investigation, the intermittent demands of match play did not appear to affect the decision-making performances of Australian football umpires, as despite a reduction in the volume of high-speed running performed during the final 20-minute period, decision-making accuracy remained consistent across each quarter (Elsworthy et al., 2014b). Nonetheless, it is perhaps inappropriate to construe these findings as indicative of physical load having no effect upon decision-making, as shortcomings in the analytical approaches adopted may have inhibited the detection of any associations. Firstly, the speed thresholds used to identify high-intensity efforts in the aforementioned studies were based upon those used with the players. It has been suggested, however, that the application of absolute speed thresholds may under- or over-estimate the true energetic demands exhibited amongst individuals, therefore leading to the erroneous evaluation of external load (Abt & Lovell, 2009). Given potential age-related disparities in the physical capacities of players and match officials (Castagna et al., 2007), the adoption of individualised speed thresholds appears particularly important within such cohorts. Additionally, as the movement profiles associated with match play are non-uniform and involve varying volumes of high-intensity activity, exploring potential associations during prolonged 10–20-minute periods are unlikely to be sensitive enough to highlight periods of transient fatigue. Although each high-intensity episode lasts only seconds (Barbero-Álvarez et al., 2012), it is plausible that such exertions may temporarily impair an official's decision-making. Thus, examining the accuracy of decisions made immediately following acute periods of high physical exertion may permit a better understanding of the impact that high physical loads have upon the decision-making performances of match officials.

Several studies have previously explored the decision-making performances of team sport officials in relation to the running speeds exhibited at the time of (or in the moments prior to) a decision. When the free kick decisions of 29 Australian Football umpires were assessed, the instantaneous velocity of the umpire at the time of the incident was found to have no effect upon their decision-making accuracy (Elsworthy et al., 2014a). Similar findings have since been revealed within a cohort of 11 soccer FR whereby no associations were observed between the running speed of the official at the time of the infringement and the correctness of their decision (Riiser et al., 2019). Interestingly, these results contradict those of Oudejans and colleagues (2005) who reported the accuracy of AR's offside decisions to be compromised at higher running speeds. Although speculative, such discrepancies may reflect differences in both the type of official (FR/AR) and the type of decisions (foul play/offside) studied. In relation to foul play situations faced by FR, these decisions typically manifest over longer periods of time whereby the official may react and position themselves accordingly. During offside decisions, however, AR must keep up with the offside line whilst simultaneously attending their attention to play. It therefore appears plausible that at higher running speeds, AR may be inclined to direct greater attention to their own position and movement, thus diverting attention away from the incident at the critical moment (Müller et al., 2007). Previous studies have also examined whether decision-making accuracy is associated with a match official's running speed (Paradis et al., 2016) or distance covered (Riiser et al., 2019) in the moments preceding a decision. Although decision-making accuracy appears to be unrelated to the average running speeds and distances covered in the 30 seconds prior to each decision (Riiser et al., 2019; Elsworthy et al., 2014a; Paradis et al., 2016), higher relative running speeds performed in the 5 seconds before a decision increased the likelihood of an error being made (Elsworthy et al., 2014a). It was suggested that this might relate to an enhanced physiological

demand; however, as no measures of internal load were obtained, this remains speculative and requires further investigation.

2.4.2. Decision-Making Accuracy of Team Sport Officials in Relation to Markers of Internal Load

To assess the relationship between the physical and decision-making performances of team sport officials, researchers have examined decision-making accuracy in relation to the specific 10–15-minute period of the match, with findings proving equivocal. Having assessed the *in-situ* decision-making performances of 10 elite soccer FR, Mallo and colleagues (2012) reported the greatest error rates to occur in the final 15-minute period of the match. Similar findings were evidenced within rugby league officials whereby the lowest accuracy was observed in the final 10 minutes (Emmonds et al., 2015). Under the assumption that levels of physical fatigue would be elevated towards the end of a match, these findings were interpreted as suggesting physical exertion to bear a negative impact upon decision-making performance. Conversely, Mascarenhas and colleagues (2009) highlighted a reduced decision-making accuracy during the opening 15-minutes of match play, with this purported to reflect some form of warm-up effect. Data obtained from soccer AR (Catteeuw et al., 2010c) and Australian football umpires (Elsworthy et al., 2014b; Larkin et al., 2014b) suggests decision-making accuracy remains relatively consistent over the course of a match. As the physical demands of sports officiating are driven primarily by contextual factors relating to the movement patterns of the players, such inconsistencies are perhaps unsurprising given the large levels of between-match variation prevalent in match physical performances (Weston et al., 2011b). To explore the impact of physical exertion upon the decision-making performances of match officials, it would therefore appear prudent to assess decision-making accuracy in relation to objective markers of physiological load.

In an analysis of 99 decisions made by a Spanish Third Division FR during four games, decisional errors were reported to occur as a result of the official's HR being in excess of 75% of their maximal (Gómez-Carmona & Pino-Ortega, 2016). Nevertheless, as HR analyses revealed 95% of match play to have been spent above 75% of maximal HR (HR_{max}), these findings should be interpreted with caution. In a cohort of rugby league officials, Emmonds et al (2015) reported trivial correlations between measures of mean HR and decision-making accuracy when averaged over each 10-minute interval of match play. These data were interpreted as suggesting no association between decision-making performance and physiological load; however, such analytical approaches are likely insufficient to adequately assess the intricacies of such a relationship. Although a 10-minute analysis within a naturalistic environment remains the most sensitive to date, aggregating data over prolonged fixed-time epochs may conceal acute periods of high physiological stress that could potentially hinder optimal decision-making. Indeed, recent evidence suggests that discrete fixed-time epochs may underestimate physical match demands when compared to rolling averages (Doncaster et al., 2019; Fereday et al., 2020). It has also been revealed that significantly greater peak values are identified for various measures of physical and physiological load when using 1-min rolling averages as compared to 3-, 5-, or 10-minute rolling averages (Delaney et al., 2017; Martin-Garcia et al., 2018; Doncaster et al., 2019). This latter finding is perhaps unsurprising and reflects the dynamic, stochastic nature of match play whereby the activity profile of players and officials are influenced by the specific match situation (Carling & Dupont, 2011). It is also important to note that the duration and frequency of high-intensity actions such as sprinting are bound by the limitations of the anaerobic energy system and the onset of fatigue (Mohr et al., 2005). To explore the impact that high physical and physiological loads have upon the correctness of the officials' decisions, it would therefore appear appropriate to assess levels of physiological load in the moments immediately preceding each decision (<60 s). Thus, whilst

extensive efforts have been made to identify the psychological stressors that influence an official's decisions, further work is required to examine the decision-making performances of soccer officials in relation to physiological performance markers obtained shortly prior to each decision, such as HR, oxygen uptake, and perceived exertion. A key objective of the present thesis was therefore to examine the correctness of foul play decisions made by a cohort of soccer FR in relation to the physical and physiological loads recorded immediately prior to the decision.

2.5. CURRENT RESEARCH PARADIGMS: STRIKING THE BALANCE BETWEEN EXPERIMENTAL CONTROL AND ECOLOGICAL VALIDITY

Significant scientific efforts have been undertaken to explore the multifaceted factors that impact the decision-making performances of soccer officials. Methodologically, this work is characterised by a dichotomy between research paradigms that emphasise either ecological validity or experimental control. In considering the findings of the aforementioned literature, it is therefore important to acknowledge the merits and limitations of each approach. To achieve an accurate reflection of the demands imposed upon soccer officials, and to ensure that the integrity of their decision-making processes remain intact, research should perhaps be conducted within the naturalistic environment. Conducting research within such settings, however, presents a number of logistical challenges as collection procedures may interfere with normal refereeing duties (Rollo et al., 2014). Drawing meaningful conclusions from competitive matches is also difficult given their highly variable nature and the infinite number of contextual factors that may confound the soccer officials' decisions. An alternative approach involves the administration of simulated representative tasks within controlled laboratory environments. Although such approaches lend themselves to enhanced experimental control, methodological criticisms are noted regarding the fidelity of simulation-

based protocols and the transferability of findings derived from laboratory settings to competition (Vilar et al., 2012; Piggott et al., 2019). To enhance the ecological validity of simulation protocols, it has been argued that the properties of the task should reflect those of the environment to which the results are to be generalised (Vilar et al., 2012). As such, a key focus of the present thesis was to design simulated tasks that closely reflect the multidimensional nature of soccer officiating.

2.6. IMPLICATIONS FOR THE TRAINING AND DEVELOPMENT OF SOCCER OFFICIALS

Although decision-making embodies the most critical component of a soccer official's role, little time is typically apportioned to specific decision-making training, with officials generally exhibiting a preference for physical conditioning with a weak culture of skill practice (MacMahon et al., 2007; Catteeuw et al., 2009b; Weston et al., 2011d). Traditionally, the decision-making training of soccer officials has focused on the development of declarative and procedural knowledge (i.e., laws of the game) through lecture-style meetings, with practical experience being limited to the officiating of competitive matches (MacMahon et al., 2007). Training and testing protocols that effectively develop the perceptual-cognitive skills of match officials under the psychological and physiological conditions experienced during match play are therefore warranted.

Given the disconnect between the importance of perceptual-cognitive expertise and the time apportioned to decision-making training, considerable interest has been paid within recent years to the development of video-based methods for developing and assessing the decision-making skills of team-sport officials (Kittel et al., 2019c). To date, video-based programmes have been devised primarily for soccer FR (Schweizer et al., 2011; Gulec & Yilmaz, 2016; Van Biemen et al., 2018), AR (Catteeuw et al., 2010b; Put et al., 2013; Put et al., 2014; Put et al.,

2016), and Australian football umpires (Larkin et al., 2014a; Larkin et al., 2018b; Kittel et al., 2019a), with such training methods suggested to result in improved performance during video-based decision-making tasks (Schweizer et al., 2011; Put et al., 2016; Van Biemen et al., 2018). Though video-based methods possess high levels of construct validity, few studies have reported measures of reliability, therefore making it difficult to draw conclusions as to their efficacy (Kittle et al., 2019b). The extent to which potential enhancements following video-based training are replicated in the form of improved decision-making on the field of play is also unclear. Whilst acknowledging the potential for such methods to be used as alternative and supplemental means of developing the decision-making capacities of match officials within off-field contexts, the fidelity and ecological validity of current methods remains to be elucidated. Typically, video-based training methods are administered within controlled off-field environments whereby the official remains in a rested state, thus isolating the physiological, psychological, and perceptual-cognitive demands. However, as decisions during match play are made under challenging psycho-physiological conditions, such conditions should perhaps be replicated during training (Weston et al., 2012).

In light of the above, a key focus of those responsible for the development and preparation of soccer officials should be to devise training interventions that enhance the soccer officials' ability to make correct decisions under high levels of stress and anxiety. When administered amongst athletes, a recent meta-analysis suggests that psychological interventions typically yield a moderate effect on sports performance, with improvements lasting up to a month following the end of the intervention (Brown & Fletcher, 2017). In relation to soccer officials more specifically, a structured programme of mental skills training was previously reported to have resulted in the improved performance of an elite Scottish FR (Mathers & Brodie, 2011). It should be noted, however, that decision-making performance wasn't explicitly measured; improved performance was simply inferred from the number and level of international match

appointments received during the following season. Psychological interventions previously administered amongst soccer officials have largely focused on the learning and rehearsal of mental skills such as imagery, positive self-talk, and relaxation techniques (Mathers & Brodie, 2011; Samuel, 2015). Notwithstanding their value, the initial learning of such skills typically involves the mental simulation and rehearsal of performance away from the applied context. Alternatively, a case exists for the concurrent performance of decision-making tasks under stressful conditions comparable to match play. The argument being that by exposing individuals to higher levels of anxiety during training, they may become accustomed and desensitised to the distracting processes that are believed to hinder performance. Whilst acknowledging the challenges of replicating the pressures of competitive match play, evidence from studies with elite basketball players suggests that the systematic exposure of athletes to punishment-conditioned stimuli during training may have a positive effect upon perceptual-cognitive performance (Oudejans & Pijpers, 2009; Oudejans & Pijpers, 2010). Nevertheless, whether such approaches would improve the match decision-making performances of soccer officials remains to be seen and represents an exciting avenue for future research. This was however beyond the scope of the present thesis.

To create additional opportunities for soccer officials to make decisions under elevated levels of physiological stress, attempts have been made to develop referee-specific protocols that impose the physical and decision-making stimuli simultaneously. Within a sample of 22 elite Israeli soccer FR, Samuel et al (2019) assessed a novel decision-making simulator whereby the perceptual-cognitive and physiological demands of match play were imposed concurrently. Specifically, officials ran at varying paces on a motorised treadmill whilst concurrently adjudicating on pre-recorded footage of match play. Nonetheless, a number of limitations warrant consideration when interpreting the fidelity of this protocol. In relation to the exercise component of this task, officials alternated 4-minute intervals at 10 $km \cdot h^{-1}$ with 1-

minute intervals at 13 km·h^{-1}, for a total of 60-minutes. Although measures of total distance (10.6 km) were comparable to those reported during actual match play (Castagna et al., 2007), the extent to which this protocol replicates the intermittent activity profile characteristic of a soccer FR is questionable. Indeed, soccer match play necessitates that officials engage in periods of both low-intensity activity such as standing (~0 km·h^{-1}) or walking (~6 km·h^{-1}) as well as periods of high-intensity activity such as high-speed running (~18 km·h^{-1}) or sprinting (>21 km·h^{-1}) (Krustrup et al., 2009; Barbero-Álvarez et al., 2012). The protocol adopted by Samuel and colleagues (2019) therefore appears to be of a relatively moderate-intensity and excludes the frequent activity changes inclusive of the high-intensity activities that may prove counterproductive to an official's optimal decision-making. Concerns also exist with regards to the fidelity of the decision-making task. As is perhaps common within off-field studies examining the decision-making performances of team sport officials, this protocol involved the presentation of match broadcast footage filmed from an elevated position in the grandstand. Such approaches have however been criticised as they do not replicate the in-game perspective of the official (Craig, 2013). Considerable scope therefore exists for the further development and refinement of simulation-based training protocols that combine decision-making tasks with the challenging physiological conditions that characterise soccer match play. As such, a key objective of the present thesis was to develop a valid and reliable protocol that enabled officials to make match-related decisions in unison with the physical and physiological conditions that are representative of competitive match play.

2.7. CONCLUSIONS AND PERSPECTIVE

To summarise, soccer officiating represents a highly challenging activity during which complex perceptual-cognitive processes are undertaken in combination with high levels of physiological and psychological stress. As highlighted by the foregoing review of the literature,

these competing demands place an additional strain on the officials' attentional resources and may therefore compromise the correctness of their decisions. A number of important implications therefore exist for the training and development of soccer officials. Firstly, to acquire the skills and capacities required to cope with the breadth of demands encountered during match play, a holistic approach to training is necessary whereby each facet of performance is developed simultaneously. It is therefore vital that those responsible for the preparation and development of soccer officials are able to monitor the training process closely as equipped with a detailed understanding of current practices, training may be adjusted and diversified accordingly. At present, however, the training practices engaged in by soccer officials are not well documented and there does not currently exist a standardised and valid method of monitoring the wide range of activities pertinent to this cohort (i.e., physical, decision-making, psychological, and technical skills training). It is therefore crucial that these gaps within the literature be addressed if researchers and practitioners are to positively impact the future training and development of soccer officials. Thus, two of the key objectives of the present thesis were to develop a valid and reliable self-report questionnaire that could be used to monitor the multifaceted and wide-ranging training practices of soccer officials, and then use this newly developed tool to provide an updated account of the training practices currently engaged in by Scottish soccer officials. Meanwhile, it also stands to reason that the officials' training should be focused on preparing for the types of situations and contexts that hinder their decision-making most. The rationale being that by increasing their exposure to these "worst-case" scenarios, officials may become accustomed and desensitised to the most challenging periods of match play. However, whilst extensive efforts have been made to identify the psychological stressors that influence an official's decisions, further work is required to elucidate the impact that heightened levels of physical and physiological load have on the decision-making performances of soccer officials. Previous literature exploring the association

between physical load and decision-making has generally been conducted in-situ. Though such approaches offer high levels of ecological validity, the array of environmental and contextual factors that characterise competitive match play render it difficult to delineate the influence of physical load on the in-match decisions of soccer officials. An alternative approach that is often used to circumvent such challenges is through the use of simulation protocols that closely mimic the demands of match play. In addition to negating the large levels of variability that confound match data, simulation protocols enable physiological responses to be monitored in a more controlled environment. Unfortunately, existing simulation protocols have been devised exclusively for soccer players, with a valid and reliable protocol yet to be established for the officials. A need therefore exists for the development of a novel simulation protocol that reflects both the physical and decision-making aspects of soccer officiating. With the above in mind, it is the aims of this thesis to firstly develop valid and reliable tools that may assist in the multidisciplinary evaluation of the training and performances of soccer officials, and secondly, to use these tools to reveal important new insights into the training and performances of soccer officials. Such insights may then be used to help inform and guide the training process.

CHAPTER 3

Development and Validation of the Referee Training Activity Questionnaire (RTAQ): Towards a Better Understanding of the Training Practices of Soccer Officials

Publications arising from this chapter:

McEwan, G. P., Unnithan, V., Easton, C., & Arthur, R. (2020). Development and validation of the Referee Training Activity Questionnaire (RTAQ): Towards a better understanding of the training practices of soccer officials. *Journal of Sport Sciences,* 38(24), 2782-2793. https://doi.org/10.1080/02640414.2020.1800371

Conference proceedings arising from this chapter:

McEwan, G. P., Unnithan, V., Easton, C., & Arthur, R. (2019). Development and validation of the Referee Training Activity Questionnaire (RTAQ): Towards a better understanding of the training practices of soccer officials. *24th annual Congress of the European College of Sport Science*, Prague, Czech Republic, 3rd – 6th July 2019.

3.1. ABSTRACT

Purpose: To develop and validate a novel self-report measurement tool for assessing the training practices of soccer officials. **Methods:** The Referee Training Activity Questionnaire (RTAQ) was developed and validated following a systematic multi-stage process involving: 1) item generation; 2) assessments of content and face validity; and 3) assessments of criterion validity. In stage 1, items were generated following a review of the literature and semi-structured interviews with an expert panel (n=8). Items were then assessed for measures of content and face validity in stage 2. These assessments were undertaken by a sample of elite soccer officials and academics experienced in questionnaire design (n=6), with the content validity index (CVI) subsequently calculated. To ascertain levels of criterion validity in stage 3, the RTAQ was completed by a cohort of officials (n=25) who subsequently recorded a detailed 7-day training diary. **Results:** Following content analyses in stage 1, the preliminary RTAQ was developed and comprised 3 primary sections (12 sub-sections) assessing: 1) attributes perceived to underpin soccer officiating performance; 2) general training information; and 3) specific training practices. In stage 2, content validity was confirmed for 8 sub-sections (CVI $\geq$ 0.78) with 5 sub-sections being deemed invalid (CVI < 0.78). Various amendments were carried out in accordance with participant feedback. In stage 3, negligible mean biases, wide 95% LOA, and significant Pearson correlations were observed between the RTAQ and training diaries for most training activities. **Conclusion:** Taken together, the present findings provide initial evidence that the RTAQ holds promise as a valid and effective method of acquiring insight into the multifaceted training practices of soccer officials.

3.2. INTRODUCTION

Soccer officiating represents a highly complex and demanding activity characterised by the adjudication of match play within physiologically and psychologically challenging conditions (Helsen & Bultynck, 2004). For instance, in addition to making ~137 observable decisions per match (Helsen & Bultynck, 2004), soccer FR typically cover total distances of 9-11 km (Krustrup et al., 2009; Weston et al., 2011c; Castillo et al., 2017), with a corresponding mean match HR of 85-95% of HR_{max} (Weston et al., 2006; Castillo et al., 2016b). Conversely, although their movement is limited to half of the length of the field, distances of 5-7 km are covered by AR, with mean match HRs of ~79% of HR_{max} achieved (Castillo et al., 2016b). Furthermore, as matches are often contested in the presence of large partisan crowds, the soccer officials' decisions receive routine scrutiny from players, managers, and spectators alike (Webb et al., 2020). Well-developed physical and psychological attributes therefore serve as important requisites of the successful match official (Castagna et al., 2007; Slack et al., 2013), whilst perceptual-cognitive skills are integral to successful decision-making within dynamic, time-constrained domains (Spitz et al., 2018b). To succeed at the top level, soccer officials must also possess sufficient technical knowledge (i.e., knowledge of the Laws of the Game, tactical awareness), whilst game-management skills are necessary for *"managing"* the game effectively (Unkelbach & Memmert, 2008). To acquire expertise within each of these various facets of performance, commitment to training and practice is important (Baker & Young, 2014). Central to the magnitude of the adaptive responses achieved from training are the volume, intensity, frequency, and type of practice undertaken (Hawley, 2008). Additionally, as the greatest improvements in performance are achieved when training simulates the demands of competition, an integrated approach to training is advised, whereby the physical, perceptual, and psychological elements of soccer officiating are addressed in unison (Weston et al., 2012). Given the association between training and performance, understanding the habitual training

practices of soccer officials is important if researchers and practitioners are to successfully evaluate and adjust training, with the intention of optimising performance (Saw, Main, & Gastin, 2015).

In comparison to the volume of literature pertaining to players, very little information is readily discernible on the training practices of soccer officials. Generally, the training of soccer officials is self-led and focuses on physical training activities, with less attention paid to the development of other skills pertinent to refereeing (Weston et al., 2012). As such, efficacious and accessible methods of measuring a wider range of training are required in order to monitor and then diversify and enhance the training of soccer officials. Self-report measures such as training diaries have long been utilised within applied sport science research as simple and inexpensive methods of acquiring insight into an athlete's training (Pugliese et al., 2014). However, from a data management and interpretation perspective, their use with large cohorts typically proves problematic given the considerable volume of data generated (Mujika, 2017). Self-administered questionnaires represent an alternative means of collecting such information (Helsen et al., 1998; Ford et al., 2010). Of note is the deliberate practice questionnaire which having been developed for soccer players (Helsen et al., 1998), has previously been implemented with soccer officials (MacMahon et al., 2007). Nonetheless, the appropriateness of the deliberate practice questionnaire for use with this cohort is questionable. Whilst the items of this questionnaire were adapted to the role of the soccer official following consultation with two referee coaches (MacMahon et al., 2007), a more robust process of item generation appears warranted to ensure their specificity to both FRs and ARs (Saw et al., 2017). Specifically, the inventory of activities listed within the deliberate practice questionnaire were categorised as on-field and off-field, with a large emphasis placed upon physical training activities. In comparison, little detail was provided surrounding activities unique to officials, such as specific decision-making, psychological, technical, and game-management skills

training. Thus, a clear need exists for the development of a valid measurement tool with the potential to assess these crucial yet often overlooked aspects of a soccer official's training and preparation.

In order to develop a valid instrument that sufficiently addresses each aspect of a soccer official's training, a participatory approach is advocated whereby the end-users (i.e., both FR and AR) are actively involved in the development and assessment of its psychometric properties (Shrier et al., 2014). Of the properties suggested to underpin the validity of an instrument are the content validity (the extent to which the questionnaire is deemed to cover all relevant topics of the construct in question) and face validity (the extent to which items appear to provide an adequate reflection of the construct in question) (Mokkink et al., 2010). It is also important that novel instruments be assessed for sufficient levels of criterion validity i.e., the degree to which the results of a questionnaire provide an adequate reflection of a *"gold standard"* (Mokkink et al., 2010). Notably, as questionnaires previously employed to assess the training practices of soccer officials have not been systematically developed and validated for use with soccer FR and AR (MacMahon et al., 2007), scope exists for a new tool to be developed and validated for specific use with such populations. The current investigation therefore aimed to develop and validate a novel measurement tool (the Referee Training Activity Questionnaire) for assessing the training practices of soccer officials.

3.3. METHODS

3.3.1. Experimental Approach

The Referee Training Activity Questionnaire *(*RTAQ*)* was developed and validated following a systematic multi-stage process (Artino et al., 2014): 1) item generation and questionnaire development; 2) assessments of content and face validity; and 3) assessments of criterion validity (Figure 3.1). During each stage, participants were provided with a detailed

information sheet outlining the purpose of the study and informed consent was obtained (see Appendices A and B for an example). Institutional ethical approval was acquired for all phases of the study.

3.3.2. Stage 1: Item Generation and Questionnaire Construction

In accordance with previous research (MacNamara & Collins, 2011), content domains were defined, and an item pool was generated following: 1) a review of the literature (please refer to Chapter 2); and 2) consultation with experts informed on the attributes and training practices pertinent to soccer officials. To conceptualise the scope of the questionnaire and identify the primary facets of officiating performance, a literature search of PubMed and Web of Science databases was performed using various combinations of issue-related search terms. Search terms included: (*"referee"* OR *"assistant referee"* OR *"match official"*) AND (*"soccer"* OR *"association football"*). Secondary searches included manual searches of the reference lists of the recovered articles (Moher et al., 2015). Following this review of the literature, a thematic analysis of the recovered articles was undertaken by the lead author (GM) whereby articles were arranged into common themes. This process resulted in the identification of five themes deemed to underpin soccer officiating performance, with these categorised as: physical, decision-making, psychological, technical, and game management. This information was subsequently used to guide semi-structured interviews with experts and build an inventory of items from which to construct the questionnaire.

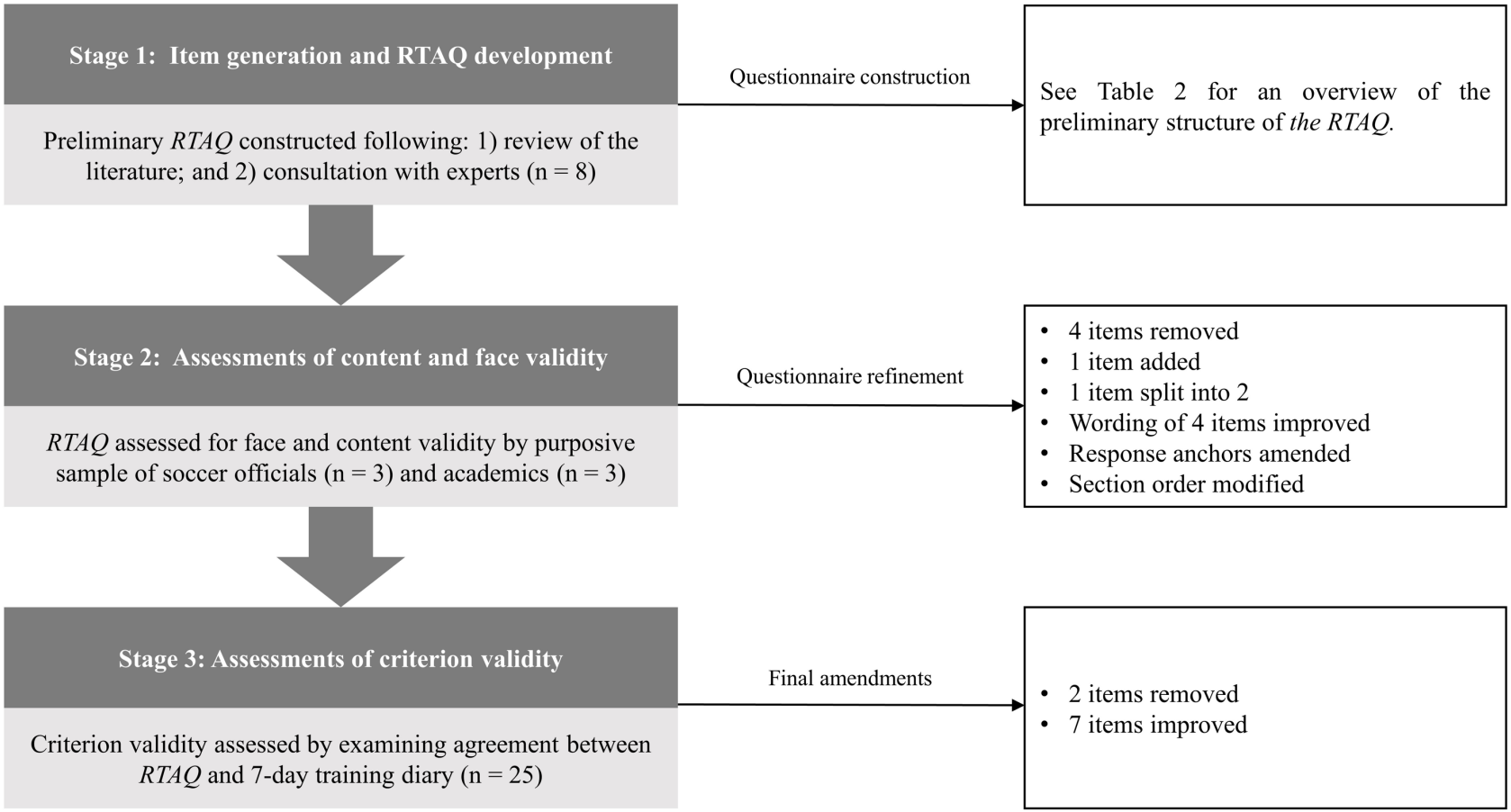


Figure 3.1. Flow chart of multi-stage process used to develop and validate the Referee Training Activity Questionnaire (RTAQ).

3.3.2.1. Participants

Eight experts, purposefully sampled to incorporate both elite soccer officials and referee managers/coaches, were invited to participate in the preliminary stage of this investigation. Although this cohort were exclusively male, it is important to note that this simply reflects our limited access to elite participants, and that no criteria existed for the exclusion of females. At the time of interviewing, all officials were active elite FR (n=2) and AR (n=2) within Scotland, had officiated within the Scottish Premiership for 12 ± 3 years, and possessed 8 ± 6 years of experience on the FIFA list. A sample of referee managers/coaches (national manager: n=1; development officer: n=1; national referee fitness instructor: n=1; and head of sports medicine: n=1) were also interviewed on the basis that they were responsible for the promotion/demotion of officials between the different professional levels, and for the provision of training support for officials.

3.3.2.2. Interview Procedure

Interviews were semi-structured in nature and were directed by an interview guide developed and revised by the research team following the aforementioned review of the literature (Appendix C). Accordingly, a list of open-ended and non-leading questions were used to assess the training practices and attributes perceived as pertinent to performance, with probes seeking elaboration or clarification implemented to follow up on promising leads (Patton, 2001). All interviews were conducted face-to-face by the lead researcher in a quiet and convenient location and lasted 27 ± 3 min (range: 23–33 min). With the participants' permission, interviews were audio recorded.

3.3.2.3. Data Analysis

Qualitative data were analysed using concurrent deductive and inductive content analysis (Sparkes & Smith, 2013) whereby the analysis was based upon two primary (attributes and training practices) and five secondary (physical, decision-making, psychological, technical, and game-management) *a priori* research themes, whilst remaining open to emergent findings within participants' responses. Firstly, audio-recordings were transcribed verbatim and were subsequently double-checked to ensure accuracy. Close reading of the text was then undertaken by the lead researcher to ensure familiarity with the data. Raw data units were then created from participants' words before being grouped into lower-order themes and then higher-order themes. During analysis, internal homogeneity (that data within a category share clear characteristics) and external heterogeneity (clear differences exist between different categories) was sought (Patton, 2001). To ensure the trustworthiness of the analyses, the authors discussed and confirmed the allocation of raw data units to specific categories through constructive debate (Elo et al., 2014).

3.3.3. Stage 2: Content and Face Validity

Based upon the findings of stage 1, items were generated directly from participants' wording to ensure that they remained in accordance with participants' perspectives and understanding of the construct in question (Artino et al., 2014). A preliminary draft of the RTAQ was subsequently constructed (Table 3.1). Stage 2 sought to assess the content and face validity of the preliminary RTAQ.

Table 3.1. Preliminary structure of the Referee Training Activity Questionnaire (RTAQ).

	Items (*n*)	Question	Response scale
Section 1. General training practices			
1.1. Domain-specific activities	6	Please indicate the number of times (*within the past two weeks*) and the average duration per session that you have engaged in each of the following activities.	Open-ended
1.2. Training environment	5		
1.3. Training objective	5		
Section 2. Specific training practices			
Section 2.1. Training activities			
2.1.1. Physical	10	Please indicate the number of times (*within the past two weeks*) and the average duration per session that you have engaged in each of the following activities.	Open-ended
2.1.2. Decision-making	4		
2.1.3. Psychological	9		
2.1.4. Technical	4		
Section 2.2. Attributes			
2.2.1. Physical	10	Please indicate how often you engage in training that is focused on developing each of the following skills and attributes.	5-point Likert
2.2.2. Decision-making	9		
2.2.3. Psychological	7		
2.2.4. Technical	6		
2.2.5. Game-management	6		

3.3.3.1. Participants

To assess the content and face validity of the RTAQ, a purposive sample of 6 participants were recruited (Polit et al., 2007). This cohort included male soccer officials (FR: n=2; AR: n=1) of varying levels of professional attainment and experience, academics experienced in questionnaire design within sport science (n=2), and an applied sport science practitioner (n=1).

3.3.3.2. Face and Content Validation

The preliminary RTAQ was distributed electronically and participants were requested to review each section based upon its relevance, readability, redundancy, and ease of interpretation. For each section of the questionnaire, participants used a 4-point Likert scale to rate: the clarity and conciseness of its wording *(1=not clear; 4=very clear)*; the relevance of the item to understanding the training practices of officials *(1=not relevant; 4=very relevant)*; the relevance of the items to officiating *(1=not relevant; 4=very relevant)*; the extent to which the items matched the descriptions provided *(1=poor match; 4=good match)*; and the degree to which the scales were clear and easy to use *(1=not easy; 4=very easy)* (Dunn et al., 1999). Participants were also invited to provide any additional comments or suggestions that they may have had (Arnold et al., 2013).

3.3.3.3. Data Analysis

Likert scale data were dichotomised into "*acceptable*" (rating of 3–4) and "*not acceptable*" (rating of 1–2) and the content validity index (CVI) of each section was calculated (Polit & Beck, 2006). Specifically, the CVI of each section was determined as the number of participants providing a rating of 3–4 (agreed) divided by the total number of participants. As recommended by Polit and Beck (2006), content validity was defined as a CVI ≥ 0.78.

3.3.4. Stage 3: Criterion Validity of the Referee Training Activity Questionnaire

To assess the RTAQ for criterion validity, estimates of training activity involvement (type, frequency, and duration) were compared to those documented during a detailed 7-day training diary (Halson, 2014). During the data collection period, officials maintained their usual training activities and habits. As part of their continued sport science support, officials utilised HR monitors (Polar H10, Finland) during their physical training sessions to record measures of internal (HR) and external (session duration, total distance covered) load. Specifically, data pertaining to the duration and type of activity was recorded, with HR data being used to help identify and confirm the type of session. Consequently, objective data obtained from the HR monitors during physical training sessions was used to corroborate the self-report data of the training diaries (Borresen & Lambert, 2009).

3.3.4.1. Participants

Scottish male FR (n=19) and AR (n=12) spanning various levels of professional attainment and experience agreed to participate in this stage of the investigation (Table 3.2).

3.3.4.2. Referee Training Activity Questionnaire (RTAQ)

The RTAQ was developed using an online survey tool (Online Surveys, Jisc, UK) and was distributed to participants via a web link towards the end of the 2017-18 competitive season. Participants were permitted two months (April and May) to complete the questionnaire. Although pertinent to consider seasonal variation in training practices (Sæther & Aspvik, 2014), the distribution of the questionnaire towards the latter stages of the season was not deemed problematic in the current context given our interest in comparing training at a single time-point.

Table 3.2. Demographic information of Referee Training Activity Questionnaire cohort (n=31).

	Mean (SD)	Range	Number (proportion who completed training diary)
Age (years)	34 ± 7	24 – 49	
Officiating experience (years)	12 ± 6	3 – 31	
Officiating role			
FR			19 (73.7)
AR			12 (91.7)
Category			
Category 1 (FIFA)			4 (75.0)
Category 1			5 (80.0)
Category 1 (Development)			4 (75.0)
Category 2			1 (0.0)
Category 2 (Development)			3 (100.0)
Category 3 SAR (FIFA)			2 (100.0)
Category 3 SAR			9 (88.9)
Category 3 SAR (Development)			1 (100.0)
Category 3			1 (0.0)
Category 3 (Development)			1 (100.0)
Highest level officiated			
Intercontinental			1 (100.0)
Continental			10 (80.0)
National			17 (82.4)
Regional			3 (66.7)

Data are presented as mean (SD), range, or number (proportion who completed training diary). FR, field referee. AR, assistant referee. SAR, specialist assistant referee.

3.3.4.3. Training Diary

Following the completion of the RTAQ, participants completed a training diary whereby they were asked to record their training activities for a period of seven consecutive days (Halson, 2014). It was deemed necessary to administer the RTAQ and training diaries in this order to: 1) avoid the considerable respondent fatigue that would likely occur had the RTAQ and training diaries been completed concurrently; and 2) minimise potential biasing and priming effects that would occur had the training diaries been completed first. Accordingly, seven 24-hour diary sheets were provided to participants who were asked to be as detailed as possible with regards to the type, duration, intensity, and conditions of training. To minimise the risk of risk of report bias, the recommendations of Podsakoff et al (2012) were followed where participants received clear instructions of how to complete the diary, including examples of the level of detail sought. Participants were asked to complete the diary sheets at the end of each day, and to be consistent with the procedure. Upon completion of the collection period, participants were asked to comment on whether the information they provided was representative of a typical training week. To avoid the confounding influence of seasonal variations in training activities, the maximal time interval between the completion of the RTAQ and the training diary was one month (Saether & Aspvik, 2014). Training diaries were subsequently analysed by the lead author (GM).

3.3.4.4. Data Reduction and Analysis

Thirty-one participants completed the RTAQ; however, 6 participants failed to complete the training diary. Consequently, the data of 25 participants were included for analyses. Prior to the use of parametric statistical test procedures, Shapiro-Wilk's tests were used to verify that the assumption of normality of distribution was satisfied. To corroborate the self-report data of the training diaries with the objective data obtained from the HR monitors, associations

between the total training volumes recorded for each physical training activity were assessed using Pearson product-moment correlations. The relationships between the RTAQ and the training diaries, with respect to the total training volumes reported for each distinct training activity listed within the RTAQ, were also examined via Pearson correlations, with uncertainty in the estimates expressed as 95% confidence intervals (CI). Correlations of 0.10–0.29, 0.30–0.49, 0.50–0.69, 0.70–0.89, and ≥0.90 were considered small, moderate, large, very large, and nearly perfect (Hopkins et al., 2009). Paired sample t-tests determined whether any systematic bias existed between the outputs of the RTAQ and the training diaries, with 95% CI also calculated. Additionally, agreement between the RTAQ and training diaries was assessed using the 95% limits of agreement (LOA) approach (Bland and Altman, 2007). Statistical significance was set at $P < 0.05$ with statistical procedures completed using Statistical Package for Social Sciences (SPSS 22.0, IBM, USA).

3.4. RESULTS

3.4.1. Stage 1: Item Generation and Questionnaire Construction

3.4.1.1. Thematic Analyses

Following content analyses, 12 lower-order themes were identified and were subsequently grouped into three higher order themes: 1) attributes perceived to underpin soccer officiating performance; 2) general training information; and 3) specific training practices. The attributes perceived to underpin soccer officiating performance were categorised as: a) physical; b) decision-making; c) psychological; d) technical; and e) and game-management. General training information comprised: f) domain-specific activities; g) training environment; and h) training objective. Finally, the specific training practices of officials were categorised as: i) physical training; j) decision-making training; k) psychological training; and l) technical training. An overview of the item pool generated during phase 1 is presented in Figure 3.2.

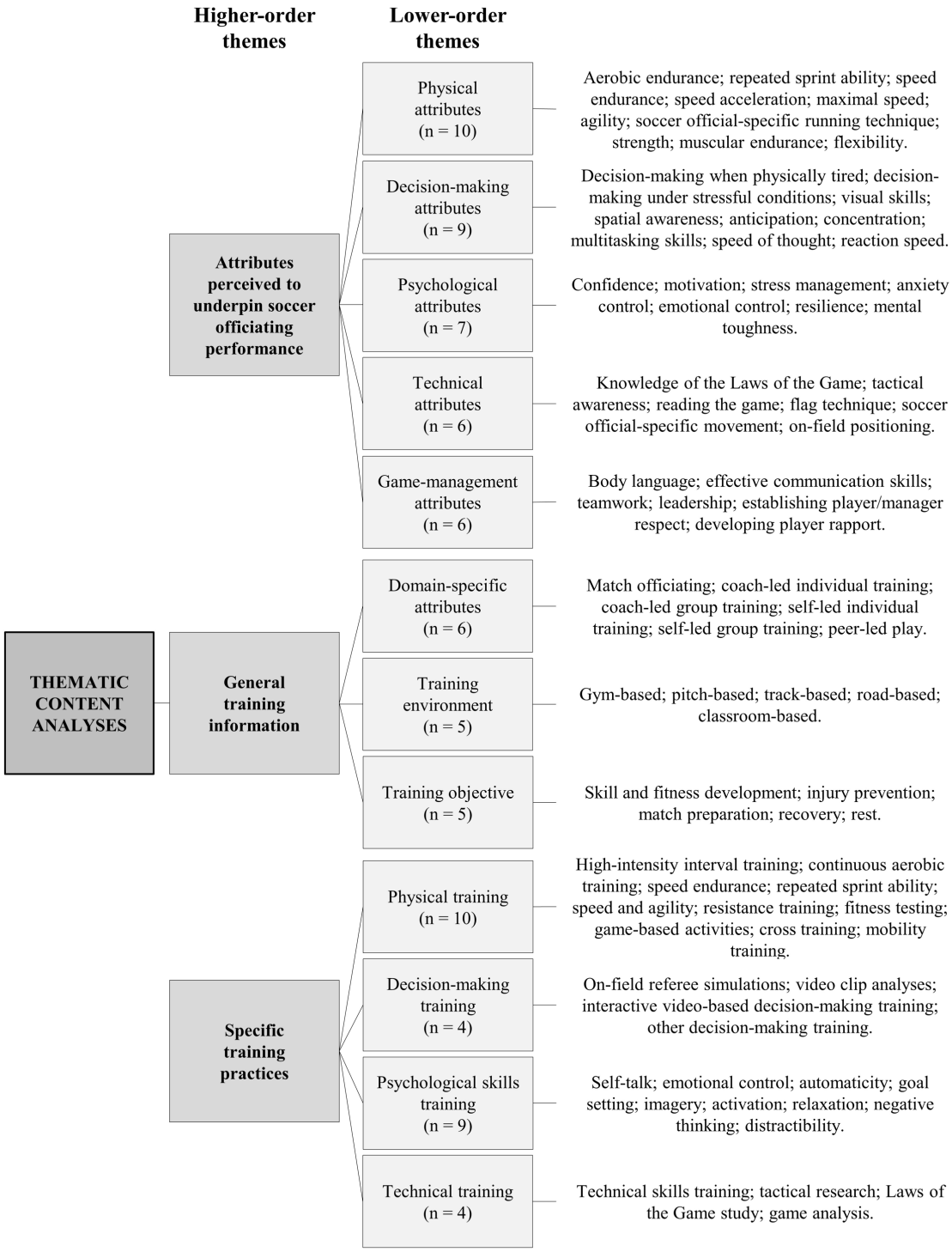


Figure 3.2. Item pool generated following thematic content analyses of semi-structured interviews.

3.4.1.2. Questionnaire Construction

An overview of the preliminary structure of the RTAQ is presented in Table 3.1. Briefly, the preliminary questionnaire comprised 2 primary sections and 12 sub-sections. To quantify the frequency and duration apportioned to various general and specific training practices, open-ended response scales were employed. To examine the regularity by which officials partook in training aimed at developing particular skills and attributes, a 5-point Likert scale (*0 = less than once a month; 4 = more than once a week*) was utilised. Response anchors were fully labelled as per previous recommendations (Wade, 2006; Krosnick & Presser, 2010).

3.4.2. Stage 2: Assessments of Content and Face Validity

3.4.2.1. Content Validity Index (CVI)

When assessed based upon the wording, relevance, and description of questions and items, each section was deemed valid as exhibited by a CVI ≥ 0.78. Of the 12 sub-sections within the questionnaire, eight received a CVI of 1.00, demonstrating complete agreement amongst experts.

When the scales within each sub-section of the questionnaire were assessed based upon their usability, eight were deemed valid (CVI $\geq$ 0.78) with complete agreement being reached amongst experts (CVI of 1.00) for six sub-sections. Of the five sub-sections pertaining to specific training practices, each received a CVI of 0.50 suggesting them not to be valid. Additional feedback received from participants suggested response anchors to be somewhat ambiguous with suggestions for the implementation of clearer response labels. The CVI for each section of the RTAQ is presented in Table 3.3.

Table 3.3. Content validity index (CVI) for each section of the Referee Training Activity Questionnaire (RTAQ) (n=6).

	Please rate the degree to which you believe the wording of the above question is clear and concise (1 = not clear; 4 = very clear)		Please rate the degree to which you believe the above question is relevant for understanding the current training practices of soccer officials (1 = not relevant; 4 = very relevant)		Please rate the degree to which you believe the items listed above are of relevance to soccer officials (1 = not relevant; 4 = very relevant)		Please rate the degree to which you believe the items listed above match the descriptions provided (1 = poor match; 4 = good match)		Please rate the degree to which you believe the scales were clear and easy to use (1 = not easy; 4 = very easy)	
	Mean (range)	CVI	Mean (range)	CVI	Mean (range)	CVI	Mean (range)	CVI	Mean (range)	CVI
1. General training practices										
1.1. Domain-specific	3.7 (3–4)	1.00	3.8 (3–4)	1.00	3.8 (3–4)	1.00	3.7 (3–4)	1.00	3.5 (3–4)	1.00
1.2. Environment	3.7 (3–4)	1.00	4.0 (4)	1.00	4.0 (4)	1.00	3.8 (3–4)	1.00	3.5 (3–4)	1.00
1.3. Objective	3.7 (3–4)	1.00	4.0 (4)	1.00	4.0 (4)	1.00	4.0 (4)	1.00	3.5 (3–4)	1.00
2. Specific training practices										
2.1. Training activities										
2.1.1. Physical	3.7 (3–4)	1.00	4.0 (4)	1.00	3.8 (3–4)	1.00	4.0 (4)	1.00	2.3 (1–4)	0.50
2.1.2. Decision-making	3.8 (3–4)	1.00	4.0 (4)	1.00	3.8 (3–4)	1.00	3.7 (3–4)	1.00	2.3 (1–4)	0.50
2.1.3. Psychological	3.7 (3–4)	1.00	3.8 (3–4)	1.00	4.0 (4)	1.00	3.7 (2–4)	0.83	2.5 (1–4)	0.50
2.1.4. Technical	3.7 (3–4)	1.00	3.5 (2–4)	0.83	3.3 (2–4)	0.83	3.3 (2–4)	0.83	2.3 (1–4)	0.50
2.2. Attributes										
2.2.1. Physical	3.7 (3–4)	1.00	3.8 (3–4)	1.00	3.7 (3–4)	1.00	3.7 (3–4)	1.00	3.5 (2–4)	0.83
2.2.2. Decision-making	3.7 (2–4)	0.83	3.8 (3–4)	1.00	3.8 (3–4)	1.00	3.7 (3–4)	1.00	3.5 (2–4)	0.83
2.2.3. Psychological	3.7 (3–4)	1.00	3.7 (3–4)	1.00	3.7 (3–4)	1.00	3.7 (3–4)	1.00	3.7 (3–4)	1.00
2.2.4. Technical	3.5 (3–4)	1.00	3.7 (3–4)	1.00	3.7 (3–4)	1.00	3.5 (2–4)	0.83	3.5 (2–4)	0.83
2.2.5. Game-management	3.7 (3–4)	1.00	3.8 (3–4)	1.00	3.8 (3–4)	1.00	4.0 (4)	1.00	3.7 (3–4)	1.00

Data are presented as mean (range). CVI, content validity index.

3.4.2.2. Questionnaire Refinement

Following the checks of content and face validity, the questionnaire was revised and refined. Items belonging to sections demonstrating a low CVI were either discarded or improved and retained in accordance with the suggestions of participants. This process resulted in several modifications being made to the initial questionnaire (Figure 3.1). Once amendments were completed, the questionnaire was re-circulated for approval with no further changes requested at this stage. The final RTAQ covered two main themes: 1) general training information (18 items) and 2) specific training practices (63 items).

3.4.3. Stage 3: Assessments of Criterion Validity

3.4.3.1. Validity of the Training Diary Data

Significant Pearson correlations ($r > 0.96$, $P < 0.001$) revealed the estimates of physical training activity derived from the training diary and the HR monitors to be strongly related, thus supporting the validity of this data.

3.4.3.2. Agreement Between the RTAQ and Training Diary

Section 1 (general training information): Negligible mean biases (range: -22 to 6 min), wide 95% LOA (range: -155 to 111 min), and significant Pearson correlations ($r > 0.53$, $P < 0.006$) were found between the RTAQ and training diaries for most items relating to domain-specific activities, training environment, and training objective. A significant systematic bias ($P = 0.015$) revealed the RTAQ to provide lower estimations of time spent in classroom-based training by 25 min (95% LOA: -119 to 69 min). A significant systematic bias ($P = 0.001$) also revealed the RTAQ to yield lower estimations of time apportioned to skill development sessions by 67 min (95% LOA: -237 to 104 min). The agreement between the RTAQ and training diaries for items pertaining to general training information are presented in Table 3.4.

Section 2 (specific training practices): Negligible mean biases (range: -10 to 16 min), wide 95% LOA (range: -122 to 153 min), and significant Pearson correlations ($r > 0.55$, $P < 0.005$) were found between the RTAQ and training diaries for most physical, decision-making, psychological skills, and technical training activities. A significant systematic bias ($P = 0.049$) revealed the RTAQ to provide higher estimations of time apportioned to game-based activities by 37 min (95% LOA: -138 to 212 min). A significant systematic bias ($P = 0.030$) also revealed the RTAQ to yield higher estimations of time apportioned to imagery by 4 min (95% LOA: -14 to 22 min), as well as time apportioned to relaxation ($P = 0.046$) by 5 min (95% LOA: -18 to 28 min). The agreement between the RTAQ and training diaries for items relating to specific training practices are presented in Table 3.5.

3.4.3.3. Final Questionnaire

Following assessments of criterion validity, several modifications were carried out (Figure 3.1). Completion time for the questionnaire was 23 ± 7 min. The final version of the RTAQ is available within Appendix D.

Table 3.4. Systematic bias, 95% limits of agreement (LOA), and Pearson correlations for training volumes of general training practices (n=25).

	Assessment method		Bias (95% CI)	LOA -	LOA +	*r* (95% CI)
	RTAQ (min/week)	Training diary (min/week)				
Section 1. Domain-specific activities						
Match officiating	142 ± 63	148 ± 70	-6 (-27 to 15)	-105	93	0.72 (0.45 to 0.87) ***
Coach-led individual training	17 ± 35	11 ± 34	6 (-5 to 18)	-49	61	0.68 (0.39 to 0.85) ***
Coach-led group training	54 ± 55	63 ± 59	-10 (-22 to 2)	-67	47	0.87 (0.73 to 0.94) ***
Self-led individual training	173 ± 91	169 ± 92	5 (-13 to 22)	-77	86	0.90 (0.78 to 0.95) ***
Self-led group training	3 ± 10	5 ± 14	-2 (-5 to 1)	-15	11	0.91 (0.81 to 0.96) ***
Peer-led play	7 ± 26	2 ± 12	5 (-2 to 12)	-28	37	0.89 (0.77 to 0.95) ***
Section 2. Training environment						
Gym-based	131 ± 85	134 ± 90	-3 (-19 to 12)	-79	72	0.91 (0.80 to 0.96) ***
Pitch-based	87 ± 92	83 ± 84	4 (-15 to 23)	-86	95	0.87 (0.72 to 0.94) ***
Track-based	14 ± 27	10 ± 25	3 (-7 to 14)	-46	53	0.53 (0.17 to 0.77) **
Road-based	32 ± 46	35 ± 53	-3 (-12 to 5)	-44	38	0.92 (0.83 to 0.97) ***
Classroom-based	27 ± 47	52 ± 70	-25 (-45 to -5) *	-119	69	0.73 (0.47 to 0.87) ***
Section 3. Training objective						
Skill development session	16 ± 36	82 ± 103	-67 (-102 to -31) **	-237	104	0.59 (0.25 to 0.80) **
Fitness development session	163 ± 72	185 ± 109	-22 (-50 to 6)	-155	111	0.79 (0.58 to 0.91) ***
Injury prevention session	27 ± 28	28 ± 27	-1 (-11 to 9)	-47	45	0.64 (0.32 to 0.83) **
Match preparation session	21 ± 25	21 ± 30	-1 (-9 to 7)	-38	36	0.78 (0.57 to 0.90) ***
Recovery session	35 ± 33	37 ± 34	-2 (-8 to 5)	-32	28	0.90 (0.77 to 0.95) ***

Data are presented as mean (SD). RTAQ, Referee Training Activity Questionnaire. CI, confidence interval. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Table 3.5. Systematic bias, 95% limits of agreement (LOA), and Pearson correlations for training volumes of specific training practices (n=25).

	Assessment method		Bias (95% CI)	LOA -	LOA +	*r* (95% CI)
	RTAQ (min/week)	Training diary (min/week)				
Physical training activities						
High-intensity interval training	48 ± 50	52 ± 51	-4 (-17 to 9)	-67	59	0.80 (0.59 to 0.91) ***
Continuous aerobic training	36 ± 38	44 ± 43	-9 (-22 to 5)	-73	56	0.68 (0.39 to 0.85) ***
Speed endurance	18 ± 24	10 ± 17	8 (-1 to 16)	30	46	0.58 (0.24 to 0.80) **
Repeated sprint ability	15 ± 23	11 ± 18	4 (-3 to 12)	-31	39	0.65 (0.34 to 0.83) ***
Speed and agility	12 ± 21	8 ± 15	5 (-1 to 11)	-24	33	0.70 (0.43 to 0.86) ***
Resistance training	59 ± 60	68 ± 59	-10 (-22 to 3)	-68	48	0.88 (0.74 to 0.94) ***
Game-based activities	37 ± 89	0 ± 2	37 (1 to 78) *	-138	212	0.04 (-0.36 to 0.43)
Cross training	28 ± 49	38 ± 58	-9 (-22 to 3)	-70	52	0.84 (0.67 to 0.93) ***
Mobility training	16 ± 27	20 ± 28	-4 (-12 to 4)	-41	33	0.76 (0.52 to 0.89) ***
Decision-making training activities						
On-field refereeing simulations	7 ± 22	2 ± 12	5 (-3 to 13)	-32	42	0.51 (0.14 to 0.75) *
Video clip analyses	64 ± 100	49 ± 69	16 (-13 to 45)	-122	153	0.72 (0.45 to 0.87) ***
Interactive video-based decision-making	9 ± 27	8 ± 25	1 (-3 to 4)	-17	19	0.94 (0.87 to 0.98) ***
Other decision-making training	9 ± 21	4 ± 11	4 (-4 to 12)	-32	40	0.49 (0.11 to 0.74) ***
Psychological skills training						
Self-talk	10 ± 32	0 ± 0	10 (-4 to 23)	-54	73	n/a§
Emotional control	2 ± 6	0 ± 2	2 (-1 to 4)	-11	14	0.02 (0.38 to 0.41)
Automaticity	5 ± 24	0 ± 0	5 (-5 to 15)	-42	52	n/a§
Goal setting	7 ± 14	3 ± 10	4 (-1 to 9)	-17	26	0.63 (0.31 to 0.82) **
Imagery	5 ± 9	0 ± 2	4 (-1 to 8) *	-14	22	0.04 (0.36 to 0.43)
Activation	3 ± 7	0 ± 0	3 (-1 to 6)	-11	17	n/a§
Relaxation	7 ± 19	2 ± 9	5 (-1 to 10) *	-18	28	0.90 (0.79 to 0.96) ***
Negative thinking	2 ± 9	0 ± 0	2 (-1 to 6)	-16	20	n/a§
Distractibility	0 ± 2	0 ± 0	0 (-1 to 1)	-3	3	n/a§
Technical training						
Technical skills training	10 ± 27	5 ± 24	6 (-1 to 11)	-21	32	0.86 (0.71 to 0.94) ***
Tactical research	31 ± 40	19 ± 27	12 (-2 to 26)	-55	78	0.55 (0.19 to 0.73) **
Laws of the game study	20 ± 28	11 ± 25	9 (-1 to 19)	-35	54	0.64 (0.33 to 0.83) **
Game analysis	44 ± 50	47 ± 55	-3 (-22 to 17)	-96	90	0.59 (0.26 to 0.80) **

Data are presented as mean (SD). RTAQ, Referee Training Activity Questionnaire. CI, confidence interval. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$. § Unable to be calculated.

3.5. DISCUSSION

The present study aimed to develop and validate a novel measurement tool to capture information pertaining to the training practices of soccer officials. The resulting product is the RTAQ, a self-report questionnaire that presents good levels of content, face, and criterion validity, and may feasibly be implemented to assess the training practices engaged in by soccer FR and AR. This presents an attractive proposition to the applied practitioner as the longitudinal monitoring of training may permit the successful evaluation, and adjustment, of training (Saw et al., 2015).

Although susceptible to report bias, support for self-report measures is growing (Saw et al., 2017). However, to ensure the reliability of the results, it is imperative that an instrument first be assessed for measures of content, face, and criterion validity (Mokkink et al., 2010). In contrast to measures previously employed within the literature (MacMahon et al., 2007), items for the RTAQ were generated and revised following a systematic process to ensure their specificity and validity to soccer officials. Firstly, an item pool was generated following consultation with a sample of elite soccer officials and their support staff (MacNamara & Collins, 2011). An expert panel subsequently examined items both quantitatively and qualitatively during stage two, with excellent levels of content and face validity being evidenced (Polit & Beck, 2006). In stage three, estimates of training activity involvement reported by the RTAQ were compared against those obtained from a detailed training diary (Halson, 2014). For most training activities, the mean bias exhibited between methods was negligible, thus highlighting a good level of agreement. Additionally, large to nearly perfect correlations were evident between the RTAQ and training diaries for most items, with such values proving comparable to previous research detailing the associations between the self-reported and actual training duration of physically active participants (Borresen & Lambert, 2006). However, relatively wide LOA were present for various training activities such as high-

intensity interval training (-67 to 59 min) and video clip analyses (-122 to 153 min). Whilst such findings are in line with those of Borresen and Lambert (2006), there exist a number of factors that may help to explain the relatively wide LOA. Firstly, although the RTAQ was found to possess high levels of face and content validity, it is possible that some officials may have misinterpreted certain items on the RTAQ. In doing so, these individuals may have under- or over-reported their engagement in certain activities on the RTAQ when compared to the training diaries. Another plausible explanation for the wide LOA presented for certain activities relates to potential disparities in the time course in which the RTAQ and training diaries were completed. To account for potential seasonal variations, the maximal time interval between completion of the various methods was limited to one month. However, despite these considerations, it is important to acknowledge the variation that may present itself amongst weekly-microcycles (Malone et al., 2015). For instance, training during 1-, 2-, and 3-match weeks is likely to be markedly different, particularly with regards to the volume of high-intensity training performed (Anderson et al., 2016). Moreover, officials in the present study were part-time and combined their officiating careers with additional occupations. It therefore appears reasonable to suggest that extenuating factors will influence and drive an official's training availability from week-to-week. Nevertheless, as participants noted their responses to be representative of a typical training week, these suggestions remain speculative. Given the implications that such variation may have in the ability to effectively prescribe training, it may therefore be important to corroborate with objective data where possible (Borresen & Lambert, 2009).

Whilst previous instruments have largely focused on physical training activities, the participatory approach adopted in the current study enabled the identification of other activities unique to soccer officials, such as specific decision-making, psychological, technical and game-management skills training. Generally, soccer officials are reported to exhibit a

propensity for physical conditioning with a weak culture of skill practice (Weston et al., 2012). Moreover, the decision-making training of soccer officials has traditionally focused on the development of declarative knowledge (i.e., Laws of the Game) through lecture-style meetings, with practical experience being limited to competitive matches (MacMahon et al., 2007). While physical training may be monitored objectively with relative ease through the use of HR monitors and GPS devices, the monitoring of self-led training activities performed off field proves more challenging. Thus, practitioners and researchers working within the applied setting may wish to implement the RTAQ to assess the engagement of soccer officials in the wider spectrum of activities pertinent to performance. Such information may provide important insights into potential shortcomings in current practices, from which action may be taken to promote a greater diversification of training. Notwithstanding, it is also important to acknowledge that whilst the RTAQ assesses the total time engaged in various different training activities, it is primarily the intensity and purposefulness of training that govern the magnitude of the adaptive responses achieved (Hawley, 2008). As such, although the RTAQ represents a novel method of addressing important areas of training often overlooked, opportunity exists for its implementation alongside current monitoring strategies in order to capture reliable estimates of both training volume and intensity.

In the context of elite soccer, the monitoring of training has become common practice, with an abundance of research now available detailing the training practices of the players (Los Arcos et al., 2017; Weston, 2018). Conversely, very little information is readily discernible on the training of soccer officials, with such discrepancies potentially related to disparities in the resources (financial and personnel) available (Weston et al., 2012; Halson, 2014). Self-report measures such as training diaries are commonly employed within both research and applied practice as simple and inexpensive methods of acquiring insight into an athletes' training (Pugliese et al., 2014). Nonetheless, training diaries often prove problematic when dealing

with large cohorts due to the considerable volume of data that can be generated (Mujika, 2017). Furthermore, owing to their self-report nature, training diaries are reliant upon the level of detail that an official provides. In the present study, disparities were revealed in the extent to which participants reported their engagement in psychological skills training, with such information appearing not to be captured by the training diaries. Although we are unable to confirm which method provides the most accurate account, we may speculate that training diaries are likely less effectual in capturing an official's engagement in more casual types of training of higher frequencies and shorter durations. Conversely, whilst not discounting the potential for social desirability response bias, the provision of pre-defined options may facilitate a greater recall of casual training activities. Nevertheless, this remains speculative and warrants further investigation. It is also worth noting the higher completion rates observed for the RTAQ in comparison to the training diaries. Indeed, whilst all 31 participants completed the RTAQ during the assessments of criterion validity, six officials failed to complete the training diary. These findings provide additional evidence that training diaries may present issues of compliance (Mujika, 2017), and further support the RTAQ as a quick and practical alternative of capturing information pertaining to the training practices of soccer officials.

Another promising feature of the RTAQ concerns its specificity to both FR and AR. Indeed, given their differing roles on the field of play, perspectives of the attributes and training practices pertinent to performance will likely differ. Consequently, to ensure the applicability of the RTAQ to each type of official, both FR (n=23) and AR (n=15) were consulted at each stage of its development and validation. Encouragingly, the RTAQ presented high levels of content, face and criterion validity amongst both FR and AR, thus emphasising its specificity to these specialised roles. Although a relatively small number of officials were included at each stage, it is important to note that our sample size was restricted to the small numbers of

elite participants available – a challenge commonly encountered when conducting research within the applied field. Nonetheless, the sample sizes present remain consistent with those of previous literature (Polit et al., 2007; Heikkilä et al., 2017) and represent a significant progression from previous measures whereby items were amended to the role of the soccer official following consultation with two referee coaches (MacMahon et al., 2007). An additional consequence of our limited access to elite officials was that our cohorts were exclusively male. Although the attributes and training practices identified in the current study are likely applicable across sexes given the comparable demands imposed during match play (Mallo et al., 2010), further research is required to confirm the validity of the RTAQ amongst female officials. Finally, as our participants were recruited from categories 1–3, our ability to extrapolate these findings to lower-level officials remains limited. Nevertheless, a wide spectrum of officiating levels (regional to intercontinental) was represented in the current investigation. Future research may therefore wish to utilise the RTAQ to examine potential differences in the training practices engaged in by soccer officials of varying levels of experience and professional attainment.

3.5.1. Conclusions and Perspective

In conclusion, the present data have demonstrated the RTAQ to be a valid method of monitoring the multifaceted training practices of soccer officials, with good levels of content, face, and criterion validity being observed. Thus, given the previous absence of a standardised and accessible measure, the present work represents a valuable contribution to the literature and possesses important implications for future research and practice. Firstly, it is crucial that those responsible for the development of soccer officials monitor the training process closely, as in doing so, current practices may be evaluated and adjusted accordingly. Practitioners working with soccer officials within the field may therefore wish to consider the

implementation of the RTAQ as a means of monitoring engagement in the wide range of training activities pertinent to performance. The RTAQ may also be utilised for research purposes and to provide an extensive and up-to-date account of the training practices currently engaged in by soccer officials. As such a study has yet to be completed, the RTAQ was administered during Chapter 4 amongst a cohort of Scottish soccer officials with the view of documenting the types and volumes of training engaged in by soccer officials of varying roles and levels of professional attainment. It was anticipated that such information would prove useful in identifying areas for improvement and steering the direction of Chapters 5 and 6 (Bishop, 2008; Drust & Green, 2013).

Chapter 4

Training Practices and Perceptions of Soccer Officials: Insights from the Referee Training Activity Questionnaire (RTAQ)

Publications arising from this chapter:

McEwan, G. P., Unnithan, V. B, Easton, C., & Arthur, R. (2022). Training practices and perceptions of soccer officials: Insights from the Referee Training Activity Questionnaire. *International Journal of Sport Science & Coaching,* online ahead of print. https://doi.org/10.1177/17479541221110707

4.1. ABSTRACT

Purpose: To: 1) document the training practices of soccer officials in relation to their role and level of professional attainment; and 2) explore the association between the officials' perceptions and training practices. **Methods:** Field (FR) and assistant (AR) referees listed at categories 1-3 (n=173) with the Scottish Football Association were invited to participate in this national cross-sectional study. Using an online version of the Referee Training Activity Questionnaire (RTAQ), officials reported the volume and type of training engaged in during a 2-week in-season period. Respondents' perceptions of the skills pertinent to performance were also explored using a 7-point Likert scale *(1=not at all important; 7=extremely important)*, with the frequency in which they targeted the development of these skills during training assessed using a 5-point Likert Scale *(1=less than once a month; 5=more than once a week)*. **Results:** A total of 91 responses were received, representing a 52.6% response rate. Irrespective of their role or level of professional attainment, the officials' training was focused mostly on physical conditioning, with significantly less time apportioned to decision-making (P<0.001; ES=1.28), psychological (P<0.001; ES=1.47), and technical (P<0.001; ES=1.23) skills training. Meanwhile, although decision-making and psychological skills were rated as *"very important"* to *"extremely important"* amongst both FR and AR, training that targeted the development of these skills was engaged in *"less than once a month"*. **Conclusions:** The present study provides important new insights into the training practices engaged in by soccer officials. As accurate decision-making is at the forefront of soccer officiating, enhancing the exposure of match officials to match-related decision-making during training should be considered a priority for future research.

4.2. INTRODUCTION

To facilitate optimal performance, soccer officials require a multifaceted skill profile encompassing physical, decision-making, psychological, and technical attributes (Castagna et al., 2007; Slack et al., 2013). While the most appropriate methods of developing such skills remain open to debate, research has long argued the role of deliberate practice – that is, effortful and structured activity with the intention of enhancing performance (Ericsson et al., 1993). As contended by the theory of deliberate practice, between-group differences in skilled performance relate predominantly to differences in the volume of training accumulated over extended periods of time (Baker and Young, 2004). Accordingly, the long-term commitment of soccer officials to training and practice appears crucial if the skills and attributes necessary to officiate at the top level of the game are to be acquired.

The magnitude and direction of adaptations achieved in response to training is dependent upon a multitude of variables. Firstly, the training load imposed over a given time plays a key role in shaping the training response (Fitzpatrick et al., 2018). Careful adjustments to the volume and intensity of training are therefore necessary in order to maximise training-performance potential (Gabbett, 2016). Quantifying the volume and intensity of training performed by soccer officials therefore represent important aspects of effective training management (Hawley, 2008, Weston, 2018). The specificity of training, or the extent to which the stimulus reflects the demands of competition, also warrants consideration. In the context of soccer officials, the principle of specificity holds that their training programmes should reflect the multidimensional nature of the role (Weston et al., 2012). It would therefore appear important to monitor each aspect of a soccer official's training and preparation, with attention paid not only to the time engaged in physical training, but that apportioned to decision-making, psychological, and technical skills training also. Equipped with such information, coaches and

practitioners working with match officials may be able to better prescribe and diversify training, with the intention of optimising performance (Borresen & Lambert, 2009).

In contrast to the volume of literature pertaining to the players, the training practices of soccer officials has gone largely unexplored. Previously, Weston and colleagues (2011d) provided an account of the training engaged in by an elite English Premier League official over an 8-year period. Notwithstanding the value of such data in detailing the longitudinal training practices of an elite official, these data were limited to physical training sessions only, and did not consider the time apportioned to other activities pertinent to soccer officials. In a series of other studies, an amended version of the deliberate practice questionnaire was administered amongst elite Belgian and English officials to document the time engaged in various on-field and off-field training activities (MacMahon et al., 2007; Catteeuw et al., 2009). Respondents were found to exhibit a preference for physical conditioning with less time devoted to skills practice (MacMahon et al., 2007; Catteeuw et al., 2009). The applicability of the deliberate practice questionnaire for use with soccer officials has however been questioned (McEwan et al., 2020). In particular, as the deliberate practice questionnaire was initially developed and validated for use with soccer players (Helsen et al., 1998), a large emphasis remained on physical training activities, with relatively little detail provided surrounding the activities unique to soccer officials, such as specific decision-making, psychological, and technical skills training. It is therefore unclear whether previous observations simply reflect the limited scope of the measurement tools utilised (MacMahon et al., 2007; Catteeuw et al., 2009b). Additional research is therefore necessary to expand upon these findings and to document engagement in the wider spectrum of activities pertinent to performance. Additionally, previous research has largely focused on elite FR, with little data currently available for AR or those officials of lower levels of professional attainment. However, given the differing demands imposed on FR and AR, and the varying levels of training support provided to officials at different stages of their

refereeing careers, training practices may well differ amongst different types and levels of soccer official. The aim of the current study was to provide an updated account of the training practices engaged in by soccer officials in relation to their roles (FR or AR) and level of professional attainment. A secondary aim was to explore the association between the perceptions of soccer officials and the training practices they engaged in.

4.3. METHODS

4.3.1. Participants

Soccer officials listed at categories 1–3 (n=173) with the Scottish Football Association were invited to participate in this national cross-sectional questionnaire study. An overview of the career pathway for match officials within Scotland is presented within Figure 4.1. Invitations were circulated via email between September 2018 and December 2018, and no incentives to complete the questionnaire were offered. Informed consent was obtained electronically on the first page of the questionnaire and the study received institutional ethical approval (24-7-17-002).

4.3.2. Referee Training Activity Questionnaire (RTAQ)

The RTAQ developed and validated in Chapter 3 (McEwan et al., 2020) was created using an online survey tool (Online Surveys, Jisc, UK) and was distributed to participants in the form of a web link. The RTAQ is available within Appendix D.

In addition to examining the frequency with which officials engaged in training that focused on the development of particular skills, it was also deemed prudent to examine the perceptions of officials and the importance they attributed to these various skills and attributes (Harper & McCunn, 2017). In doing so, mismatches between the perceptions and practices of officials could potentially be highlighted. Accordingly, a supplementary section was completed

following the RTAQ whereby respondents were asked to indicate the level of importance they attributed to a range of physical (n=10), decision-making (n=9), psychological (n=5), technical (n=6), and game-management (n=4) skills. The skills listed in this section of the questionnaire were consistent with those identified during the development and validation of the RTAQ within Chapter 3 (McEwan et al., 2020), and responses were provided using a 7-point Likert scale (*1=not at all important; 7=extremely important*) (Wade, 2006). In total, the questionnaire took 16 ± 7 minutes to complete.

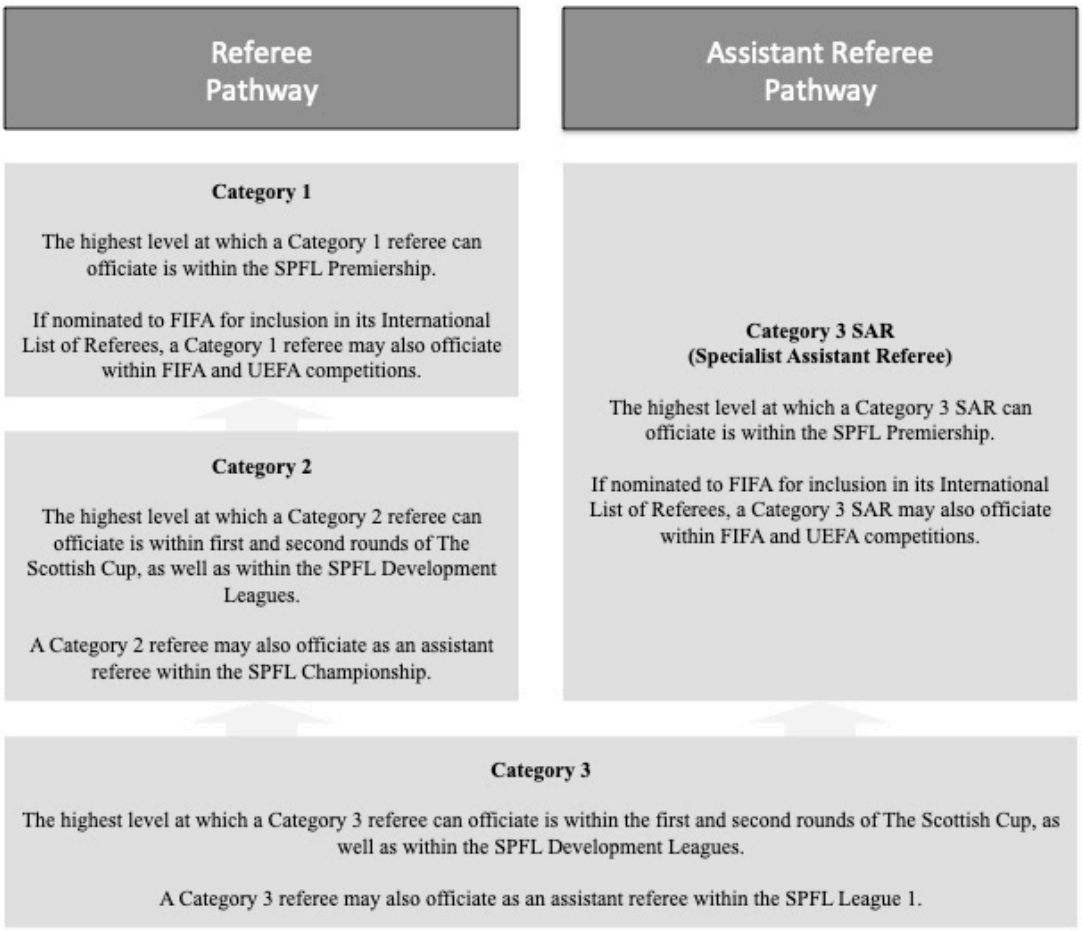


Figure 4.1. Overview of the career pathway for match officials within the Scottish Football Association.

4.3.3. Data Analysis

Demographic information derived from multiple choice questions were converted into a proportion of the total number of responses per category, with descriptive statistics reported as mean ± SD. Measures of training volume were ascertained by multiplying the frequency and duration of training. Prior to the use of parametric statistical test procedures, Shapiro-Wilk's tests were used to verify that the data were normally distributed. Between-category differences in the volume of training engaged in were analysed using factorial ANOVAs (two-way mixed), with referee category adopted as the between-participants factor and training activity as the within-participants factor. Any violations to sphericity were corrected using Huynh-Feldt procedures when the Greenhouse-Geisser value was greater than 0.75. When the Greenhouse-Geisser value was less than 0.75, any violations to sphericity were corrected using the Greenhouse-Geisser procedure (Field, 2009). Mean standardised differences are reported as Cohen's effect sizes (ES) with the following criteria used to interpret the practical significance of findings: trivial, <0.2; small, 0.21-0.6; moderate, 0.61-1.2; large, 1.21-1.99; and very large, >2.0 (Hopkins et al., 2009). Likert scale responses concerning the ratings of importance and frequency of training are presented in a descriptive manner as the mean response expressed as an integer (Hopkins, 2010). Statistical procedures were performed using Statistical Package for Social Sciences (SPSS 22.0, IBM, USA), with significance set at $P < 0.05$.

4.4. RESULTS

4.4.1. Respondents

A total of 91 responses were received, representing an overall response rate of 52.6%. Five responses were excluded from the final analyses due to incomplete answers. Subsequently, a total of 86 responses were included in the final analysis. An overview of respondents' demographic information is presented in Table 4.1.

Table 4.1. Demographic information of participants (n=86) who completed the Referee Training Activity Questionnaire (RTAQ).

	Cat 1 (n=25)	Cat 3SAR (n=28)	Cat 2 (n=15)	Cat 3 (n= 18)
Age (years)	34.7 (8.8)	37.9 (6.1)	30.7 (5.5)	28.1 (5.7)
Officiating experience (years)	15.3 (5.2)	14.5 (4.5)	8.3 (2.1)	7.1 (2.5)
Highest level officiated				
Intercontinental	1 (4.0)	4 (12.3)	0 (0.0)	0 (0.0)
Continental	10 (40.0)	12 (42.9)	0 (0.0)	0 (0.0)
National	14 (56.0)	12 (42.9)	10 (66.7)	12 (66.7)
Regional	0 (0.0)	0 (0.0)	5 (33.3)	6 (33.3)

Data presented as mean ± SD or number (%). SAR, specialist assistant referee.

4.4.2. General Training Information

4.4.2.1. Domain-Specific Activities

There was a significant main effect for activity ($F_{(2.095,171.828)} = 137.25$, $P<0.001$) with more time apportioned to self-led individual training than match officiating ($P=0.001$; 95% CI, 26–157 min, ES=0.46), coach-led individual training ($P<0.001$, 95% CI, 223–367 min, ES=1.34), coach-led group training ($P<0.001$, 95% CI, 134–276 min, ES=0.95), self-led group training ($P<0.001$, 95% CI, 298–378 min, ES=1.43), and peer-led play ($P<0.001$, 95% CI, 239–374 min, ES=1.48) (Table 4.2). More time was apportioned to match officiating than coach-led individual training ($P<0.001$; 95% CI, 166–239 min, ES=1.81), coach-led group training ($P<0.001$; 95% CI, 74–153 min, ES=0.93), self-led group training ($P<0.001$, 95% CI, 183–248 min, ES=2.17), and peer-led play ($P<0.001$; 95% CI, 182–248 min, ES=2.11). Time spent in coach-led group training was greater than time spent in coach-led individual training ($P<0.001$, 95% CI, 57–122 min, ES=0.90), self-led group training ($P<0.001$, 95% CI, 71–134 min, ES=1.06), and peer-led play ($P<0.001$, 95% CI, 72–131 min, ES=1.14). A significant activity x group interaction ($F_{(6.286,171.828)} = 3.178$, $P=0.005$) was observed. Post-hoc tests revealed Category 1 officials to take part in significantly more self-led individual training ($P=0.013$, 95%

CI, 30–368 min, ES=0.34), but less coach-led group training (P=0.003, 95% CI, 24–159 min, ES=0.40), than category 2 officials.

4.4.2.2. Training Environment

There was a significant main effect for activity ($F_{(2.676,219.438)}$ = 56.05, P<0.001) with more time apportioned to gym-based training than pool-based (P<0.001; 95% CI, 170 – 274 min; ES=1.39), pitch-based (P<0.001; 95% CI, 39–179 min; ES=0.51), track-based (P<0.001; 95% CI, 126–239 min; ES=1.05), road-based (P<0.001; 95% CI, 109–238 min; ES=0.88), and classroom-based (P<0.001; 95% CI, 127–226 min; ES=1.16) training (Table 4.2). Time spent in pitch-based training was greater than that apportioned to pool-based (P<0.001; 95% CI, 67–158 min; ES=0.81), track-based (P<0.001; 95% CI, 25–120 min; ES=0.50), road-based (P=0.002; 95% CI, 15–114 min; ES=0.43), and classroom-based training (P<0.001; 95% CI, 22–113 min; ES=0.49). Time apportioned to pool-based training was less than track-based (P<0.001; 95% CI, 16–63 min; ES=0.55), road-based (P<0.001; 95% CI, 26–70 min; ES=0.71), and classroom-based (P<0.001; 95% CI, 23–67 min; ES=0.68) training. The activity x category interaction was not significant ($F_{(8.028,219.438)}$ = 0.632, P=0.751).

4.4.2.3. Training Objective

There was a significant main effect for training objective ($F_{(1.685,138.200)}$ = 116.76, P<0.001) with significantly more training focused on fitness development than skill development (P<0.001; 95% CI, 174–284 min; ES=1.30), injury prevention (P<0.001; 95% CI, 162–266 min; ES=1.28), match preparation (P<0.001; 95% CI, 171–269 min; ES=1.40), and recovery (P<0.001; 95% CI, 158–263 min; ES=1.24) (Table 4.2). The activity x category interaction was significant ($F_{(5.056,138.200)}$ = 3.081, P=0.011) with category 3SAR officials engaging in more fitness development than category 3 officials (P=0.011; 95% CI, 25–297 min; ES=0.35).

Table 4.2. Total time (min) apportioned by soccer officials to general training practices across a 2-week in-season period.

	Cat 1 (n = 25)	Cat 3SAR (n = 28)	Cat 2 (n = 15)	Cat 3 (n = 18)	ALL (n = 86)
Domain-specific activities					
Match officiating	215 ± 97	209 ± 78	252 ± 85	240 ± 87	229 ± 87
Coach-led individual training	24 ± 60	27 ± 73	26 ± 54	27 ± 52	26 ± 60
Coach-led group training	68 ± 74	115 ± 66	159 ± 105¥	121 ± 68	116 ± 78
Self-led individual training	406 ± 211§	325 ± 172	207 ± 127	346 ± 226	321 ± 185
Self-led group training	24 ± 73	6 ± 24	3 ± 13	19 ± 33	13 ± 36
Peer-led play	10 ± 48	17 ± 36	5 ± 21	24 ± 35	14 ± 35
Training environment					
Gym-based	228 ± 182	254 ±144	226 ± 137	198 ± 121	227 ± 26
Pool-based	15 ± 40	3 ± 10	3 ± 12	0 ± 0	5 ± 17
Pitch-based	145 ± 164	114 ± 129	113 ± 108	99 ± 74	117 ± 38
Track-based	37 ± 73	56 ± 68	33 ± 60	53 ± 55	45 ± 8
Road-based	84 ± 69	55 ± 65	16 ± 45	58 ± 21	53 ± 23
Classroom-based	84 ± 75	63 ± 76	30 ± 41	71 ± 31	50 ± 23
Training objective					
Skill development	27 ± 63	39 ± 67	22 ± 27	23 ± 37	28 ± 8
Fitness development	268 ± 176	349 ± 218⌘	222 ± 79	188 ± 95	257 ± 70
Injury prevention	48 ± 52	49 ± 68	35 ± 46	40 ± 34	43 ± 7
Match preparation	55 ± 49	41 ± 94	30 ± 31	23 ± 47	37 ± 14
Recovery	73 ± 59	42 ± 46	33 ± 31	38 ± 40	46 ± 18

Data are presented as mean ± SD. SAR, Specialist Assistant Referee.

¥ Significant difference from Cat 1 (P<0.01). § Significant difference from Cat 2 (P<0.05). ⌘ Significant difference from Cat 3 (P<0.05).

4.4.3. Specific training practices

There was a significant main effect for activity ($F_{(1.537,126.053)}$ = 132.048, P<0.001) with officials engaging in significantly more physical training than decision-making (P<0.001; 95% CI 251–399 min; ES=1.28), psychological skills (P<0.001; 95% CI, 290–434 min; ES=1.47), and technical skills training (P<0.001; 95% CI, 213–346 min; ES=1.23) (Figure 4.2). Participants reported greater engagement in technical skills training than decision-making (P=0.002; 95% CI, 13–78 min; ES=0.41) and psychological skills training (P<0.001; 95% CI, 49–116 min; ES=0.75), whilst time spent in decision-making training was greater than that apportioned to the development of psychological skills (P=0.004; 95% CI, 9–65 min; ES=0.39). The activity x category interaction was not significant ($F_{(4.612,126.053)}$ = 1.125; P=0.350).

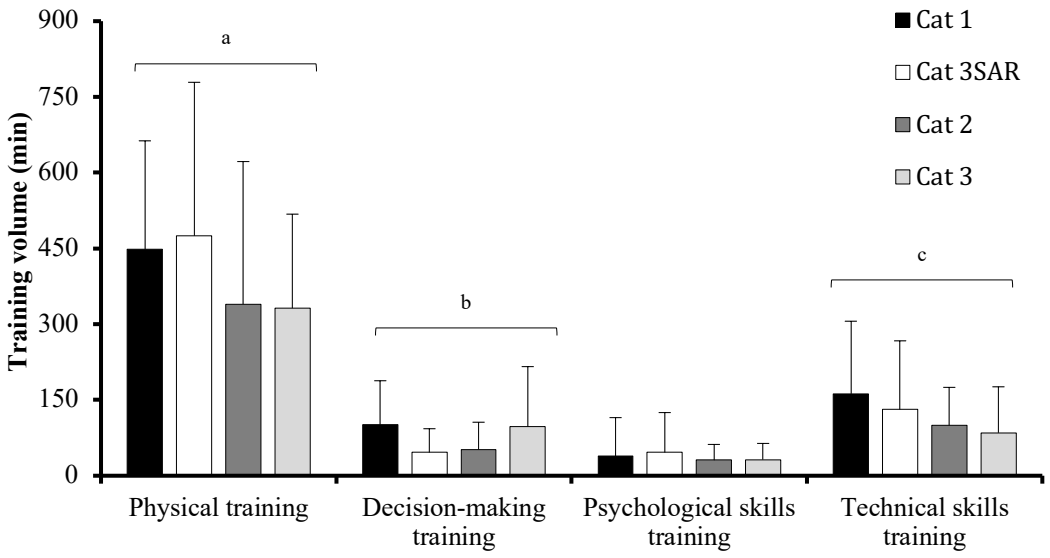


Figure 4.2. Total time apportioned by soccer officials to physical, decision-making, psychological, and technical skills training across a 2-week in-season period (mean ± SD). [a] Significant difference from decision-making, psychological, and technical skills training (all P<0.001). [b] Significant difference from psychological skills training (P<0.01) [c] Significant difference from decision-making (P<0.01) and psychological (P<0.001) skills training.

4.4.3.1. Physical Training

There was a significant main effect for activity ($F_{(4.658,381.952)} = 18.049$, $P<0.001$) with more time apportioned to high-intensity interval training, continuous aerobic training, and resistance training than speed endurance ($P\leq0.028$; ES≥0.38), repeated sprint ($P=0.000$; ES≥0.51), speed and agility ($P<0.001$; ES≥0.55), game activities ($P\leq0.016$; ES≥0.40), cross-training ($P\leq0.029$; ES≥0.38), and mobility training ($P<0.001$; ES≥0.53) (Table 4.3). Post-hoc tests also revealed more time to be apportioned to speed endurance than repeated sprint ($P=0.026$; 95% CI, 1–26 min; ES=0.38) and speed and agility ($P=0.004$; 95% CI, 3–33 min; ES=0.44) training. The activity x category interaction was not significant ($F_{(13.974,381.952)} = 0.767$, $P=0.705$).

4.4.3.2. Decision-Making Training

There was a significant main effect for activity ($F_{(1.477,121.092)} = 50.28$, $P<0.001$) with post-hoc tests revealing more time to be apportioned to video-clip analyses than on-field simulations ($P<0.001$; 95% CI, 39–79 min; ES=0.85), interactive decision-making ($P<0.001$; 95% CI, 36–78 min; ES=0.78), and other types of decision-making training ($P<0.001$; 95% CI, 40–82 min; ES=0.86) (Table 4.3). A significant activity x category interaction ($F_{(4.430,121.092)} = 2.625$, $P=0.033$) was observed with category 3SAR officials spending more time in interactive decision-making than category 1 officials ($P=0.049$; 95% CI, 0–48 min; ES=0.29).

4.4.3.3. Psychological Skills Training

There was a significant main effect for activity ($F_{(4.566,374.416)} = 4.85$, $P<0.001$) with more time apportioned to goal setting than emotional control ($P=0.011$; 95% CI, 1–15 min; ES=0.41), negative thinking ($P=0.019$; 95% CI, 1–17 min; ES=0.39), and distractibility training ($P=0.023$; 95% CI, 1-17 min; ES=0.38) (Table 4.3). More time was apportioned to imagery than negative thinking training ($P=0.016$; 95% CI, 1–14 min; ES=0.40) and distractibility training ($P=0.017$;

95% CI, 1–14 min; ES=0.39). The activity x category interaction ($F_{(13.698,374.416)}$ = 0.628, *P*=0.838) was not significant.

4.4.3.4. Technical Skills Training

There was a significant main effect for activity ($F_{(2.418,198.285)}$ = 15.916, *P*<0.001) with officials apportioning more time to game analysis than technical skills training (*P*<0.001; 95% CI, 23–63 min; ES=0.63) or tactical research (*P*=0.003; 95% CI, 7–43 min; ES=0.40) (Table 4.3). Significantly less time was apportioned to technical training than tactical research (*P*=0.003; 95% CI, 4–31 min; ES=0.39) or laws of the game study (*P*<0.001; 95% CI, 12–34 min; ES=0.60). The activity x category interaction ($F_{(7.254,198.285)}$ = 1.694, *P*=0.109) was not significant.

4.4.4. Perceived Importance and Frequency of Training of Attributes Deemed Relevant to Soccer Officiating

4.4.4.1. Physical Attributes

Physical attributes were rated as *moderately important* to *very important* amongst FR and were trained *once a week* to *once every two weeks* (Figure 4.3). Physical attributes were rated as *moderately important* to *extremely important* amongst AR and were trained *more than once a week* to *once every two weeks*.

Table 4.3. Total time apportioned by soccer officials to specific training practices across a 2-week in-season period.

	Cat 1 (n = 25)	Cat 3SAR (n = 28)	Cat 2 (n = 15)	Cat 3 (n = 18)	ALL (n = 86)
Physical training					
High-intensity interval training	118 ± 69	92 ± 68	86 ± 83	81 ± 59	94 ± 70
Continuous aerobic training	75 ± 96	100 ± 108	37 ± 62	68 ± 78	70 ± 86
Speed endurance training	38 ± 37	34 ± 31	29 ± 34	32 ± 32	33 ± 34
Repeated sprint training	28 ± 31	29 ± 36	12 ± 20	11 ± 15	20 ± 25
Speed and agility training	13 ± 35	26 ± 53	14 ± 21	8 ± 15	15 ± 31
Resistance training	107 ± 134	106 ± 133	71 ± 78	79 ± 85	91 ± 107
Game-based activities	13 ± 51	38 ± 68	33 ± 93	36 ± 74	30 ± 71
Cross-training	25 ± 62	42 ± 112	34 ± 71	10 ± 42	28 ± 72
Mobility training	31 ± 64	9 ± 25	25 ± 44	6 ± 13	18 ± 37
Decision-making training					
On-field refereeing simulations	13 ± 19	6 ± 17	3 ± 12	0 ± 0	5 ± 12
Video clip analyses	95 ± 88	63 ± 72	41 ± 43	52 ± 54	63 ± 64
Interactive video-based decision-making	0 ± 0	24 ± 56$^{\Psi}$	0 ± 0	0 ± 0	6 ± 14
Other	0 ± 0	4 ± 16	2 ± 8	0 ± 0	2 ± 6
Psychological skills training					
Self-talk	8 ± 29	5 ± 13	4 ± 10	6 ± 13	6 ± 16
Emotional control	0 ± 0	3 ± 10	0 ± 0	0 ± 0	1 ± 2
Automaticity	0 ± 0	5 ± 23	0 ± 0	0 ± 0	1 ± 6
Goal setting	6 ± 14	13 ± 31	5 ± 7	13 ± 22	9 ± 19
Imagery	5 ± 10	7 ± 15	13 ± 36	5 ± 9	8 ± 17
Activation	6 ± 24	5 ± 19	2 ± 6	0 ± 0	3 ± 12
Relaxation	14 ± 39	7 ± 24	7 ± 14	9 ± 18	9 ± 24
Negative thinking	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Distractibility	0 ± 0	1 ± 3	0 ± 0	0 ± 0	0 ± 0
Technical skills training					
Technical skills training	13 ± 42	10 ± 26	9 ± 21	3 ± 8	9 ± 24
Tactical research	43 ± 58	31 ± 36	26 ± 28	6 ± 12	27 ± 33
Laws of the Game study	27 ± 49	39 ± 51	34 ± 32	28 ± 24	32 ± 39
Game analysis	79 ± 74	51 ± 63	30 ± 26	47 ± 77	52 ± 60

Data are presented as mean ± SD. SAR, Specialist Assistant Referee.
Ψ Significant difference from Cat 1 ($P<0.05$).

4.4.4.2. Decision-Making Attributes

Decision-making attributes were rated as *very important* to *extremely important* amongst both FR and AR (Figure 4.4). However, FR reported training decision-making attributes *less than once a month*, whilst AR trained *less than once a month* to *more than once a month*.

4.4.4.3. Psychological Attributes

Although psychological attributes were rated as *very important* to *extremely important* amongst FR, respondents reported that they engage in psychological skills training *once a month* (Figure 4.5). Similar observations were reported amongst AR whereby psychological attributes were rated as *very important* to *extremely important* but trained at a frequency of *less than once a month* to *once a month*.

4.4.4.4. Technical Attributes

With the exception of flag technique whereby FR rated it as *moderately important* and trained this skill *less than once a month*, FR rated technical attributes as *very important* to *extremely important* and trained these *once every two weeks* (Figure 4.6). Technical attributes were rated as *moderately important* to *extremely important* amongst AR and were trained *once a month* to *once every two weeks*.

4.4.4.5. Game-Management Attributes

Game-management attributes were rated as *extremely important* amongst FR but were trained *less than once a month* (Figure 4.7). Similar findings were observed amongst AR, whereby game-management attributes were deemed *very important* to *extremely important* but trained *less than once a month* to *once a month*.

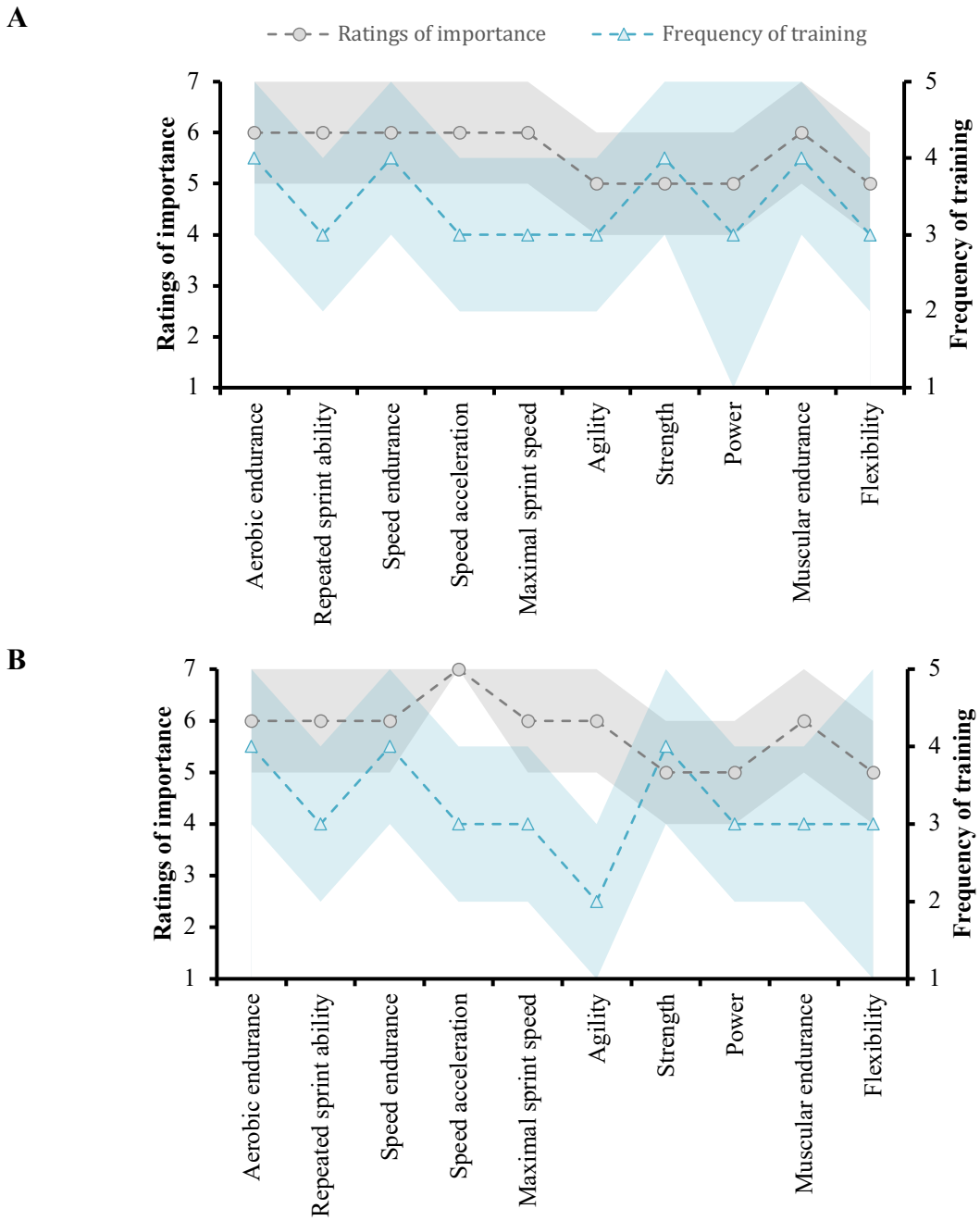


Figure 4.3. Perceptions and practices of field (A; n=58) and assistant (B; n=26) soccer officials in relation to physical attributes. Perceptions were assessed using a 7-point Likert scale (1=not at all important; 7=extremely important). Practices were assessed using a 5-point Likert scale (1=less than once a month; 5=more than once a week).

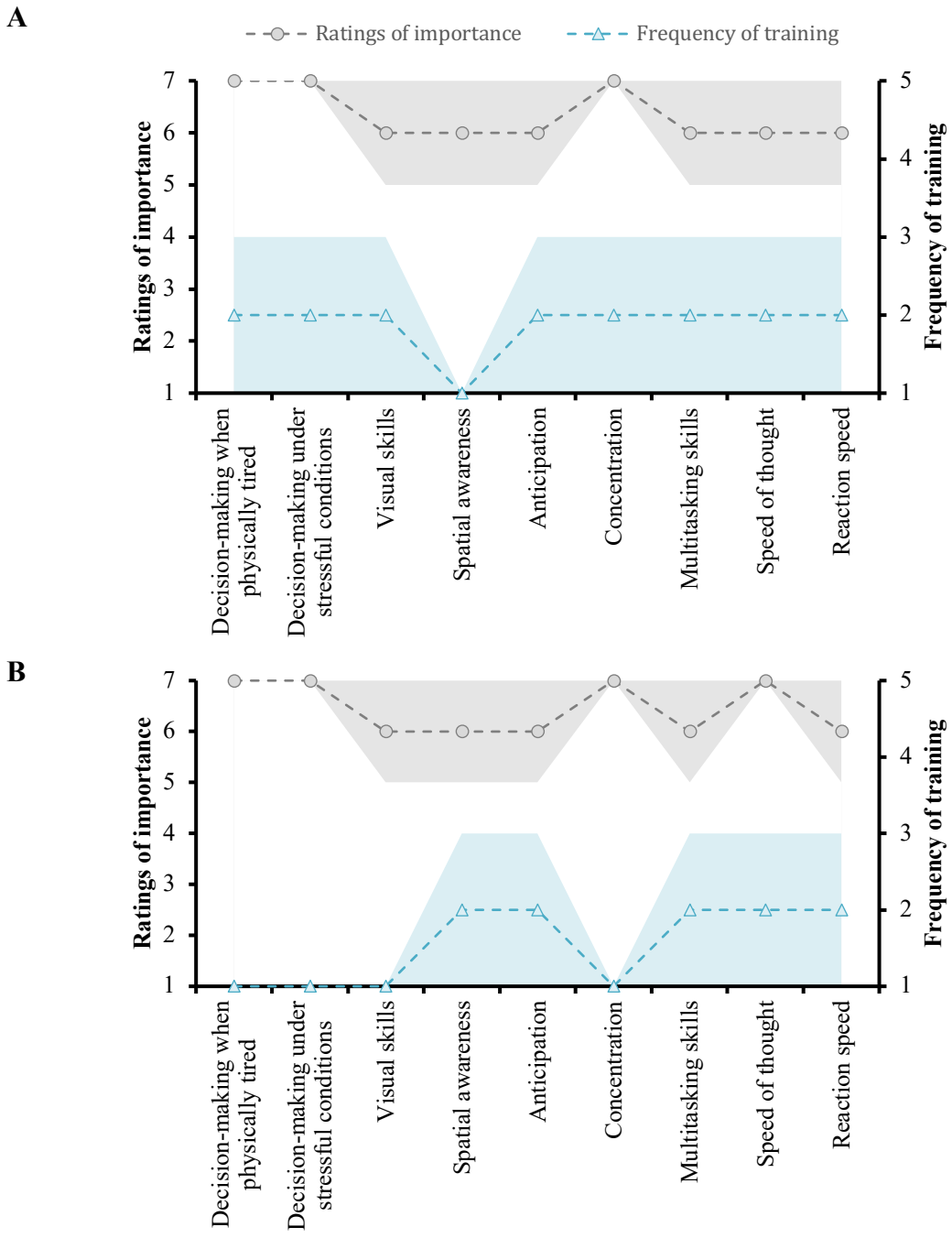


Figure 4.4. Perceptions and practices of field (A; n=58) and assistant (B; n=28) soccer officials in relation to decision-making attributes. Perceptions were assessed using a 7-point Likert scale (1=not at all important; 7=extremely important). Practices were assessed using a 5-point Likert scale (1=less than once a month; 5=more than once a week).

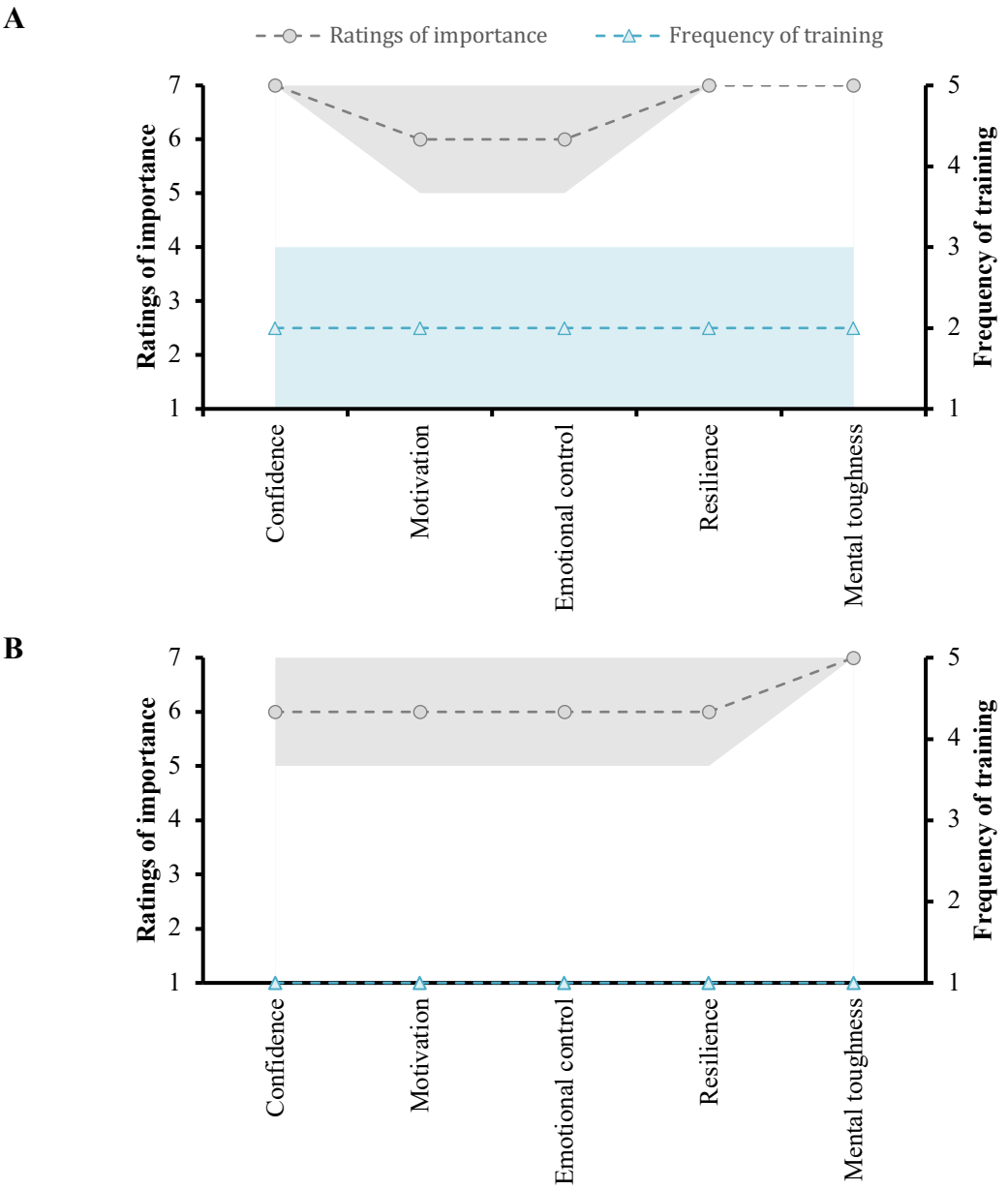


Figure 4.5. Perceptions and practices of field (A; n=58) and assistant (B; n=28) soccer officials in relation to psychological attributes. Perceptions were assessed using a 7-point Likert scale (1=not at all important; 7=extremely important). Practices were assessed using a 5-point Likert scale (1=less than once a month; 5=more than once a week).

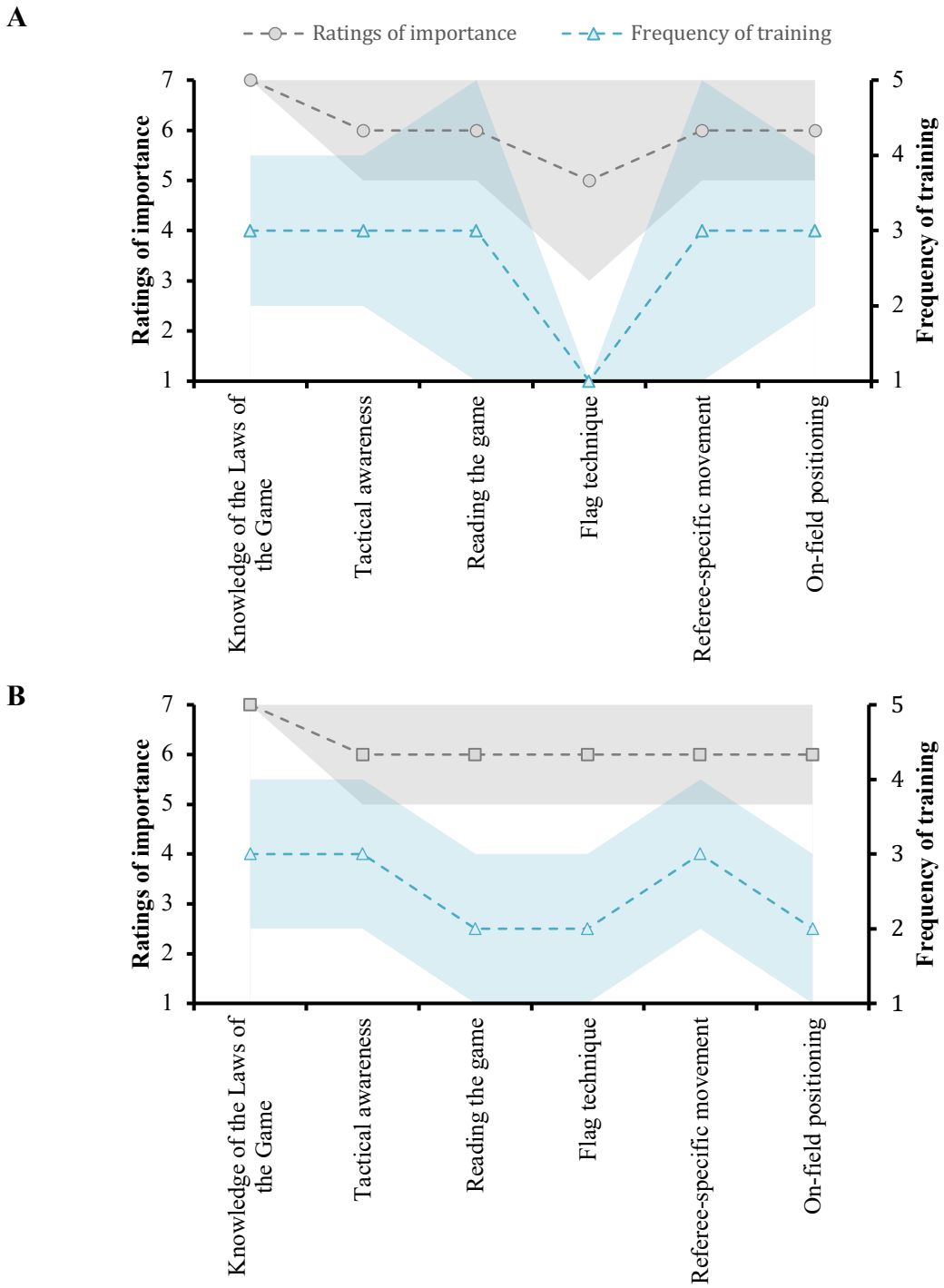


Figure 4.6. Perceptions and practices of field (A; n=58) and assistant (B; n=28) soccer officials in relation to technical attributes. Perceptions were assessed using a 7-point Likert scale (1=not at all important; 7=extremely important). Practices were assessed using a 5-point Likert scale (1=less than once a month; 5=more than once a week).

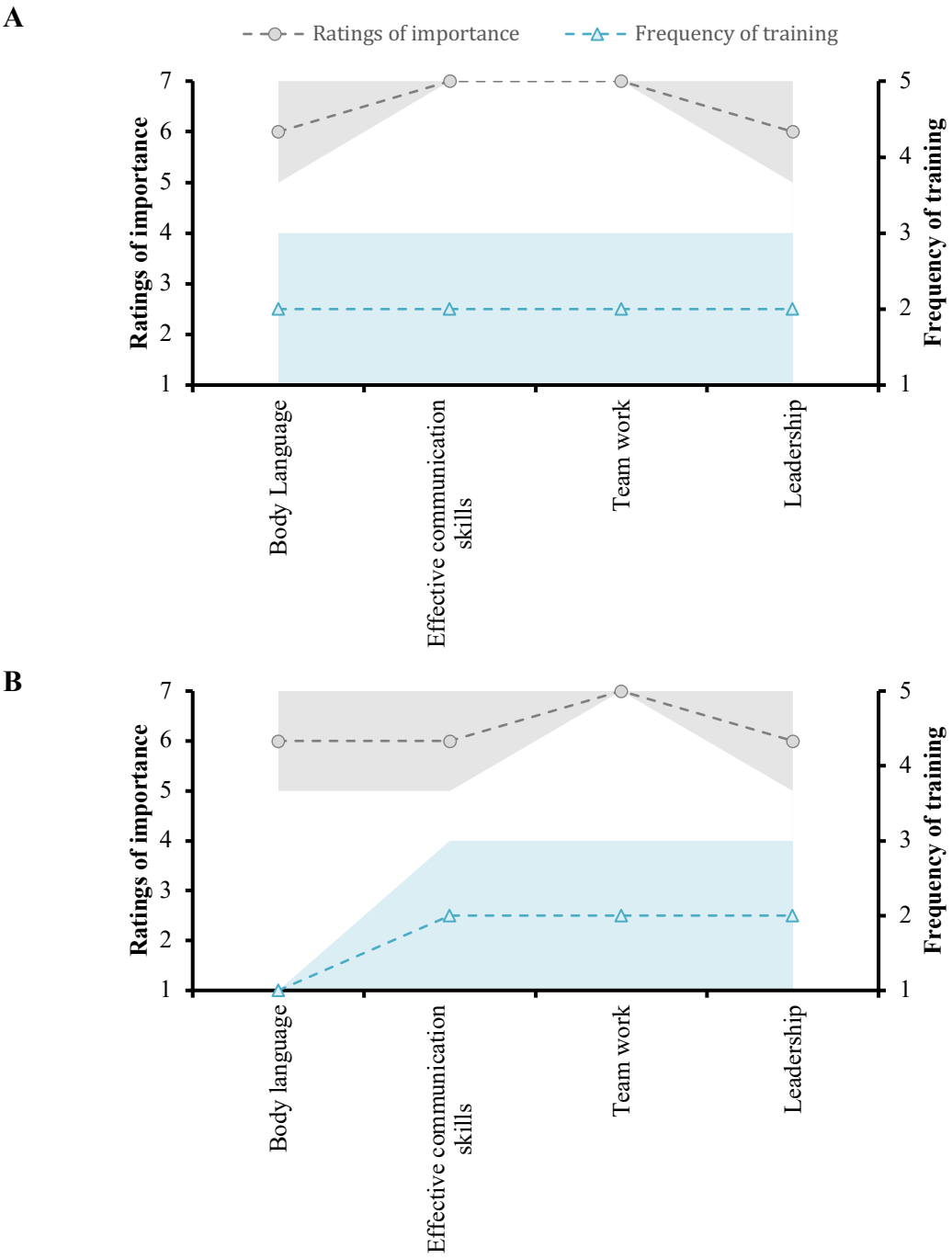


Figure 4.7. Perceptions and practices of field (A; n=58) and assistant (B; n=18) soccer officials in relation to game-management attributes. Perceptions were assessed using a 7-point Likert scale (1=not at all important; 7=extremely important). Practices were assessed using a 5-point Likert scale (1=less than once a month; 5=more than once a week).

4.5. DISCUSSION

This national cross-sectional study sought to examine and document the training practices of soccer officials. Our findings broadly highlight that soccer officials, irrespective of their role or level of professional attainment, largely focus on physical conditioning with relatively little attention directed towards other facets of performance, such as decision-making, psychological, and technical skills training. We also sought to explore relationships between respondents' perceptions of the skills and attributes pertinent to optimal performance, and how frequently they engaged in training that targeted the development of these skills. Whilst decision-making, psychological, and game management skills were rated as *"very important"* to *"extremely important"* amongst both FR and AR, training that targeted the development of these skills was engaged in less than once a month. In considering the multidimensional nature of soccer officiating, such findings are likely to have important implications for both the development of soccer officials and their preparedness to meet the demands of match play.

Consistent with the observations of previous literature (MacMahon et al., 2007; Catteeuw et al., 2009b), relatively low volumes of decision-making training were recorded in the present investigation, with significantly more time apportioned to the development of physical fitness. Considering the substantial physical and physiological loads imposed during match play (Krustrup et al., 2009; Castillo et al., 2017), the propensity of officials to engage in high volumes of physical training is perhaps unsurprising. Indeed, as soccer has become a progressively faster and more dynamic game, growing attention has been directed towards the physical conditioning of soccer officials, with routine fitness testing becoming an integral component of the match selection criteria adopted by national and international referee governing bodies (Weston et al., 2009; Castagna et al., 2012). Additionally, as match officials generally reach elite status at an age where physical capacities begin to decline, the emphasis placed on physical training may reflect their necessity to ensure that levels of physical fitness,

and therefore their ability to meet match demands, are maintained (Castagna et 2005; Weston et al., 2010).

Notwithstanding the importance of physical fitness, well-developed decision-making skills are central to the successful performances of soccer officials (Spitz et al., 2018b). Interestingly, although decision-making attributes were perceived as *"very important"* to *"extremely important"* amongst both FR and AR, respondents targeted the development of these skills *"less than once a month"*. Although speculative, such findings may reflect the current absence of naturalistic training methodologies that provide officials the opportunity to develop their decision-making abilities under match-like conditions (Weston et al., 2012). In the context of soccer players, small-sided games are often employed as a means of enhancing decision-making skills in unison with the tactical, technical, and physical aspects of the game (Light et al., 2014). In contrast, the present findings demonstrate that the decision-making and technical skills training of soccer officials are largely limited to video clip and game analyses. Whilst video-based training may result in improved performance during video-based decision-making tasks (Schweizer et al., 2011; Put et al., 2016), the fidelity and ecological validity of existing methods remains unclear. Typically, video-based training is performed within controlled off-field environments during which officials remain in a rested state. However, such conditions do not reflect the realities of competitive matches whereby decisions are undertaken in combination with elevated levels of physiological and psychological stress. It has therefore been suggested that to prepare officials for the multidimensional nature of match play, such conditions should be reproduced within training (Weston et al., 2012). Whether such approaches would improve the accuracy of the in-match decisions made by soccer officials remains to be seen however and represents an important avenue for future research.

To create additional opportunities for soccer officials to make decisions under elevated levels of physiological stress, referee governing bodies occasionally programme on-field

simulations during training whereby players are recruited to recreate game-like scenarios (Spitz et al., 2016; Helsen et al., 2019). Although the value of specific match practice is clear, such methods necessitate substantial amounts of organisation and compliance from clubs. With this in mind, it is perhaps unsurprising that officials in the present study apportioned relatively little time to on-field simulations, with *in-situ* decision-making experience being limited to official matches. Such findings may have important implications for the long-term development of soccer officials. The primary concern being that by limiting their decision-making experience to the finite number of match appointments available per season, it will take officials significantly longer to accrue the hours of deliberate practice required to attain decision-making expertise (Catteeuw et al., 2009b). In addition, competitive match play represents a practice-poor environment as the opportunity to receive feedback and engage in reflective practice is delayed. Thus, a key challenge facing those responsible for the development of soccer officials is to create novel training methodologies that facilitate the more rapid acquisition of decision-making expertise. A potential avenue in which this may be achieved is through the continued development and refinement of video-based simulation protocols that provide soccer officials with the opportunity to engage in decision-making under match-like conditions.

Officials engaged in significantly more gym-based than pitch-based training, with a large emphasis placed upon resistance training. These findings are consistent with those of Weston and colleagues (2011d) who reported gym-based strength training to predominate the training schedule of an elite English Premier League FR. In keeping with the large aerobic demands imposed during match play, officials also engaged in large volumes of high-intensity interval training and continuous aerobic training. Conversely, relatively little time was apportioned to on-field speed and agility training, with these observations remaining consistent across refereeing categories. As officials occupy an age bracket of 10-15 years greater than that of their playing counterparts (Weston et al., 2010) and given the progressive declines that occur

in speed and power performance with increasing age (Castagna et al., 2005), this latter finding may have particular implications for older officials. Firstly, as isolated and repeated bouts of high-speed running frequently precede crucial moments during match play (Faude et al., 2010), the ability of officials to keep up with play and position themselves appropriately is vital. Additionally, as part of the match selection criteria introduced by FIFA, officials must attain minimum standards during a repeated sprint test comprising six 40-m sprints interspersed with 90-s recovery (Weston et al., 2009). As the fastest and mean sprint times achieved during this test are related to an official's match-related running capacity, repeated sprint ability has been identified as an important discriminator amongst soccer officials (Weston et al., 2009). The time allocated to repeated-sprint training, however, was relatively low. To ensure that soccer officials are able to meet the demands of match play and fitness testing, greater attention should perhaps be paid to the development of speed and repeated sprint ability. In particular, repeated sprint training has obvious appeal for soccer officials, as such training represents a time-efficient method of enhancing several components of match-related fitness, such as speed, power, and high-intensity running performance (Taylor et al., 2015).

To elicit an optimal training stimulus and enhance their readiness for competition, the training of soccer players is typically periodised in relation to match day (Malone et al., 2015). In practical terms, this involves the deliberate manipulation of loading patterns to incorporate periods of high-intensity training and low-intensity recovery (Kelly et al., 2020). As the match activities and competitive schedules of soccer officials mirror those of the players (Weston et al., 2011c), it may be expected that similar strategies are adopted amongst officials. However, relatively little training was apportioned to low intensity activities such as injury prevention, match preparation, and recovery activities. Such findings are likely to have important implications for soccer officials as mismatches between loading and recovery processes, particularly during periods of fixture congestion, may impair physical capacities and

predispose officials to a heightened risk of injury and illness (Gabbett, 2016). Nonetheless, it is important to acknowledge that the data presented relates to self-reported training volumes only, and no measures of training intensity were obtained. Although previous studies have shed light on the training loads of soccer players (Anderson et al., 2016; Stevens et al., 2017), the training loads exhibited by soccer officials remains largely unexplored. Future research may therefore wish to examine the loading patterns and intensity distribution of soccer officials across a competitive season.

Although contingent upon the appraisal mechanisms employed by the individual, excessive increases in stress and anxiety are associated with the impaired decision-making performances of soccer officials (Balmer et al., 2007). Thus, given the challenging psychosocial environments in which they operate, calls have been made for psychological skills training to become an integral component of the training and preparation of match officials (Mathers & Brodie, 2011; Samuel, 2015). Nonetheless, our findings suggest that soccer officials, irrespective of their officiating category, currently engage in little training focused on the development of psychological skills. Previous investigations have found the successful application of sport psychology within applied sport environments to be impeded and made more challenging by a number of internal and external factors (Johnson et al., 2011; McDougall et al., 2015). With respect to elite soccer, for example, negative perceptions of sport psychology have previously been reported whereby an insufficient understanding of sport psychology and scepticism of the field were cited as significant barriers (Pain & Harwood, 2004; Johnson et al., 2010). Nevertheless, officials in the current study acknowledged that psychological skills were *"very important"* to *"extremely important"* to successful performance. Alternative barriers that may preclude the implementation of psychological support programmes within applied performance environments include a lack of funding and time (Pain & Harwood, 2004; Brink et al., 2018). Indeed, as sport psychology support

programmes have traditionally been delivered on an *ad hoc* and consultancy basis (McDougall et al., 2015), the implementation of such programmes may prove challenging within the context of smaller football associations with limited resources.

A response rate of 52.6% was achieved in the present investigation; a rate that is higher or comparable to previous surveys exploring the perceptions and practices of those working within elite soccer (Towlson et al., 2013; Akenhead & Nassis, 2016; Harper et al., 2016a; McCunn et al., 2018). As is often the case with descriptive studies of a cross-sectional nature, it does however remain a possibility that the current findings may have been subject to self-selection or non-response bias. Nonetheless, in considering the diversity of the current cohort, we believe that the responses received provide a fair and accurate reflection of current training practices within Scottish match officials. For instance, whereas previous studies have focused exclusively on small samples of elite FR (MacMahon et al., 2007; Catteeuw et al., 2009b), we focused on both FR (n=58) and AR (n=28) of different levels of professional attainment. This notwithstanding, readers should remain cognisant that the findings of the present study reflect the practices and perceptions of soccer officials from a single football association. The ability to generalise these findings to other contexts is therefore limited, as different training practices and perceptions may well be present within different sporting cultures and football associations. Examining the training practices and perceptions of officials from different countries and football associations therefore represents a potential avenue for future research.

4.5.1. Conclusions and Perspective

The present study has provided a comprehensive and up-to-date account of the current training practices of Scottish soccer officials and has helped highlight the specific areas of their training that warrant improvement. As per previous suggestions, officials were found to engage in large volumes of physical training, with considerably less attention directed towards

preparing for the other demands encountered during match play. In particular, respondents were found to engage in low volumes of decision-making skills training, with their exposure to match-related decision-making being largely limited to competitive matches or video clip analyses. As accurate decision-making is at the forefront of the officials' performance, such findings are likely to have important implications for their development and ability to fulfil their duties on the field on play. Enhancing the exposure of soccer officials to match-related decision-making during training should therefore be considered a priority for future research and practice.

Chapter 5

Validity and Reliability of Physiological and Perceptual Responses During a Treadmill-Based Soccer Referee Simulation (SRS)

Conference proceedings arising from this chapter:

McEwan, G. P., Unnithan, V., Easton, C., Glover, A. J., & Arthur, R. (2021). Validity and reliability of physiological and perceptual responses during a treadmill-based Soccer Referee Simulation (SRS). *26th annual Congress of the European College of Sport Science*, 8th – 10th September 2021.

5.1. ABSTRACT

Purpose: To explore the validity and reliability of the physiological and perceptual responses elicited during a novel treadmill-based Soccer Referee Simulation (SRS). **Methods:** Following the collection of baseline measures and habituation procedures, eight sub-elite soccer referees completed a single trial of the SRS whereby measures of heart rate (HR), oxygen uptake ($\dot{V}O_2$), blood lactate concentrations, and differential ratings of perceived exertion (RPE) were obtained. Referees' HR responses were also monitored during a series of competitive matches (5 match observations per referee). For the reliability aspect of the investigation, eight well-trained males were initially habituated, and thereafter completed three separate trials of the SRS during which the reliability of the selected outcome variables were ascertained. Trials were separated by 3-7 days and performed under standardised conditions. **Results:** No differences were evidenced between the SRS and match play in relation to measures of mean HR (P=0.444; ES=0.29), peak HR (P=0.074; ES=0.74), or HR-based training impulse (P=0.498; ES=0.25). Additionally, no systematic differences were detected between reliability trials for any of the measured outcome variables (P≥0.293), whilst good levels of reliability were observed for measures of mean HR (ICC=0.94; CV=3.1%), peak HR (ICC=0.93; CV=2.2%), HR-based training impulse (ICC=0.95; CV=10.0%), mean $\dot{V}O_2$ (ICC=0.95; CV=2.6%); blood lactate concentrations (ICC≥0.89; CV≤11.5%), and differential RPE (ICC≥0.94; CV≤15.1%). **Conclusions:** The SRS represents a valid and reliable protocol that closely replicates the physiological and decision-making demands of soccer refereeing.

5.2. INTRODUCTION

Soccer officiating represents a highly demanding activity, requiring officials to perform complex decision-making processes under challenging physiological conditions (Weston et al., 2012). Growing interest therefore exists with respect to the physical and physiological demands imposed during match play (Castagna et al., 2007), and their association with the officials' decision-making performances (Mascarenhas et al., 2009). However, as large levels of inter-match variability are evident in the volumes of high-speed running performed by match officials, the detection of real systematic changes in performance outcomes between matches is difficult (Weston et al., 2011b). Additionally, whilst the continuous monitoring of HR within applied settings is relatively simple, assessing an official's respiratory, metabolic, and perceptual responses during match play is impractical as collection procedures may interfere with normal refereeing duties (Rollo et al., 2014). Drawing meaningful conclusions on the *in-situ* impact of match play on the physical and decision-making performances of soccer officials therefore remains a challenge.

Simulation protocols that mimic match play represent alternative approaches that have been used to standardise the internal and external demands imposed on players (Russell et al., 2011). In addition to negating the contextual factors that confound match play data, simulation protocols also enable researchers to assess physiological and perceptual responses in a more controlled environment. To date, several soccer player simulations have been developed utilising treadmill-based (Aldous et al., 2014; Page et al., 2015) and free-running (Small et al., 2010; Russell et al., 2011) protocols. Whilst the validity and reproducibility of such protocols are well established, previous simulations are based upon the movement patterns elicited by the players during match play. Although the activity profiles of soccer FR are interrelated with those of the players, distinct differences exist (Weston et al, 2007; Weston et al., 2011c). Notably, FR generally cover greater distances comprising more low intensity running but less

sprinting, than their playing counterparts (Weston et al., 2011c). As the physical and decision-making demands of soccer officiating are imposed concurrently, rationale also exists for the inclusion of match-specific decision-making into the physical training and testing protocols of sport officials (Kittle et al., 2018). Indeed, findings in Chapter 4 suggested that soccer officials consider decision-making training to be of upmost importance; however, have little access to naturalistic decision-making training methods and protocols. Valid and reliable protocols that closely replicate both the physical and decision-making elements of soccer officiating are therefore warranted.

In a recent investigation, Samuel et al (2019) assessed the efficacy of a soccer official simulation whereby the physical and decision-making demands were imposed concurrently. Specifically, FR ran at varying paces on a motorised treadmill whilst simultaneously adjudicating on pre-recorded footage of match play. Whilst acknowledging the welcomed departure from previous experimental approaches whereby the officials' physical and decision-making performances have been explored in isolation, a number of limitations warrant consideration when interpreting the fidelity of this protocol. In relation to the exercise component of this task, FR alternated 4-minute intervals at 10 $km \cdot h^{-1}$ with 1-minute intervals at 13 $km \cdot h^{-1}$, for a total of 60 minutes. Although measures of total distance were comparable to those achieved during match play (Castagna et al., 2007), the extent to which this protocol sufficiently replicates the intermittent activity profile of soccer FR is questionable. Indeed, soccer match play necessitates that officials engage in periods of both low-intensity activity such as standing or walking as well as periods of intense activity such as high-speed running ($\geq$18 $km \cdot h^{-1}$) (Krustrup et al., 2009). The protocol developed by Samuel and colleagues (2019) therefore appears to be of a relatively moderate-intensity as it excludes the frequent bouts of high-intensity activity that are commonplace during matches. The reproducibility of the physiological responses elicited during this protocol also remains unclear. Concerns also exist

with regards to the fidelity of the decision-making task as officials were presented with match broadcast footage (filmed from an elevated position in the grandstand) on a small treadmill-mounted monitor. Such approaches have however been criticised as they do not replicate the in-game perspective of the official (Craig, 2013). Thus, scope exists for the further development and refinement of simulation-based protocols that more closely replicate the physical and decision-making demands of soccer officiating. Accordingly, the aims of this study were to: 1) evaluate the validity of the physiological and perceptual responses elicited during a novel Soccer Referee Simulation (SRS) and compare these to the internal loads recorded during competitive match play; and 2) ascertain the levels of absolute and relative reliability associated with this protocol.

5.3. METHODS

5.3.1. Experimental Approach

The present investigation comprised two parts: 1) determining the validity of the SRS in relation to the internal loads recorded during competitive matches; and 2) determining the test-retest reliability of the physiological and perceptual responses elicited during the SRS (Figure 5.1). To establish the validity of the SRS, a cohort of sub-elite FR attended the laboratory on two separate occasions. Following initial preliminary measurements and habituation procedures, a single trial of the SRS was completed whereby officials' HR, oxygen uptake ($\dot{V}O_2$), blood lactate concentrations, and differential ratings of the perceived exertion (RPE) were assessed. These outcome variables were selected on the basis that they are regularly adopted within applied settings and have previously been used to assess the physiological and perceptual responses of soccer FR during match play (Castagna et al., 2007; Weston et al., 2012). The officials' HR responses were also monitored during a total of 40 competitive matches (5 match observations per official) and compared against those elicited during the SRS.

To assess the test-retest reliability of the SRS, a repeated measures design was employed comprising an initial habituation session, followed by three separate trials of the SRS. As the officials' competitive schedules did not permit their participation in multiple trials within a short period of time, a cohort of well-trained males with soccer experience and a comparable physiological profile participated in the test-retest aspect of the investigation. Although these participants were not qualified soccer officials, previous playing experience and a comparable physiological profile was deemed sufficient, as we were interested in exploring the reliability of the physiological and perceptual responses elicited during the SRS, and not decision-making accuracy per se. The absolute and relative reliability of the selected physiological and perceptual outcome variables were subsequently ascertained.

5.3.2. Participants

Eight male soccer FR (age: 30.1 ± 3.8 years; stature: 178.4 ± 8.8 cm; body mass: 77.1 ± 10.7 kg; $\dot{V}O_{2max}$: 53.2 ± 4.1 mL·kg·min^{-1}) participated in the first part of this study. At the time of investigation, participants were enrolled on the Scottish Centre of Refereeing Excellence (SCORE) programme – a two-year developmental pathway designed to accelerate the development of high-potential soccer officials. Officials possessed 7.0 ± 1.4 years of officiating experience and had officiated at a national level for 2.3 ± 1.5 years. On average, the FR trained for 3.7 ± 0.8 h/week and officiated 1-2 competitive matches per week. For the test-retest reliability element of the investigation, eight well-trained males (age: 25.1 ± 4.2 years; stature: 177.6 ± 9.0 cm; body mass: 79.6 ± 12.0 kg; $\dot{V}O_{2max}$: 50.6 ± 4.8 mL·kg·min^{-1}) participated. The maximal oxygen uptake ($\dot{V}O_{2max}$) of these participants was similar to that reported previously amongst elite soccer FR (Casajus & Castagna, 2007; Castagna et al., 2019). Participants engaged in regular high-intensity intermittent training (2-3 times per week) and possessed previous soccer playing experience. No contraindications to high-intensity exercise

were disclosed, as ascertained by a health-screening questionnaire (Appendix E), and informed written consent was obtained from all participants prior to testing (Appendix A and B). The study received institutional ethical approval (Appendix F) and conformed to the Declaration of Helsinki.

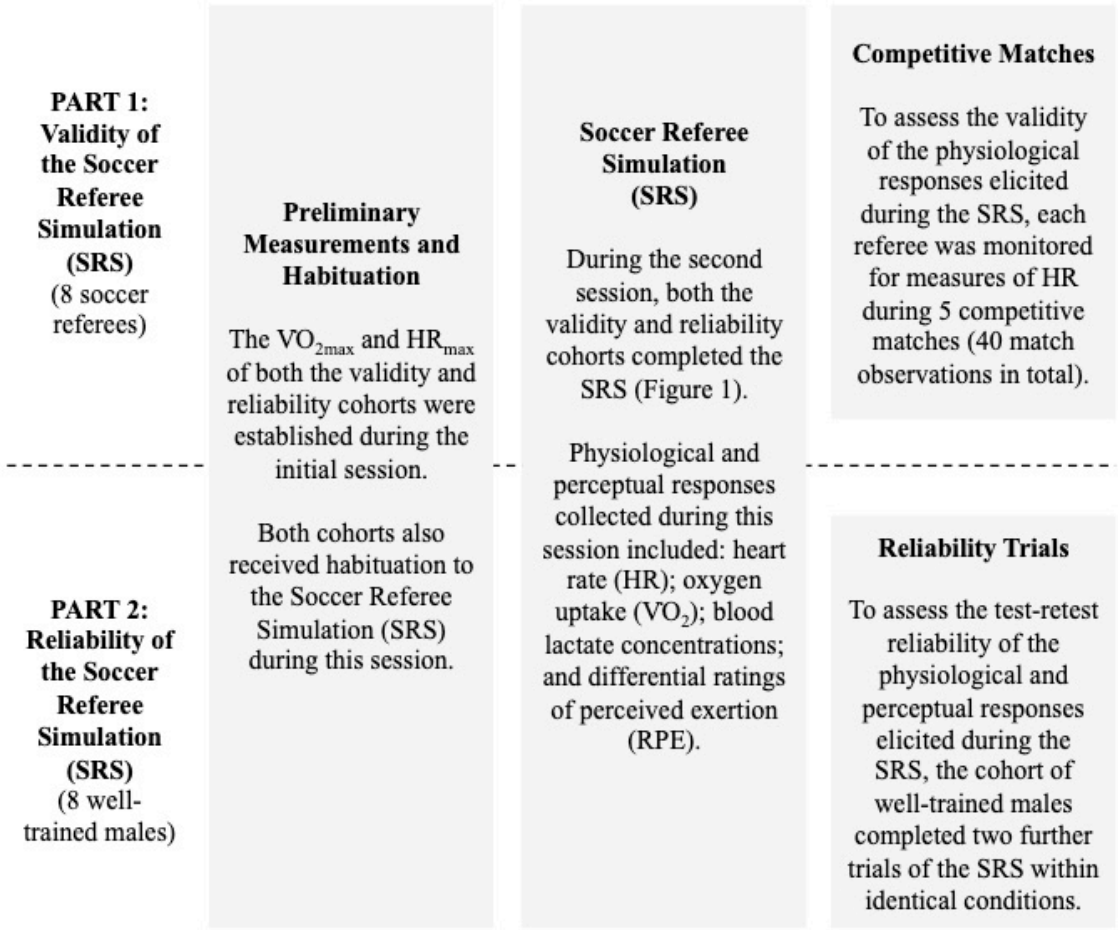


Figure 5.1. Schematic representation of study design.

5.3.3. Procedures

During the first visit of both parts of the investigation, the $\dot{V}O_{2max}$ and HR_{max} of participants were established. This initial visit also served to habituate participants to the SRS. During the subsequent trials, the SRS was performed under identical environmental conditions (temperature: 19°C; humidity: 40%). Each session lasted no longer than 45 minutes, with a minimum and maximum of 3- and 7-days separating trials, respectively. Successive trials were

performed at the same time of day (± 1h) to eliminate diurnal influence on intermittent exercise performance (Aloui et al., 2017). Participants were requested to refrain from strenuous exercise, alcohol, and caffeine intake in the 24h preceding each trial, and to standardise food and fluid intake prior to each visit (Jeacocke & Burke, 2010). All participants provided verbal confirmation of compliance with these instructions.

5.3.3.1. Preliminary Measurements and Habituation Procedures

Participants' $\dot{V}O_{2max}$ and HR_{max} were established during a ramp incremental test performed on a motorised treadmill (Woodway PPS 55sport-I, USA). Following a 5-minute warm-up, participants commenced running at 8 $km \cdot h^{-1}$ for the initial two minutes, with the treadmill speed increased by 1 $km \cdot h^{-1}$ every minute until reaching 15 $km \cdot h^{-1}$. Thereafter, running speed remained constant but the treadmill gradient was increased by 1% every minute until volitional exhaustion (Sperlich et al., 2015). Participants were instructed to perform to the best of their ability and received verbal encouragement throughout. Respiratory measurements were obtained throughout the test by means of breath-by-breath gas analysis (Jaeger Oxycon Pro, Germany) and HR was monitored via a chest-worn HR monitor (Polar H10, Finland). $\dot{V}O_{2max}$ was taken as the single highest $\dot{V}O_2$ value recorded using 15-breath rolling averages (Scheadler et al., 2017) with HR_{max} defined as the highest value recorded. A maximal effort was considered to have been given upon achievement of at least two of the following criteria: 1) respiratory exchange ratio (RER) ≥1.1; 2) plateau in $\dot{V}O_2$ (increase of <2 $ml \cdot kg \cdot min^{-1}$) despite an increasing speed; and 3) HR within ±10 $beats \cdot min^{-1}$ of age-predicted HR_{max} (Midgley et al., 2007). Upon completion of the test, participants were provided ~15 minutes recovery before being habituated to the SRS. During the habitation process, a single 16-minute block of the SRS was completed, during which participants were accustomed to the stochastic velocity profile of the SRS. Given the externally regulated nature of the SRS, this habitation process

was deemed adequate, with all participants providing verbal confirmation of familiarity with the protocols.

5.3.3.2. Soccer Referee Simulation (SRS)

During the experimental trials, participants completed the SRS on a motorised treadmill (Woodway PPS 55sport-I, USA). The protocol, which represents a modified version of previous team-sport simulations (Sirotic & Coutts, 2007; Aldous et al., 2014), was adapted to replicate the non-uniform movement patterns exhibited by elite soccer FR during match play. Accordingly, two ~16-minute blocks interspersed with a 90-s passive recovery period were performed. This recovery period mimicked the extended breaks in play that frequently occur during match play (Spitz et al., 2021) and created a window for the collection of physiological performance markers. Whilst future research may wish to extend the SRS to 45 or 90 minutes, we considered a shortened or condensed protocol to be more appropriate in the current context given the challenges associated with conducting research with athletes during the in-season. That is, given the busy match schedules of soccer officials during the in-season, it was not possible for officials to complete multiple full-duration match simulations within a short period of time. Meanwhile, in anticipating the potential for the SRS to be adopted as a novel training tool, we considered that it would be more practical to implement a shorter protocol into the routine training schedules of soccer officials during the in-season. The protocol incorporated varying periods of standing (0 $km \cdot h^{-1}$), walking (6 $km \cdot h^{-1}$), jogging (11 $km \cdot h^{-1}$), cruising (15 $km \cdot h^{-1}$), and high-speed running (18 $km \cdot h^{-1}$), with the frequency and duration of each occurrence based upon previous literature (Krustrup et al., 2009; Barbero-Alvarez et al., 2012). Due to the impracticalities associated with changing speed every few seconds on a motorised treadmill, the frequency and duration of occurrences were manipulated by a factor that resulted in an activity change every 6-23 s (Thatcher & Batterham, 2004). This resulted in 145 activity

changes, with the rate of acceleration and deceleration programmed between each activity change set at 2 $m \cdot s^{-2}$. Match play data suggest that ~36% of the accelerations and decelerations performed by elite FR during match play are performed at rates of 1.5-2.5 $m \cdot s^{-2}$ (Castillo et al., 2018). Activity changes were communicated to participants in the form of a visual countdown displayed prior to each activity on a large 40-inch monitor (NEC MultiSync LCD4010, Japan) positioned ~2 m in front of the treadmill. Based upon the modelling of the SRS and the clustering of high-speed running, the peak running demand elicited during the SRS was 222 $m \cdot min^{-1}$. The frequency and duration of each activity change are outlined within Table 5.1, with a schematic representation of the SRS activity profile presented in Figure 5.2.

Table 5.1. Frequency and duration of each activity and relative time spent within each movement category during the Soccer Referee Simulation (SRS).

	Treadmill speed ($km \cdot h^{-1}$)	Number of occurrences (n)	Duration per occurrence (s)	Relative time spent in each activity (%)
Stand	0	20	10	13.9
Walk	6	34	13	22.9
Jog	11	35	23	37.3
Cruise	15	32	12	18.8
High-speed run	18	24	6	7.1

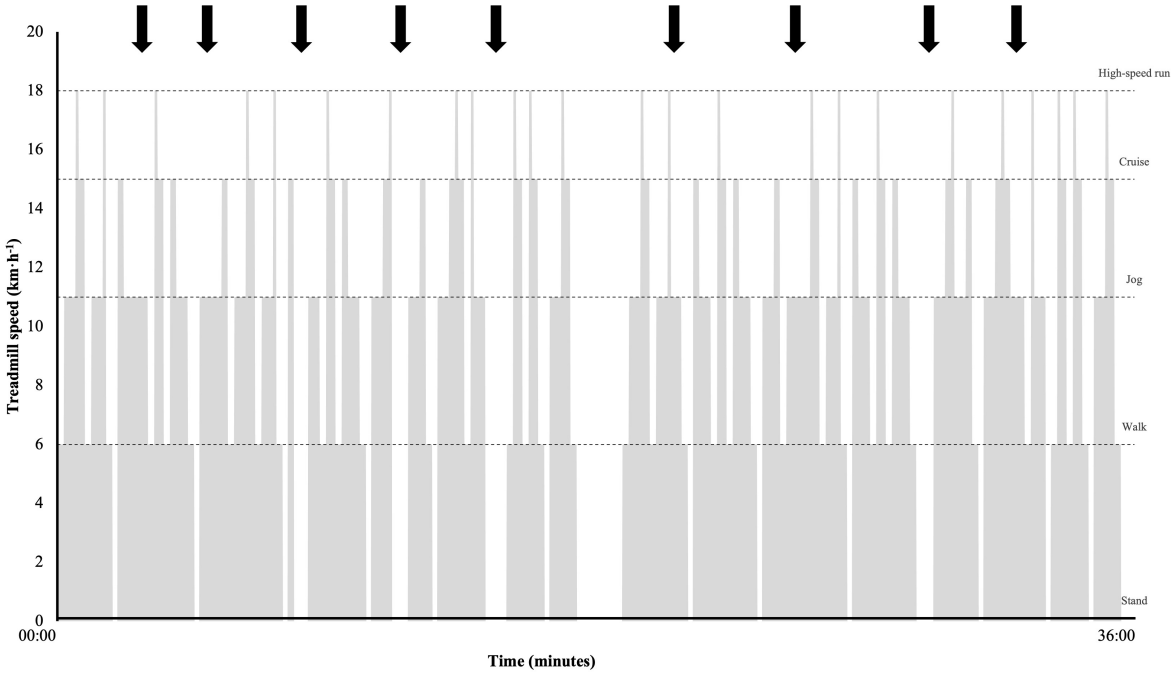


Figure 5.2. Schematic representation of the Soccer Referee Simulation (SRS) activity profile. Downward arrows indicate where each of the decision-making clips (n=10) were presented.

To replicate the dual-task nature of soccer officiating whereby the physical and decision-making demands are imposed concurrently, a series of soccer-specific decision-making clips were presented sporadically throughout the SRS. Video clips (n=10) were sourced from the Refereeing Department at the Scottish Football Association and represented foul play situations that occurred during club and international matches of the Union of European Football Associations (UEFA) (Spitz et al., 2018a). Video clips were edited to 14-29 s with the elapsed time of the match, the score line, and the background sound removed to control for the potential impact of contextual information on participants' decisions (Catteeuw et al., 2010b). Each clip represented a foul play situation in the central area of the field whereby assistance from the AR would likely be limited. The frequency with which clips were presented remained consistent with the relative number of fouls that occur during a match (Weston et al., 2011b; Mallo et al., 2012). Upon viewing each clip, participants provided a

verbal judgment on whether an offence had been committed (i.e., foul/no foul) and whether to caution the player (i.e., no caution/yellow card/red card). Clips were administered during either a stand (n = 5) or jog (n = 5) phase, with clip placement ensuring that a clip succeeded each of the five different movement activities (i.e., stand, walk, jog, cruise, and high-speed run) on at least two occasions. In the context of the present study, we simply sought to ascertain the validity and reliability of the physiological and perceptual responses elicited during the SRS, with the decision-making component of the SRS included to maintain the integrity of the protocol. Thus, the accuracy of participants' decisions was not assessed in the present study.

5.3.3.3. Physiological and Perceptual Outcome Variables

Throughout each experimental trial, participants' HR was measured continuously at a sampling rate of 1 H_Z via HR telemetry (Polar H10, Finland), whilst $\dot{V}O_2$ was ascertained by means of breath-by-breath analysis (Jaeger Oxycon Pro, Germany). Intensities of effort were subsequently calculated and expressed in relative terms as percentages of participants' HR_{max} and $\dot{V}O_{2max}$. A HR-based training impulse (TRIMP) was also calculated for each trial using Banister's TRIMP (Banister, 1991).

To assess blood lactate concentrations, capillary samples were obtained pre-, mid-, and immediately post-trial from the fingertip and were analysed within 1h of collection using a commercially available bench top analyser (Biosen C Line, Germany). In line with manufacturer guidelines, calibration procedures were performed prior to each trial. The inter-assay coefficient of variation (CV) for blood lactate concentrations was 2.0%.

Using the CR100 scale (Appendix G), participants provided differential RPE to differentiate between central (RPE-B), local muscular (RPE-M), and total (RPE-T) exertion (Weston et al., 2015). To control for the potential influence of acute fatigue upon perceptual responses, differential RPE were collected during the jog phase prior to each of the 10 decision-making

clips (epochs; E1-E10) and were obtained in a counterbalanced manner to eliminate order effects (Weston et al., 2015). Participants were fully habituated with the correct use of this scale during the preliminary session. To obtain an indication of the global intensity associated with the SRS, the average of the 10 ratings was calculated to obtain a single value for each measure.

5.3.3.4. Match play

To quantify the internal loads imposed during actual match play, each FR was monitored for measures of HR during 5 competitive matches, resulting in 40 match observations in total (Figure 5.1). Data were obtained during the first five months of the 2019-20 season (August to December), and were collected during Scottish Championship (n=20) and Scottish League 1 (n=20) matches. Each FR officiated at both levels of competition with measures of HR being similar between levels (Scottish Championship: 80.5 ± 6.9% of HR_{max}; Scottish League 1: 81.1 ± 6.3% of HR_{max}). Data were recorded continuously at a sampling rate of $1H_Z$ using a chest-worn HR monitor (Polar H10, Finland), with each official wearing the same monitor for the entirety of the study. To account for the condensed duration of the SRS, match HR corresponding to the first 36 minutes of each match were retained for analyses, with measures of mean HR, peak HR, and TRIMP subsequently calculated. Additionally, given the large levels of match-to-match variation that underpin the physical performance of soccer FR (Weston et al., 2011b), the mean HR responses recorded during the 5 matches were retained for comparison to the SRS.

5.3.3.5. Semi-Structured Interviews

To further assess the ecological validity of the SRS and gain an understanding of participants' appraisal of its psychological fidelity, a series of short semi-structured interviews (6 ± 1 min)

were conducted with the cohort of FR (n = 8). Interviews were performed face-to-face by the lead researcher and were carried out in a quiet location ~15 minutes following completion of the SRS. A schedule of open-ended and non-leading questions was used to guide interviews and explore participants' perceptions of the SRS (Appendix H). Specifically, participants were asked to comment on both the exercise (e.g., *what is your assessment of the exercise protocol in relation to your activity during a match?*) and decision-making (*e.g., what is your assessment of the decision-making task?*) components of the SRS. Where necessary, clarification and elaboration probes were implemented to ensure that an accurate and in-depth understanding of participants' responses was obtained. With participants' permission, interviews were audio recorded.

5.3.4. Data Analysis

Prior to the use of parametric statistical test procedures, normality of distribution and homogeneity of variance were verified using Shapiro-Wilk's and Levene's tests, respectively. Paired sample t-tests determined whether any differences were present in measures of HR between the SRS and match play, with Cohen's ES used to determine the magnitude of differences. The following criteria were used to interpret the practical significance of findings: trivial, <0.2; small, 0.21-0.6; moderate, 0.61-1.2; large, 1.21-1.99; and very large, >2.0 (Hopkins et al., 2009). One-way ANOVAS with repeated measures examined whether any systematic differences existed in physiological and perceptual responses between reliability trials, with two-way repeated measures ANOVAs (trial x time) used to examine measurement variables expressed over multiple time points (i.e., blood lactate concentrations and differential RPE). Where appropriate, post hoc comparisons (least significant difference) were employed. Intraclass correlation coefficients (ICC) were calculated to ascertain relative reliability and were interpreted as *small* (0.10-0.29), *moderate* (0.30-0.49), *large* (0.50-0.69), *very large*

(0.70-0.89), and *nearly perfect* (≥0.90) (Hopkins et al., 2009). To establish absolute reliability, the typical error of measurement (TEM) and CV were calculated (Atkinson & Nevill, 1998). All reliability statistics (ICC, TEM, and CV) are reported alongside their respective 95% CI. In addition, the smallest worthwhile change (SWC) was calculated for each outcome variable as the between-participant standard deviation multiplied by 0.2 (Buchheit, 2016). Data are presented as means and standard deviations (mean ± SD). Statistical procedures were completed using Statistical Package for Social Sciences (SPSS 26.0, IBM, USA) and statistical significance was set at $P<0.05$.

Qualitative data obtained during semi-structured interviews were analysed inductively by means of thematic content analysis (Sparkes & Smith, 2013). Thematic analyses were performed by the lead author (GM). Firstly, audio files were transcribed verbatim using commercially available software (Otter V2.3.27, United States) and were subsequently double checked to ensure accuracy. To create a clear understanding of the content and ensure familiarity with the data, interview transcripts were thoroughly read and re-read by the lead researcher. Following this, emerging themes were outlined through line-by-line analysis of the transcripts with quotations assigned to categories of higher generality. During analysis, internal homogeneity (that data within a category share clear characteristics) and external heterogeneity (clear differences exist between different categories) was sought (Patton, 2001).

5.4. RESULTS

5.4.1. Validity of the Physiological and Perceptual Responses Elicited During the SRS

Physiological and perceptual responses elicited during the SRS are outlined in Table 5.2.

Table 5.2. Physiological and perceptual responses elicited during the single trial of the Soccer Referee Simulation (SRS) completed by the cohort of field referees (n=8).

Heart Rate	
Mean HR (% of HR_{max})	79.6 ± 6.6
Peak HR (% of HR_{max})	90.0 ± 7.1
TRIMP (au)	58.3 ± 20.2
Oxygen uptake	
$\dot{V}O_2$ (% of $\dot{V}O_{2max}$)	61.5 ± 6.2
Blood lactate concentrations	
Pre-trial ($mmol \cdot l^{-1}$)	0.95 ± 0.17
Mid-trial ($mmol \cdot l^{-1}$)	2.04 ± 0.39
Post-trial ($mmol \cdot l^{-1}$)	3.18 ± 0.55
Differential RPE	
RPE-B (au)	38 ± 14
RPE-M (au)	28 ± 12
RPE-T (au)	37 ± 14

Data presented as mean ± SD. HR_{max}, maximal heart rate; TRIMP, Banister's Training Impulse (Banister, 1991); $\dot{V}O_{2max}$, maximal oxygen uptake capacity; RPE-B, breathlessness; RPE-M, muscular; RPE-T, total.

An illustration of a representative official's HR profile during the SRS and the corresponding time during the first half of a competitive match is presented in Figure 5.3. In relation to the HR responses elicited during the SRS and competitive match play, no differences were observed in measures of mean HR (*P*=0.444; 95% CI, -4.2–2.0%; ES=0.29), peak HR (*P*=0.074; 95% CI, -8.1–0.5%; ES=0.74), or TRIMP (*P*=0.498; 95% CI, -14.5–7.8 au; ES=0.25; Figure 5.4).

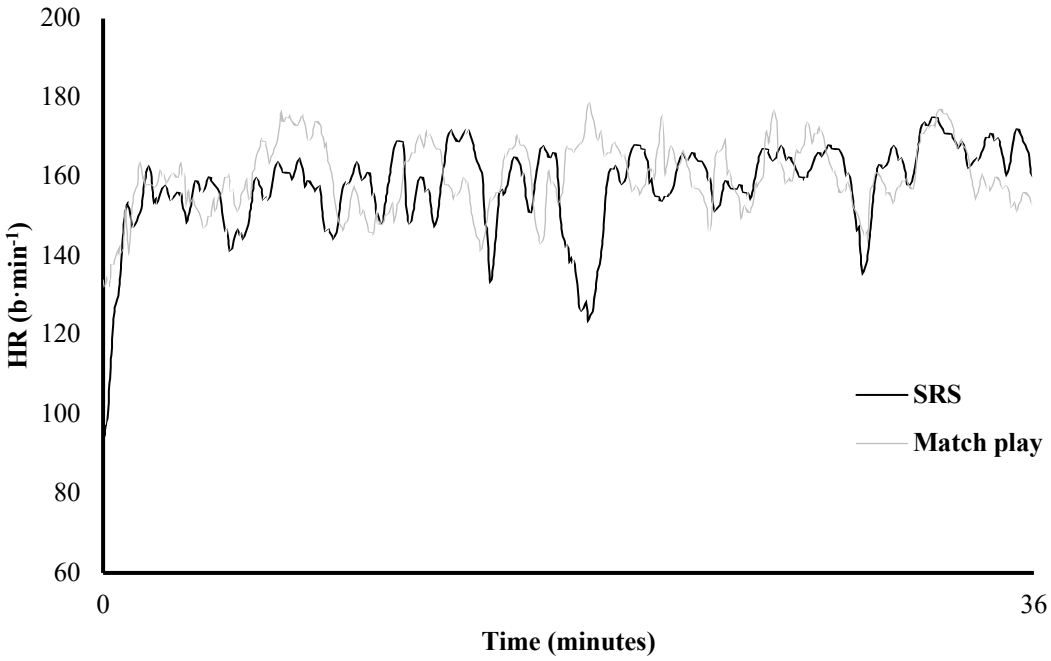


Figure 5.3. An illustration of a representative official's heart rate (HR) profile during the Soccer Referee Simulation (SRS) and the corresponding time during the first half of a competitive match.

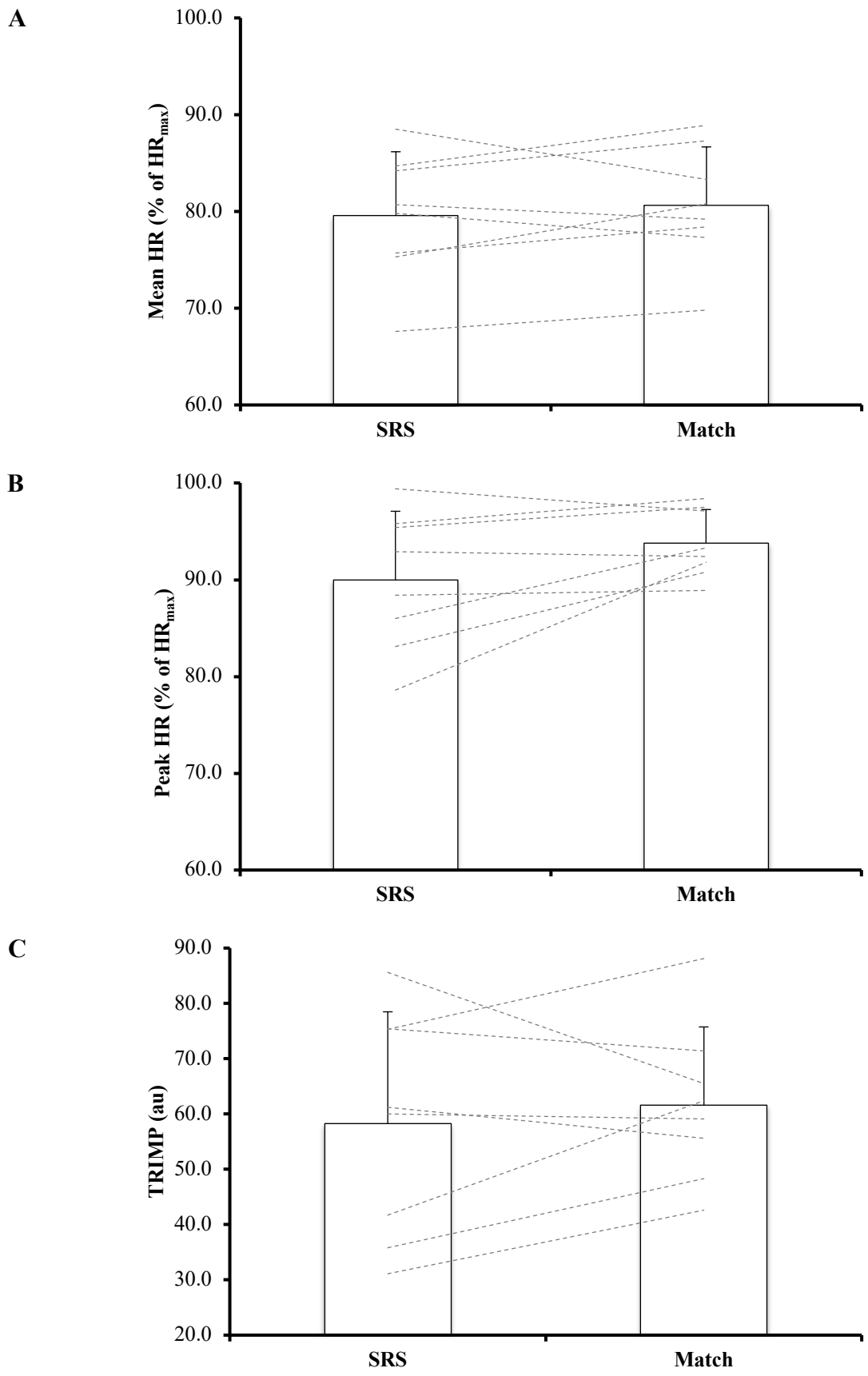


Figure 5.4. Mean heart rate (A), peak heart rate (B), and Banister's TRIMP (C) elicited during the Soccer Referee Simulation (SRS) and competitive match play.

5.4.2. Reliability of the Physiological and Perceptual Responses During the SRS

The relative and absolute reliability of all physiological and perceptual outcome variables are outlined in Table 5.3.

Measures of mean HR (P=0.391), peak HR (P=0.836), TRIMP (P=0.660), and submaximal $\dot{V}O_2$ (P=0.670) were similar between the three reliability trials. Progressive increases in blood lactate concentrations were observed during the SRS ($F_{(2,14)}$=30.317; $P \leq 0.001$) with post-hoc analyses revealing increases to occur from pre- (0.97 ± 0.25 mmol·l^{-1}) to mid-trial (3.01 ± 1.23 mmol·l^{-1}; P=0.001), with further increases observed from mid- to post-trial (3.89 ± 1.47 mmol·l^{-1}; P=0.003). No between-trial differences were detected at any time point ($F_{(2,14)}$=1.342; P=0.293) whilst the pattern of response remained similar between trials ($F_{(4,28)}$=2.195; P=0.095).

No between-trial differences were detected at any time-point for any differential RPE (RPE-B: $F_{(2,14)}$=0.854; P=0.447; RPE-M: $F_{(2,14)}$=0.010; P=0.990; RPE-T: $F_{(2,14)}$=0.732; P=0.498). Progressive increases occurred between E1-E10 in RPE-B ($F_{(9,63)}$=15.316; P=0.000), RPE-M ($F_{(9,63)}$= 3.526; $P \leq 0.001$), and RPE-T ($F_{(9,63)}$=14.931; $P \leq 0.001$). The pattern of response remained similar for all measures between trials (RPE-B: $F_{(18,126)}$=0.547; P=0.930; RPE-M: $F_{(18,126)}$=0.335; P=0.995; RPE-T: $F_{(18,126)}$=0.601; P=0.893).

Measures of mean HR, peak HR, TRIMP, and $\dot{V}O_2$ exhibited good levels of absolute reliability (CV≤10.0%) and nearly perfect levels of relative reliability (ICC≥0.928). Although good levels of absolute reliability were exhibited for blood lactate concentrations assessed post-trial (CV=8.8%), levels were moderate for blood lactate concentrations assessed pre- and mid-trial (CV≥11.3%), with very large to nearly perfect ICC (≥0.891) also present. In relation to d-RPE, moderate levels of absolute reliability (CV≥13.8%) and nearly perfect ICC (≥0.937) were exhibited.

Table 5.3. Absolute and relative reliability of physiological and perceptual responses during the Soccer Referee Simulation (SRS) within the cohort of well-trained males (n = 8).

	Trial 1	Trial 2	Trial 3	ICC (95% CI)	CV (95% CI)	TEM	SWC
Heart Rate							
Mean HR (% of HR_{max})	80.8 ± 7.0	79.1 ± 5.2	79.4 ± 6.6	0.94 (0.79–0.99)	3.1 (2.2–4.1)	2.6	1.3
Peak HR (% of HR_{max})	91.5 ± 6.5	91.2 ± 4.7	90.8 ± 4.6	0.93 (0.76–0.98)	2.2 (1.2–3.2)	2.2	1.1
TRIMP (au)	63.4 ± 17.2	60.4 ± 14.8	62.0 ± 17.6	0.95 (0.83–0.99)	10.0 (6.0–13.9)	5.0	3.3
Oxygen uptake							
$\dot{V}O_2$ (% of $\dot{V}O_{2max}$)	63.4 ± 4.1	63.4 ± 4.0	64.1 ± 5.1	0.95 (0.83–0.99)	2.6 (1.4–3.7)	1.7	0.9
Blood lactate concentrations							
Pre-trial ($mmol \cdot l^{-1}$)	1.0 ± 0.3	1.0 ± 0.2	1.0 ± 0.3	0.89 (0.63–0.98)	11.5 (6.6–16.3)	0.2	0.1
Mid-trial ($mmol \cdot l^{-1}$)	3.2 ± 1.4	3.0 ± 1.3	2.9 ± 1.1	0.96 (0.87–0.99)	11.3 (6.2–16.4)	0.6	0.3
Post-trial ($mmol \cdot l^{-1}$)	4.1 ± 1.8	3.6 ± 1.3	3.9 ± 1.3	0.97 (0.90–0.99)	8.8 (5.1–12.5)	0.4	0.3
Differential RPE							
RPE-B (au)	36 ± 15	33 ± 14	33 ± 10	0.95 (0.61–0.97)	14.0 (9.2–18.8)	5	3
RPE-M (au)	33 ± 14	32 ± 14	32 ± 9	0.94 (0.79–0.99)	13.8 (7.4–20.1)	5	2
RPE-T (au)	35 ± 14	33 ± 14	32 ± 10	0.94 (0.80–0.99)	15.1 (10.3–19.8)	5	3

Data presented as mean ± SD. ICC, intraclass correlation coefficient; CV, coefficient of variation; TEM, typical error of measurement; SWC, smallest worthwhile change; CI, confidence interval; HR_{max}, maximal heart rate; TRIMP, Banister's Training Impulse (Banister, 1991); $\dot{V}O_{2max}$, maximal oxygen uptake capacity; RPE-B, breathlessness; RPE-M, muscular; RPE-T, total.

5.4.3. Fidelity of the SRS

Eight officials (100%) suggested that the activity profile of the SRS was comparable to that of a competitive match, whilst six officials (75%) considered the physiological load to be representative of match play.

> *"It's pertinent to a match, to the movement in terms of sprinting and walking, all that kind of stuff. It's varied because we don't just sprint constantly. For example, we don't just operate constantly at 90% heart rate. It flexes up and down, so in terms of what we did there, it's probably…albeit not necessarily the full 90 minutes but squashed into a sizeable chunk, so I thought it was pertinent"* (Participant 2)

> *"I think it closely replicates the type of thing that we do on the pitch…both with the activity level, changing all the time, going up and down, and also the intensity in terms of our heart rate"* (Participant 5)

Two officials (25%) commented that the high-speed run could perhaps be faster and of a longer duration, with one official (13%) adding that there is perhaps a greater occurrence of low-intensity activities (i.e., standing and walking) during match play. Meanwhile, one official (13%) suggested that their HR would perhaps be slightly more elevated during a match (as compared to the SRS); however, they acknowledged that this might be due to the heightened levels of psychological stress experienced in competitive situations.

> *"You might be sprinting for a bit longer than six seconds…but again, you probably get more standing and walking in a game"* (Participant 3)

"I think it compares quite well. I think that activity in a match is probably slightly elevated, like heart rate. But then a lot of that is psychological which pushes your heart rate up…adrenaline, adrenaline doesn't exist on that in the same manner, but it's the closest thing I've ever come across" (Participant 7)

Six officials (75%) suggested that the decision-making clips were relevant and reflected the types of situations that occur frequently during a match. Meanwhile, four officials (50%) highlighted that the requirement to make decisions in combination with elevated levels of physiological stress was reflective of what happens during competitive match play, with five officials (63%) adding that the dual-task nature of the SRS required increased levels of concentration (as compared to video-based decision-making alone).

"They were very difficult clips…they were decisions that we would have to make on a routine basis, so I think I was quite tested in that regard" (Participant 6)

"I felt it was very realistic to the game. The fact that the running mirrored what you would do in a game and then a decision appeared, similar to how you would be feeling physically in a match situation. So, I thought overall it was very specific to a game" (Participant 4)

"It is something that I think genuinely would mimic a game in terms of having to be switched on constantly" (Participant 5)

The main differences that were identified between the SRS and actual match play were related to the absence of the psychological stressors that may also influence their decisions. For instance, two officials (25%) noted that the absence of crowd noise alleviated their levels

of psychological stress, whilst two officials (25%) stated that the fear of making a decisional error would be greater during a competitive match.

"I wasn't maybe as stressed. You have nobody in your ear…and there's no external factors other than that decision. There's no like crowd, no dissent, or there's no quick changing of play or so on" (Participant 3)

"You know, there's no comeuppance, there's no retribution for any decision I'd get wrong" (Participant 7)

5.5. DISCUSSION

The current study sought to: 1) determine the validity of the physiological and perceptual responses elicited during the SRS in relation to those recorded during competitive matches; and 2) ascertain the levels of absolute and relative reliability associated with these responses. In relation to the selected HR measures, no differences were detected between the SRS and actual match play, whilst the physiological and perceptual responses observed during the SRS were also aligned with those reported amongst elite soccer FR during match play (Weston et al., 2006; Barbero-Alvarez et al., 2012; Castillo et al., 2016b). The SRS was also found to yield high levels of reproducibility between trials. Overall, our findings have demonstrated the SRS to be a valid and reliable protocol that closely replicates the demands imposed upon soccer FR during match play.

The mean and peak HR observed during the SRS were 79.6% and 90.0% of HR_{max}, respectively, with no differences evidenced in any measure of HR between the SRS and actual match play. Albeit marginally lower, these values are comparable to published match play data whereby elite-level FR attain mean and peak HRs of ~84% and ~97%, respectively

(Weston et al., 2006; Krustrup et al., 2009; Barbero-Alvarez et al., 2012; Castillo et al., 2016b). Similarly, the mean $\dot{V}O_2$ observed during the SRS was 63.5% of $\dot{V}O_{2max}$ with slightly higher values (~68%) achieved during competitive matches (D'Ottavio & Castagna, 2002). Although the slightly reduced cardiorespiratory responses observed in the current study could be attributable to the lower competitive level of the studied cohort, it is worth noting that the SRS was developed based upon the activity profiles exhibited by elite-level FR during match play (Krustrup et al., 2009; Barbero-Alvarez et al., 2012). Further, the peak running demand elicited during the treadmill-based SRS (222 $m \cdot min^{-1}$) exceeds the ~190 $m \cdot min^{-1}$ recently identified amongst elite English (Fereday et al., 2020) and Italian (Riboli et al., 2020) soccer players during actual match play.

Previous accounts detailing the blood lactate concentrations of elite FR following the first half of match play have reported values of 3.5 $mmol \cdot l^{-1}$ (Krustrup et al., 2009). In the current investigation, significant increases in blood lactate concentrations were observed from pre- to mid-trial, with blood lactate concentrations peaking at post-trial (3.18 ± 0.55 $mmol \cdot l^{-1}$). Thus, despite the reduced duration, the SRS elicits metabolic responses that are consistent with those observed during official matches. In relation to differential RPE, the grand mean for RPE-T, RPE-B, and RPE-M were 37 ± 14 au, 38 ± 14 au, and 28 ± 12 au, respectively. As per the CR100 scale, these ratings corresponded to the verbal anchor *"somewhat strong"*. Weston and colleagues (2006) previously observed values of 7.8 au (*"very strong"*) on the 10-point Borg scale within English Premier League FR, whilst the RPE-B and RPE-M reported by Spanish FR were 6.6 au and 7.1 au (both *"very strong"*), respectively (Castillo et al., 2017). Unfortunately, drawing direct comparisons between these data is challenging, as aside from the use of different scales, previous studies have traditionally obtained one-off ratings between 10-30 min following the conclusion of match play (Weston et al., 2006; Castillo et al., 2017). In the context of the current investigation, however, differential RPE were collected at regular

intervals throughout the SRS (i.e., prior to each decision), with the grand mean of each measure calculated to provide an indication of the global intensity associated with the SRS. Whilst collection procedures often prove impractical within competitive settings, the ability to assess physiological and perceptual responses at frequent intervals represents an important benefit of the SRS. Assessing differential RPE throughout the SRS may yield interesting new insights into how perceptions of exertion develop during simulated match play and provide further evidence of the internal demands associated with soccer officiating.

The physical match performances of soccer FR are subject to large levels of match-to-match variation (Weston et al., 2011b), with the number of high-intensity activities engaged in being particularly variable (CVs of ~20-54%). Such observations bear important implications for the comparison of data derived during matches, as in the absence of considerable sample sizes, real systematic changes in outcome variables can often prove difficult to detect. A promising feature of the current data therefore relates to the extent to which physiological and perceptual responses were reproduced during the SRS. Firstly, we failed to detect any systematic bias between trials for any of the measured outcome variables ($P \geq 0.288$). Additionally, excellent levels of relative reliability were demonstrated for all physiological and perceptual responses (ICC≥0.891), whilst repeat measures of mean HR, peak HR, and $\dot{V}O_2$ exhibited good levels of absolute reliability (CV≤3.1%). Similar levels of variation have been reported previously for measures of HR (CVs of ~1-3%) and $\dot{V}O_2$ (CVs of ~3-5%) during intermittent team sport simulations (Sirotic & Coutts, 2007; Harper et al., 2016b). In relation to differential RPE and blood lactate concentrations collected pre- and mid-trial, moderate levels of absolute reliability were present. Although the CV of these measures exceed the 10% threshold commonly considered acceptable (Atkinson & Nevill, 1998), our data support previous reports whereby blood lactate concentrations generally appear less reliable than other physiological measures such as HR and $\dot{V}O_2$ (Sirotic & Coutts, 2007; Russell et al., 2011; Harper et al., 2016b).

Interestingly, good levels of absolute reliability were observed for post-trial blood lactate concentrations (CV=8.8%), with these levels comparing favourably to the ~20-34% observed following other intermittent team sport simulations (Sirotic & Coutts, 2007; Russell et al., 2011; Harper et al., 2016b). Such discrepancies are likely explained by the externally regulated nature of the SRS. Whereas the external loads performed during the SRS were replicated between trials, simulations that utilise free-running protocols possess some inherent variation in the test itself, which in turn will influence the variability of physiological responses.

Whilst the SRS represents a valid and reliable protocol that induces physiological responses representative of match play, a number of limitations warrant consideration. Firstly, the absence of changes of direction and unorthodox movements (i.e., sideways and backwards running) may compromise the ecological validity of treadmill-based protocols (Aldous et al., 2014). Notwithstanding, the modelling of treadmill simulations to incorporate frequent activity changes of high intensities can elicit biomechanical responses that are representative of those during match play (Page et al., 2015). The inability to express maximal running velocities represents another potential limitation of motorised treadmill protocols (Williams et al., 2010). It is noteworthy, however, that only ~2% of the total distance covered by elite FR during match play is by means of sprinting, whilst maximal velocities are not commonly achieved during official matches (Krustrup et al., 2009; Castillo et al., 2016a). The inability of such protocols to incorporate very high-intensity activities should therefore be considered with respect to their capacity to reliably impose physiological loads consistent with match play. We must also acknowledge that whilst previous simulations have reported physiological and perceptual responses to stabilise after familiarisation and one trial (Sirotic & Coutts, 2007), a relatively short habituation period was provided to participants in the present investigation. Nevertheless, given the externally regulated nature of the SRS, the high levels of reproducibility observed for physiological and perceptual outcome variables, and the absence of any systematic

differences between trials 1-3, the habituation procedures adopted in the present study appear suitable.

5.5.1. Conclusion and Perspective

The SRS produced highly reproducible physiological and perceptual responses that are consistent with those elicited by elite FR during competitive match play. A number of potential applications therefore exist for the SRS. As identified within Chapter 4, soccer officials currently engage in low volumes of decision-making training, with their exposure to match-related decision-making being largely limited to competitive matches. The SRS may therefore hold promise as a potential training tool and could be used to provide officials with additional opportunities to make decisions under game-like conditions. Nonetheless, whether regular engagement within the SRS would improve the decision-making abilities of soccer officials remains unclear and requires further investigation. The reproducibility of the physical and physiological responses produced during the SRS also lends support to its use within research settings, with a potential application being to assess the association between the physiological and decision-making performances of soccer FR. As identified within Chapter 2, previous literature has thus far failed to delineate the influence of physical load on in-match decisions of soccer officials. The SRS would allow this association to be explored in a more controlled environment and in the absence of the contextual factors that confound match decisions.

Chapter 6

Decision-Making Accuracy of Soccer Referees in Relation to Markers of Internal and External Load

6.1. ABSTRACT

Purpose: To examine the decision-making performances of soccer referees (FR) in relation to markers of internal and external load. **Methods:** Following the collection of baseline measures and habituation procedures, 13 sub-elite male FR completed the Soccer Referee Simulation (SRS). At various stages throughout the SRS, a video clip depicting a potential foul play situation was presented, with FR requested to make a technical and disciplinary decision. The correctness of each decision was subsequently assessed in relation to the mean heart rate (HR), respiratory rate (RR), minute ventilation ($\dot{V}E$), and running speeds recorded in the 10-s and 60-s preceding decisions. Associations between the decision-making accuracy of FR and perceptions of central (RPE-B) and local muscular (RPE-M) exertion were also explored. **Results:** A significant association was observed between decision-making accuracy and the mean HR (*P*=0.042; ES=0.272) and RR (*P*=0.024, ES=0.239) recorded in the 10-s preceding the decision, with significantly more incorrect decisions observed when HR ≥90% of HR_{max} (OR, 5.39) and RR ≥80% of RR_{peak} (OR, 3.34). Decision-making accuracy was also significantly associated with the mean running speeds performed in the 10-s (*P*=0.003; ES=0.320) and 60-s (*P*=0.016; ES=0.253) prior to the decision, with workloads of ≥250 $m \cdot min^{-1}$ increasing the likelihood of a decisional error (OR, 3.84). Finally, a significant association was found between decision-making accuracy and RPE-B (*P*=0.021; ES=0.287), with a disproportionate number of incorrect decisions occurring when RPE-B was rated as *"very strong"* to *"maximal"* (OR, 7.19). **Conclusions:** Collectively, the current data offer novel insights into the detrimental effects that high workloads may have upon the decision-making performances of soccer FR.

6.2. INTRODUCTION

Soccer officials are charged with ensuring that match play is contested in accordance with the Laws of the Game (FIFA, 2016). A key aspect of this responsibility relates to the identification of incidents of foul play, with FR awarding on average ~26 fouls per match (Weston et al., 2011b; Mallo et al., 2012). As decisional errors may impact a match's outcome, the correctness of such judgments is of utmost importance (Červený et al., 2018). To enhance their perception of potential infringements, FR endeavour to keep up with play and adopt viewing positions that afford both an unobstructed view and limit their interference with the ball or players (Mallo et al., 2012). Consequently, elite FR cover distances of 9-11 km during competitive matches, with high-speed running (≥18 $km \cdot h^{-1}$) accounting for ~1800 m (Krustrup et al., 2009; Weston et al., 2011b; Castillo et al., 2017; Weston et al., 2011c; Barbero-Álvarez et al., 2012). The internal loads elicited during match play are also considerable with elite FR reported to attain mean match HR of 85-95% of HR_{max} (Weston et al., 2006; Castillo et al., 2016b). As the physical and decision-making demands of soccer officiating are imposed concurrently, the relationship between these facets of performance warrants consideration (Weston et al., 2012).

To date, several investigations have explored the decision-making performances of team sport officials in relation to the physical and physiological demands of match play; however, findings have proven equivocal (Mascarenhas et al., 2009; Mallo et al., 2012; Elsworthy et al., 2014b; Emmonds et al., 2015; Larkin et al., 2014b). Traditionally, researchers have ascertained the decision-making accuracy of the in-game official across distinct fixed-time epochs and contextualised these against the internal and external loads imposed during that same period. In an investigation by Mascarenhas and colleagues (2009), the decision-making accuracy of seven soccer FR was found to be unrelated to the average running speed and total distance accumulated during each 15-minute period of match play. Similarly, when explored across 10-

minute intervals, no association was observed between the decision-making performances of 8 Rugby League match officials and measures of mean HR and high-intensity running distance (Emmonds et al., 2015). While these data would appear to indicate that physical load bears no effect upon decision-making performance, the analytical approaches adopted are likely insufficient to adequately assess the intricacies of such a relationship. Indeed, recent research suggests that discrete fixed-time epochs may underestimate physical match demands when compared to rolling averages (Doncaster et al., 2019; Fereday et al., 2020). Aggregating data over prolonged epochs (i.e., 10-15 min) may also conceal acute periods of high physiological stress that could potentially hinder optimal decision-making. Although high-intensity episodes last only seconds (Barbero-Álvarez et al., 2012), it is plausible that such exertions may temporarily impair an official's decision-making. Thus, examining the correctness of decisions immediately following acute periods of high physical exertion may permit a better understanding of the relationship that exists between the physical and decision-making performances of match officials.

Previous research with soccer FR observed no association between the accuracy of the officials' decisions and the total distance covered in the 30-s preceding an incident (Riiser et al., 2019). The decision-making accuracy of 29 Australian Football umpires was also found to be unrelated to the mean running speeds performed in the 30-s to 5 min prior to the decision (Elsworthy et al., 2014a). Conversely, higher relative running speeds performed in the 5-s before a decision was related to an increase in decisional errors (Elsworthy et al., 2014). While it was suggested that such findings could relate to an enhanced physiological demand, this remains speculative as no measures of internal load were obtained. As isolated and repeated bouts of high-speed running frequently precede crucial moments within match play (Faude et al., 2012), additional research is required to further explore the impact of such moments on the decision-making accuracy of soccer officials. The purpose of the present study was therefore

to assess the decision-making performances of soccer FR in relation to the internal and external loads recorded in the moments immediately preceding decisions.

6.3. METHODS

6.3.1. Participants

Thirteen sub-elite male FR (age: 30.4 ± 4.1 years; stature: 177.5 ± 7.5 cm; body mass: 76.9 ± 10.2 kg; $\dot{V}O_{2max}$: 53.5 ± 3.5 ml·kg·min^{-1}) from the Scottish Football Association (SFA) participated in the current investigation. Eight of these officials also participated in Chapter 5. At the time of the study, participants were enrolled on the SCORE developmental programme. Participants possessed 7.9 ± 2.0 years of officiating experience and had officiated at a national level for 4.5 ± 1.9 years. On average, participants trained for 3.7 ± 0.8 h/week and officiated 1-2 matches per week. Prior to the commencement of the study, a health-screening questionnaire (Appendix E) was completed to confirm each official's eligibility for intensive exercise testing, with informed consent subsequently obtained from all participants (Appendices A and B). Institutional ethical approval was received (Appendix I), and all experimental protocols conformed to the Declaration of Helsinki.

6.3.2. Preliminary Measurements and Habitation

Officials attended the laboratory on two separate occasions during the in-season (October to December). A maximum of 7 days separated trials with participants abstaining from strenuous exercise in the 48-h preceding each session. To account for the effects of circadian variation on cardiorespiratory responses during exercise (Knaier et al., 2019), sessions 1 and 2 were completed at a similar time of day (±1h) for each participant and were performed under standardised environmental conditions (temperature: 19°C; relative humidity: 40%).

The preliminary measurements and habituation procedures were identical to those detailed within Chapter 5 (please refer to Section 5.3.3.1). Briefly, participants performed a ramp incremental test until volitional exhaustion, with their HR and respiratory variables monitored continuously by means of HR telemetry (Polar H10, Finland) and breath-by-breath gas analysis (Jaeger Oxycon Pro, Germany). In addition to the data analyses performed to identify participants' HR_{max} and $\dot{V}O_{2max}$, respiratory rate (RR) and minute ventilation ($\dot{V}E$) were averaged on a 5-s basis, with the highest value recorded in a 20-s period retained as RR_{peak} and $\dot{V}E_{peak}$, respectively (Buchheit et al., 2009).

A single 16-minute block of the SRS was completed during the habitation process of Chapter 6, during which participants were accustomed to the stochastic velocity profile of the SRS and completed 5 practice video clips. Crucially, all participants were familiar with the video-based testing having performed this regularly as part of their training within the 2 years previously. Although relatively brief, the findings of Chapter 5 suggest these procedures to be sufficient, with physiological and perceptual responses found to stabilise after a short habitation period.

6.3.3. Main Trial

During the second session, the SRS developed in Chapter 5 was performed on a programmable motorised treadmill (Woodway PPS 55sport-I, USA). The validity and reliability of the SRS have been described previously in Chapter 5.

At varying intervals throughout the SRS, a series of soccer-specific decision-making clips (1280 x 720 pixels) were presented to participants by means of a 40-inch monitor (NEC MultiSync LCD4010, Japan) positioned ~2 m in front of the treadmill. Video clips (n = 10) were sourced from the Refereeing Department at the Scottish Football Association and represented foul play situations that occurred during club and international competitive

matches of the Union of European Football Associations (UEFA) (Spitz et al., 2018a). The number of clips presented during the SRS was consistent with the number of fouls that occur during match play (Weston et al., 2011b; Mallo et al., 2012). Video clips were selected following consultation with two former international standard FR (a combined total of 26 years on the FIFA list) and were deemed acceptable if they: 1) represented a foul play situation in the central area of the field whereby input from the AR would be limited; 2) were of high visual quality and were presented from an in-game perspective; and 3) omitted the in-game official's decision from view (Figure 6.1). To control for the potential impact of contextual factors upon participants' decisions, information concerning the elapsed time of the match and the score were removed, as was the background sound. To identify the reference decision for each foul play situation, the two former FR independently assessed each video clip as per Law 12 (Fouls and Misconduct) of the Laws of the Game (FIFA, 2018). To facilitate their decision-making, video clips were able to be viewed multiple times in both real time and slow motion (Pizzera et al., 2016; Spitz et al., 2017). Complete agreement was exhibited between experts for each clip. Moreover, although the Laws of the Game leave room for the individual interpretation of each official, clips were deemed to be of a similar level of difficulty, as per the opinion of the expert panel. Video clips were administered during either a stand (n = 5) or jog (n = 5) phase, with a different clip succeeding each of the five different activity categories (i.e., stand, walk, jog, cruise, high-speed run) on at least two occasions. Upon viewing each clip, participants were afforded 6-s to make a technical (no foul; indirect free kick; direct free kick; or penalty) and disciplinary (no card; yellow card; or red card) decision (Catteeuw et al., 2010b; Schweizer et al., 2011). Participants' decisions were assessed against the reference decisions made by the expert panel and were subsequently categorised as correct or incorrect.

Figure 6.1. Example of a foul-play situation for which officials made a technical (no foul; indirect free kick; direct free kick; or penalty) and disciplinary (no card; yellow card; or red card) decision.

Throughout each experimental trial, participants' HR was measured continuously at a sampling rate of 1 H_Z via HR telemetry (Polar H10, Finland). As discussed in Chapter 2, fixed epochs have been reported to underestimate physiological demands (Varley et al., 2012; Doncaster et al., 2019; Fereday et al., 2020), whilst it has also been shown that measures of relative intensity increase as the epoch duration decreases (Delaney et al., 2017; Martin-Garcia et al., 2018; Doncaster et al., 2019). Thus, in an attempt enhance the sensitivity of our analyses and increase the likelihood of capturing the most intense periods during the SRS, HR data were processed using 10-s and 60-s rolling averages, with the mean HR recorded prior to each decision retained for analysis. Mean HR data were subsequently expressed in relative terms as a percentage of participants' HR_{max} before being classified into four distinct zones: 60-69% HR_{max}; 70-79% HR_{max}; 80-89% HR_{max}; and >90% HR_{max} (Edwards, 1993). Participants' respiratory responses during the SRS were assessed by means of breath-by-breath gas analysis (Jaeger Oxycon Pro, Germany). The mean RR and $\dot{V}E$ recorded in the 10-s and 60-s preceding

each decision was subsequently calculated and retained as a percentage of participants' RR_{peak} and $\dot{V}E_{peak}$, respectively (Buchheit et al., 2009).

Using the CR100 scale (Appendix G), participants provided differential RPE to delineate between central (RPE-B) and local muscular (RPE-M) exertion (Weston et al., 2015). In contrast to the more commonly employed CR10 scale, previous research has reported the CR100 scale to provide more precise measures of exercise intensity (Franchini et al., 2016). To control for the potential influence of acute fatigue upon perceptual responses, differential RPE were collected during the jog phase preceding each decision and were obtained in a counterbalanced manner to eliminate order effects (Weston et al., 2015). Participants were fully habituated with the correct use of this scale during the preliminary session and were given instruction on how to appraise differential RPE. Specifically, participants were informed that RPE-B depends mainly on breathing rate and/or heart effort, and RPE-M depends mainly on the strain and exertion in the lower limbs (McLaren et al., 2020). For the differential RPE obtained prior to each decision, ratings were classified into four arbitrary categories: *"nothing at all* to *moderate"* (0-25 au); *"moderate* to *strong"* (26-50 au), *"strong* to *very strong"* (51-75 au), and *"very strong* to *maximal"* (76-100) (Franchini et al., 2016; Lovell et al., 2020).

The average running speeds performed in the 10-s and 60-s preceding each decision were also calculated. To aid the practical application of the findings, running speeds were expressed as a relative measure ($m \cdot min^{-1}$) and were categorised as: <150 $m \cdot min^{-1}$; 150-199 $m \cdot min^{-1}$; 200-249 $m \cdot min^{-1}$; and ≥250 $m \cdot min^{-1}$.

6.3.4. Data Analysis

To explore associations between decision-making accuracy and markers of internal and external load, the number of correct and incorrect decisions per category was calculated for each variable, with chi-squared tests of independence subsequently performed (Nevill et al.,

2002). The magnitudes of associations were assessed using Cramer's V ES and were interpreted as: trivial (<0.1); small (0.1-0.29); moderate (0.3-0.49); and large (≥0.5) (Cohen, 1992). To identify cells where observed counts deviated from an expected equal distribution, standardised residuals (SR) were calculated with values less than -2 and greater than 2 considered statistically significant (Agresti, 2007). Odds ratios (OR), with uncertainty expressed as 95% CI, were subsequently calculated to examine the likelihood of making a decisional error associated with each variable. Statistical procedures were completed using Statistical Package for Social Sciences (SPSS 26.0, IBM, USA), with statistical significance set at $P<0.05$.

6.4. RESULTS

6.4.1. Decision-Making Accuracy

Of the 130 decisions assessed, 31 errors were made, representing an overall decision-making accuracy of 76.2%. When analysed per participant, decision-making accuracy ranged from 60-90%. Of the 31 errors, 29 resulted from an incorrect disciplinary decision, with 2 being the result of an incorrect technical decision.

6.4.2. Decision-Making Accuracy in Relation to Heart Rate (HR)

A significant association was found between decision-making accuracy and mean HR in the 10-s preceding a decision ($\chi^2_{(3)} = 8.181$, $P = 0.042$, ES = 0.272), with significantly more incorrect decisions observed when HR ≥90% of HR_{max} (SR, 2.6; OR, 5.39; 95% CI, 1.70–17.07) (Figure 6.2). No association was observed between decision-making accuracy and mean HR in the 60-s preceding a decision ($\chi^2_{(3)} = 0.325$, $P = 0.955$, ES = 0.051).

6.4.3. Decision-Making Accuracy in Relation to Respiratory Variables

A significant association was found between decision-making accuracy and mean RR in the 10-s preceding a decision ($\chi^2_{(2)} = 7.445$, $P = 0.024$, ES = 0.239), with significantly more incorrect decisions observed when RR ≥80% of RR_{peak} (SR, 2.1; OR, 3.34; 95% CI, 1.32–8.45) (Figure 6.3). No association was observed between decision-making accuracy and mean RR in the 60-s preceding a decision ($\chi^2_{(2)} = 1.085$, $P = 0.581$, ES = 0.091).

No association was observed between decision-making accuracy and mean $\dot{V}E$ in the 10-s ($\chi^2_{(2)} = 4.144$, $P = 0.126$, ES = 0.179) or 60-s ($\chi^2_{(2)} = 0.955$, $P = 0.620$, ES = 0.086) preceding a decision (Figure 6.4).

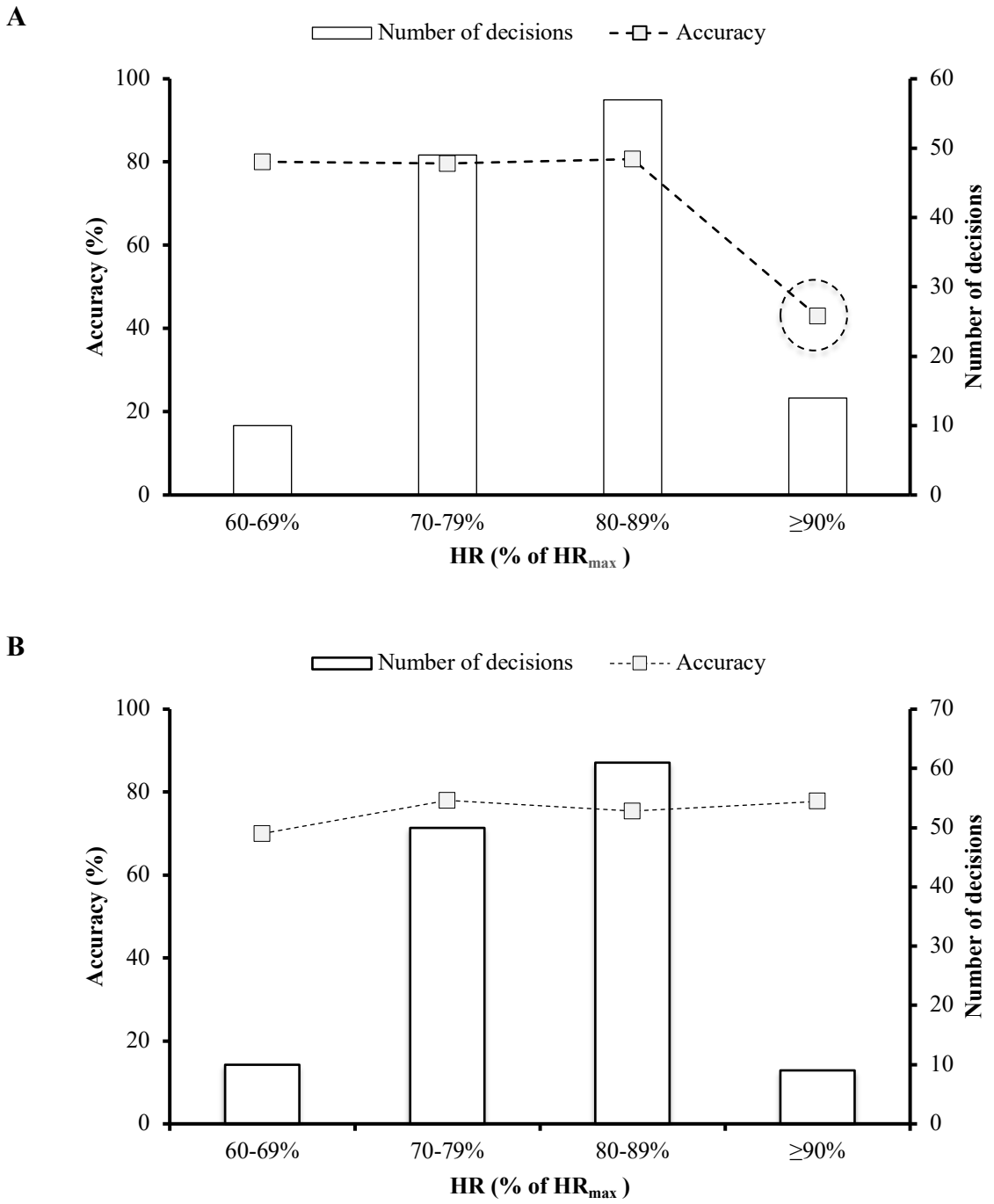


Figure 6.2. Decision-making accuracy in relation to mean heart rate (HR) in the 10-s (A) and 60-s (B) preceding the decision. Dashed circle signifies that significantly more incorrect decisions were observed when HR ≥90% of HR_{max}.

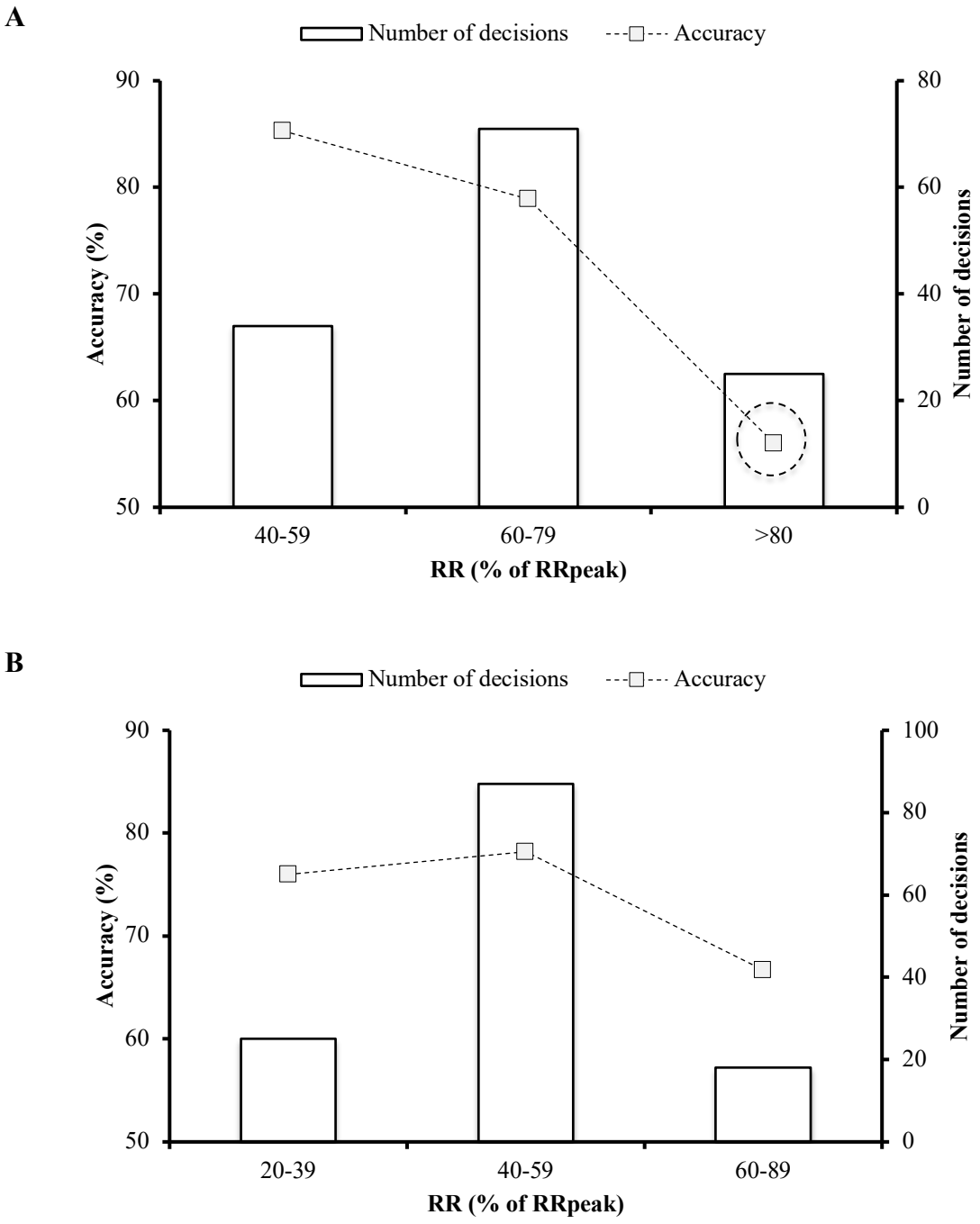
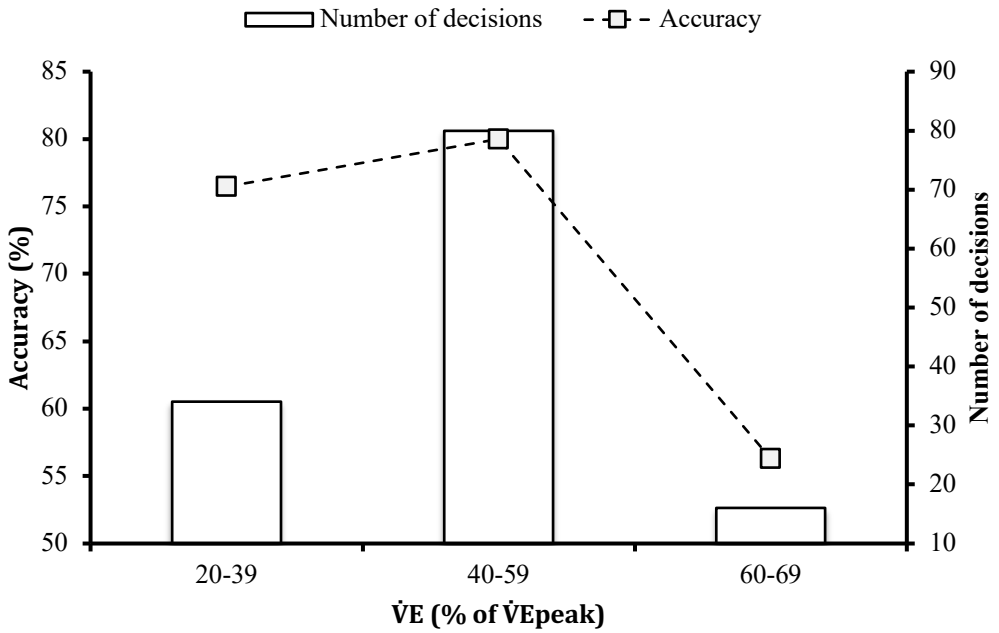


Figure 6.3. Decision-making accuracy in relation to mean respiratory rate (RR) in the 10-s (A) and 60-s (B) preceding the decision. Dashed circle signifies that significantly more incorrect decisions were observed when RR ≥80% of RR_{peak}.

A



B

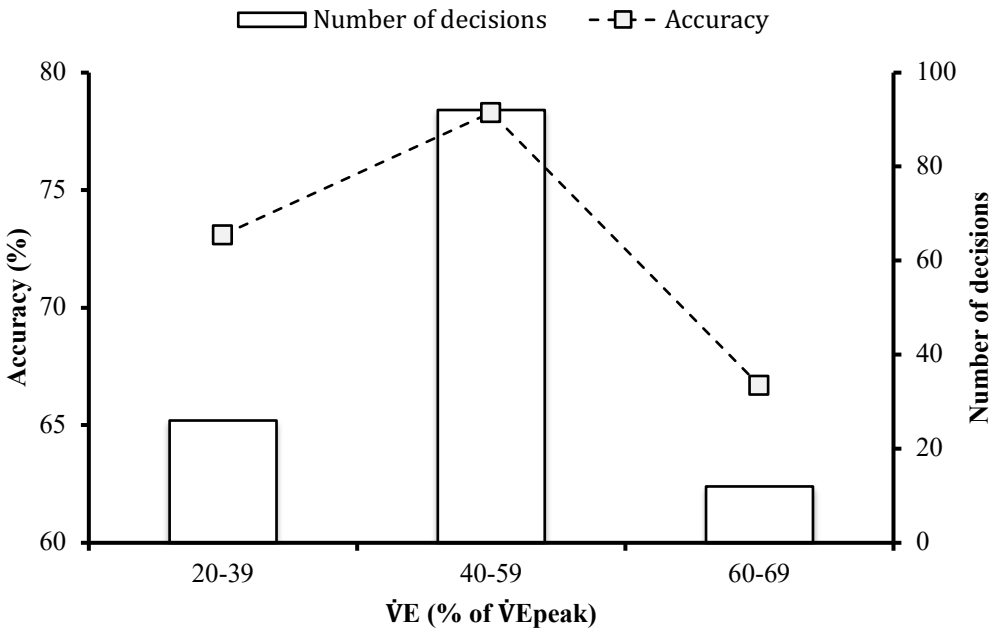


Figure 6.4. Decision-making accuracy in relation to mean minute ventilation (V̇E) in the 10-s (A) and 60-s (B) preceding the decision.

6.4.4. Decision-Making Accuracy in Relation to Differential RPE

A significant association was found between decision-making accuracy and RPE-B ($\chi^2_{(3)}$ = 9.773; P = 0.021; ES = 0.287), with significantly more incorrect decisions observed when RPE-B was rated "*very strong*" to "*maximal*" (Figure 6.5). No association was observed between decision-making accuracy and RPE-M ($\chi^2_{(3)}$ = 6.296; P = 0.098; ES = 0.225).

6.4.5. Decision-Making Accuracy in Relation to the Activity Performed Prior to Decision

A significant association was found between decision-making accuracy and mean running speed in the 10-s preceding a decision ($\chi^2_{(3)}$ = 13.651; P = 0.003; ES = 0.320), with significantly more incorrect decisions observed when running speeds were ≥250 m·min^{-1} (SR, 2.3; OR, 3.84; 95% CI, 1.53–9.60) (Figure 6.6). Decision-making accuracy was also significantly associated with mean running speed in the 60-s preceding a decision ($\chi^2_{(3)}$ = 8.301, P = 0.016, ES = 0.253).

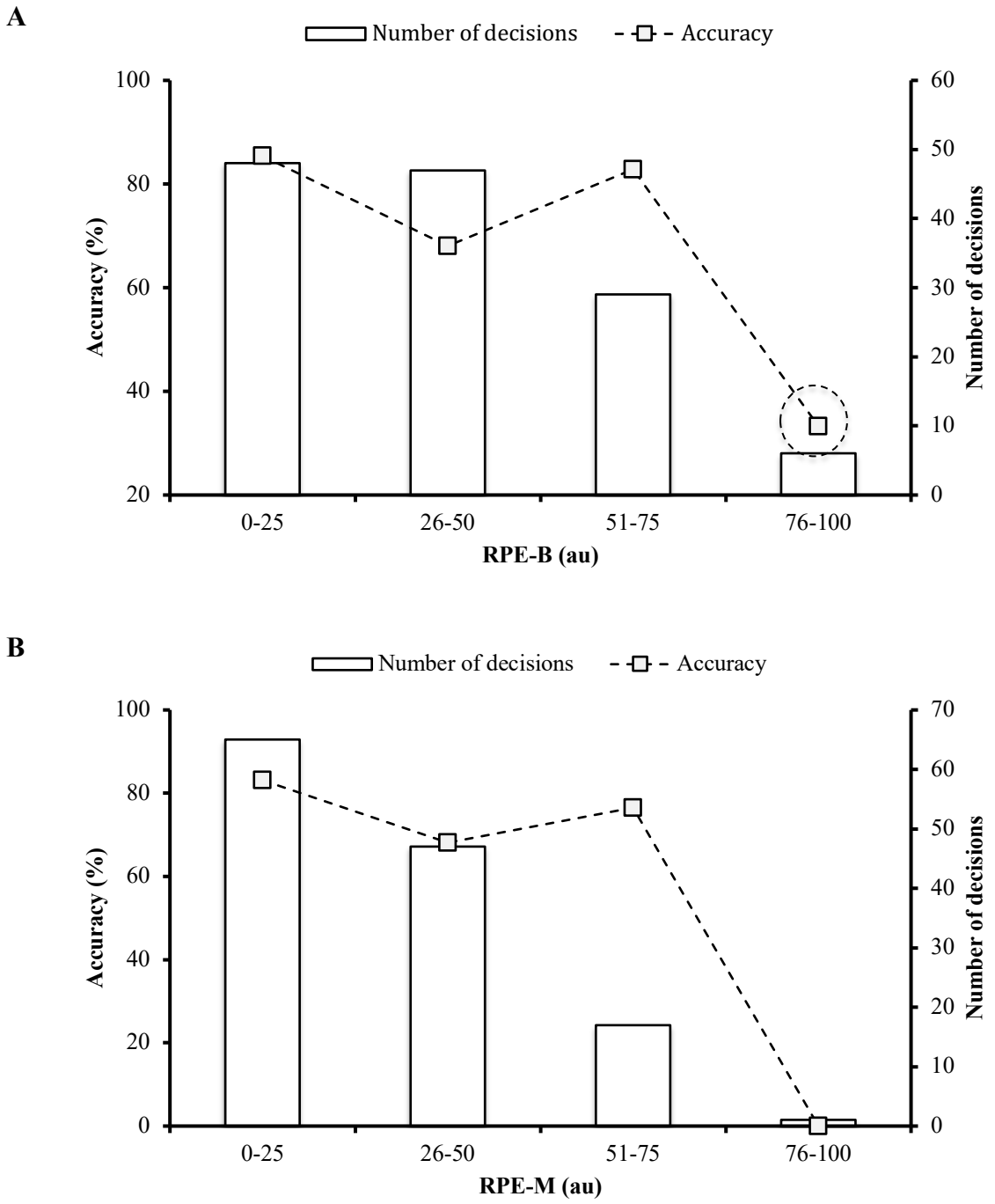


Figure 6.5. Decision-making accuracy in relation to perceptions of central (RPE-B; A) and local muscular (RPE-M; B) exertion. Dashed circle signifies that significantly more incorrect decisions were observed when RPE-B was rated as 76-100 au.

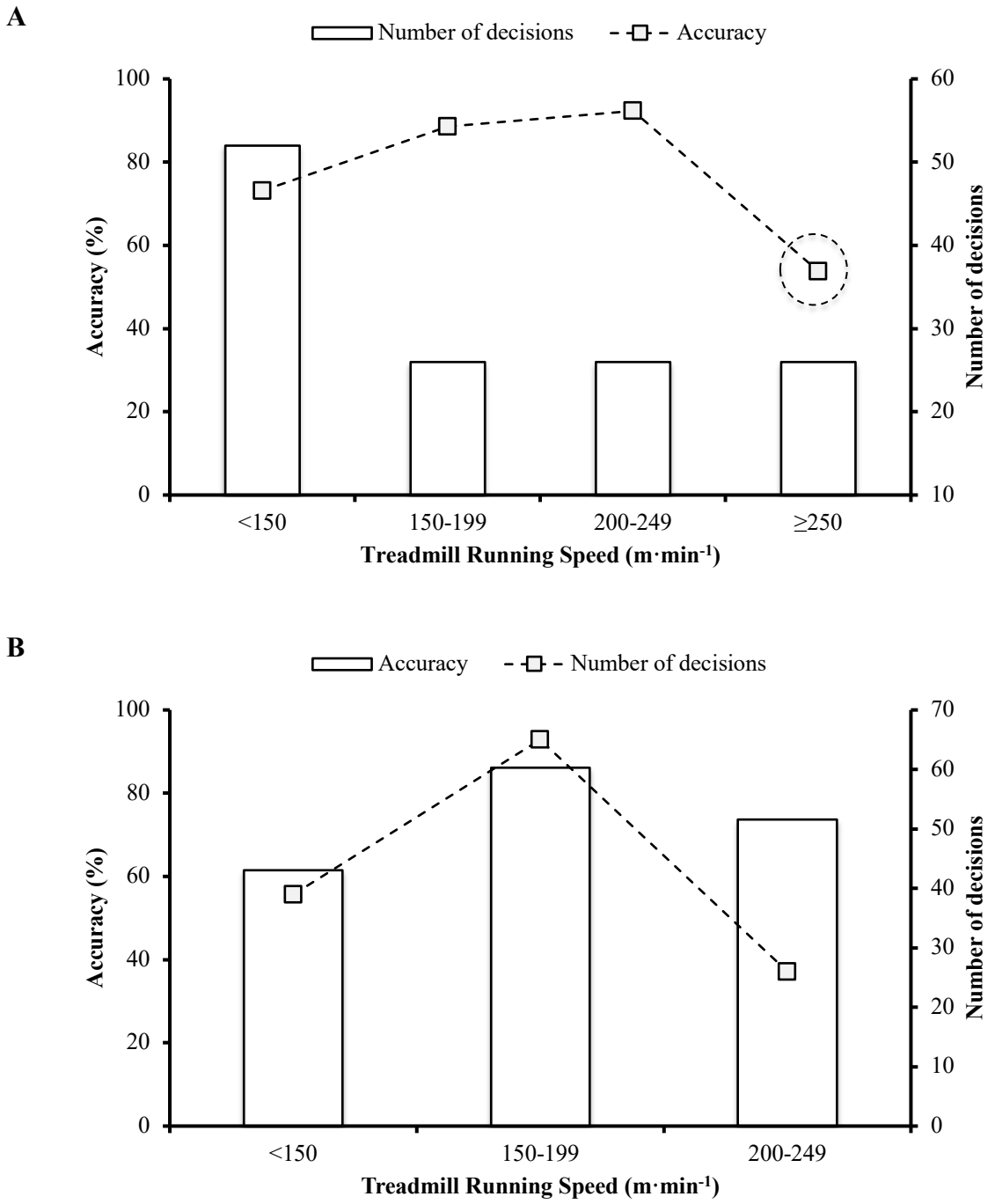


Figure 6.6. Decision-making accuracy in relation to the external workloads performed in the 10-s (A) and 60-s (B) preceding the decision. Dashed circle signifies that significantly more incorrect decisions were observed when running speeds were ≥250 m·min^{-1}.

6.5. DISCUSSION

Utilising the SRS protocol developed in Chapter 5, the present study explored the association between various markers of internal and external load and the accuracy of foul play decisions of a cohort of sub-elite soccer FR. In contrast to previous studies that have utilised prolonged fixed-time epochs to examine the decision-making and physical performances of match officials, a more sensitive analysis was performed in the current investigation whereby the correctness of each decision was assessed in relation to the physical and physiological demands imposed during the 10-s and 60-s preceding the decision. Collectively, the current data offer novel insights into the detrimental effects that high workloads may have upon the decision-making performances of soccer FR. Specifically, officials appeared more likely to make an incorrect decision when: 1) HR ≥90% of HR_{max}; 2) RR ≥80% of RR_{peak}; 3) RPE-B was rated as *very strong* to *maximal* (75–100 au); and 4) running speeds ≥250 $m \cdot min^{-1}$ were performed. Such information may be used to inform the design and delivery of training programmes aimed at preparing soccer officials for the periods of match play that prove most problematic to their decision-making.

As per the findings of the present study, the accuracy of participants' decisions was not related to the mean HR observed in the 60-s prior to the decision. When explored in relation to the HR exhibited in the 10-s prior, however, officials were 5.4 times more likely to make a decisional error when HR ≥90% of HR_{max}. These observations are in contrast to those of previous literature whereby the decision-making accuracy of a cohort of Rugby League officials was found to be unrelated to measures of mean and peak HR (Emmonds et al., 2015). Although speculative, such discrepancies may potentially be attributed to differences in the analytical approaches adopted. Whereas the analysis performed by Emmonds and colleagues (2015) involved the aggregation of data over prolonged 10-minute periods, we explored the association between the officials' HR and decision-making accuracy by assessing the

correctness of each decision in relation to the internal loads exhibited in the moments immediately preceding the decision. The relationship between the physiological and decision-making aspects of officiating performance therefore appears to be a transient one, whereby high internal loads experienced at the time of an infringement may compromise the correctness of the decision. As peak HRs of ~95% of HR_{max} are attained during competitive matches (Weston et al., 2006), such findings have important implications for the in-game decision-making performances of soccer FR, particularly in the event whereby the most intense periods of match play are aligned with the requirement to make a decision.

Another novel aspect of the present study concerns the assessment of decision-making accuracy in relation to perceived levels of central and peripheral exertion. When RPE-B was rated as *very strong* to *maximal* (75-100 au), officials were 7.2 times more likely to make a decisional error. As RPE-B is driven by sensory and affective cues such as RR and HR (Borg et al., 2010), these findings are perhaps unsurprising given the reduced decision-making accuracy observed when RR and HR were ≥80% and ≥90% of RR_{peak} and HR_{max}, respectively. Moreover, previous research has indicated sensations of breathlessness to induce heightened levels of anxiety amongst highly trained endurance athletes, with the magnitude of this association particularly prominent at greater levels of ventilation (Faull et al., 2016). Thus, it appears plausible that when subjected to high levels of cardiorespiratory stress and acute breathlessness, attention may have been diverted away from the decision at the critical moment, with officials instead focusing on the stabilisation of their breathing rate (McEwan et al., 2018). Nevertheless, this remains speculative and additional research is required to identify the mechanisms by which high internal loads may compromise the decision-making performances of soccer officials.

In comparison, the correctness of the officials' decisions was not related to measures of RPE-M; however, it is important to acknowledge that only one decision was made when RPE-

M was rated as *very strong* to *maximal*. Indeed, whilst match officiating is typically associated with a greater perceived peripheral demand (Castillo et al., 2017), officials in the current study reported higher levels of perceived respiratory exertion. As RPE-M depends mainly on the strain and exertion in the lower limbs, these findings may partially reflect the exclusion of referee-specific utility movements (i.e., sideways and backwards running) and changes of direction (Hader et al., 2014). Moreover, moderate correlations have previously been reported between RPE-M and GPS-derived measures of external load such as high-speed running (Weston et al., 2015). The reduced RPE-M observed in the current study may therefore simply be a result of the shorter duration of the SRS protocol in comparison to match play. Whilst measures of HR remain relatively stable across each 15-minute period of match play (Krustrup et al., 2009), levels of neuromuscular fatigue may develop progressively over the course of a match (Goodall et al., 2017). While it was beyond the scope of the present study, future research may wish to explore the effect that match-related fatigue has on the decision-making performances of soccer officials, particularly towards the latter stages of match play.

Another key finding from our data concerns the significant associations observed between decision-making accuracy and the mean running speeds performed in the 10-s and 60-s prior to a decision. Specifically, officials were 3.8 times more likely to make a decisional error having performed external workloads of ≥250 $m \cdot min^{-1}$ in the 10-s preceding the decision. These findings are consistent with those of Elsworthy and colleagues (2014a) who noted higher running speeds performed by Australian Football umpires in the 5-s before a decision to have increased the likelihood of an error being made. Conversely, the correctness of the foul play decisions made by 11 Norwegian soccer FR appeared unrelated to the total distances covered during the 10-s and 30-s before a decision (Riiser et al., 2019). Nonetheless, it is important to acknowledge that only 6 of the foul play decisions made by these officials were deemed incorrect. This represents an error rate (1.7%) that is substantially lower than the 14.2%

previously reported amongst elite FIFA-listed FR (Mallo et al., 2012). Although the reasons for this low error rate remain unclear, the failure of Riiser and colleagues (2019) to observe any associations between the physical and decision-making performances of their cohort may therefore simply indicate a lack of statistical power.

There were, however, a number of limitations in the present study that should be acknowledged. Most important is the performance of a simulated period of match officiating and the reduced level of ecological validity. Although the physiological responses elicited during the SRS are comparable to those observed during competitive matches (Chapter 5), the externally regulated nature of the SRS impeded the officials' ability to express maximal sprinting velocities. The impact of isolated or repeated bouts of maximal sprinting on the decision-making performances of soccer officials therefore remains unclear. Nonetheless, only ~2% of the total distance covered by elite FR is by means of sprinting, with ~19% instead covered by high-speed running (≥18 $km.h^{-1}$) (Krustrup et al., 2009). Another limitation relates to the relatively small sample of pre-selected video clips; however, it is worth noting that the frequency in which clips were presented remained in line with the relative number of fouls that occur during match play (Weston et al., 2011b; Mallo et al., 2012). Additionally, video-based testing has previously been reported to possess high levels of construct and discriminant validity amongst team sport officials (Spitz et al., 2018b). Whilst enhanced levels of ecological validity may be achieved within naturalistic environments, the approach of the present study allowed for the physical and decision-making performances of soccer officials to be assessed in isolation from contextual match factors (i.e., crowd noise) that may further exacerbate decision-making accuracy. Finally, as the officials in the current study were of a sub-elite level, future research is required to confirm whether our findings are reproducible within elite populations.

6.5.1. Conclusions and Perspective

The present findings have demonstrated for the first time that the decision-making performances of soccer FR may be compromised when decisions are made in the presence of high levels of cardiorespiratory stress. Specifically, officials were found to be significantly more likely to make a decisional error when their HR was ≥90% of HR_{max} in the 10-s preceding the decision and when their RR was ≥80% of RR_{peak}. A disproportionate number of errors were also made when RPE-B was rated as *very strong* to *maximal* (75–100 au) and when performing running speeds of ≥250 $m \cdot min^{-1}$ in the 10-s prior to the decision. These observations therefore reaffirm the importance of high levels of aerobic fitness amongst soccer FR (Castagna et al., 2019). Indeed, whilst well-developed physical capacities enable soccer officials to keep up with play and adopt suitable viewing positions, our findings suggest that high levels of physical fitness may also be necessary if officials are to cope with or avoid the sharp increases in physiological stress that may impede their optimal decision-making. It therefore stands that soccer officials should continue to apportion ample time to their physical conditioning. However, in considering the direct impact that acute periods of high physiological stress had on the correctness of the officials' decisions, it would also appear important that officials prepare for such challenges during training (i.e., the requirement to undertake complex perceptual-cognitive processes in the presence of elevated levels of physiological stress). Yet, as highlighted within Chapter 4, the physical and decision-making abilities of soccer officials are typically developed in isolation. With this in mind, practitioners may wish to use the present data to inform the design of combined physical and decision-making training sessions that help to prepare soccer officials for the periods of match play that prove most problematic to their accurate decision-making.

Chapter 7

Synthesis of Findings

7.1 INTRODUCTION

The purpose of the following chapter is to provide an overview of the experimental findings of this thesis in relation to the original aims and objectives outlined in Chapter 1. A general discussion will then follow, with the key findings contextualised in relation to the extant literature and their contribution to this body of knowledge. Thereafter, the limitations of the thesis will be identified from which recommendations for future research will be suggested. The chapter will then culminate with the discussion of the practical implications of the findings.

7.2 EVALUATION OF AIMS AND OBJECTIVES

The overarching aims of this thesis were to: 1) develop valid and reliable measurement tools that could assist in the multidisciplinary testing and evaluation of soccer officials; and 2) use these newly developed methods to collect information from which to inform and guide the future training prescription of soccer officials. These aims were achieved through the completion of four separate but interrelated studies (Chapters 3, 4, 5, and 6) that addressed the following objectives.

Objective 1: To develop and validate a novel measurement tool for assessing the training practices of soccer officials *(Chapter 3, Study 1)*.

Prior to exploring the training practices of soccer officials, it was necessary to first establish a method of collecting information pertaining to the broad range of activities unique to this cohort (i.e., specific physical, decision-making, psychological, and technical skills training). The purpose of Chapter 3 was therefore to develop and validate a self-report questionnaire that could accurately capture training information. To achieve the above objective, a purposeful sample of soccer officials (both FR and AR) and other refereeing experts were consulted and engaged in a systematic, multi-stage process encompassing: 1) item generation and

questionnaire development; 2) assessments of content and face validity; and 3) assessments of criterion validity. The resulting product was the RTAQ; a self-report questionnaire that presents good levels of content, face, and criterion validity. Collectively, the findings of Chapter 3 highlighted the RTAQ to be a valid and effective method of acquiring insight into the training practices of soccer officials and provided evidence to support its use within the subsequent chapter.

Objective 2: To document the current training practices of soccer officials in relation to their roles (FR and AR) and level of professional attainment. An additional objective was to explore the association between the perceptions and training practices of soccer officials *(Chapter 4, Study 2)*.

Following the development and validation of the RTAQ within Chapter 3, Chapter 4 sought to explore the types and volumes of training engaged in by soccer officials. The data presented broadly highlight that soccer officials, irrespective of their role or level of professional attainment, largely focus on physical conditioning with little attention directed towards other facets of performance. Perhaps most notable were the low volumes of decision-making training reported by the respondents, with the officials' exposure to combined physical and decision-making training found to be limited. Such observations, however, did not reflect the perceptions of the officials, with decision-making skills and attributes rated as *"very important"* to *"extremely important"*. Though speculative, disparities between the practices and perceptions of these officials may highlight the presence of potential barriers to this form of training. An important avenue for future research therefore concerns the development of effective and accessible decision-making assessment and training methods for soccer officials.

Objective 3: To develop a valid and reliable treadmill-based simulation that replicates the physiological and decision-making demands of soccer officiating *(Chapter 5, Study 3)*.

As revealed within Chapter 4, the exposure of soccer officials to match-related decision-making was largely limited to video-clip analyses, with little decision-making training performed under elevated levels of physiological stress. To address this, a need therefore existed for the development of a simulation protocol that combined match-related decision-making with the physiological loads commensurate of match play. Accordingly, the purpose of Chapter 5 was to assess the validity and reliability of a novel treadmill-based Soccer Referee Simulation (SRS). In relation to the validity of the physiological responses elicited during the SRS, no differences were detected between the SRS and actual match play for any of the selected HR measures, whilst the physiological and perceptual responses recorded were also aligned with those observed previously amongst elite soccer FR within the literature. The SRS was also found to yield highly reproducible physiological and perceptual responses, with good levels of test-retest reliability observed for all outcome measures. The data presented within Chapter 5 have therefore demonstrated the SRS to be a valid and reliable protocol that closely replicates the physiological and perceptual responses elicited by soccer FR during match play.

Objective 4: To assess the decision-making performances of soccer officials in relation to the internal and external loads recorded in the moments immediately preceding decisions *(Chapter 6, Study 4)*.

As outlined in Chapter 2, a key question yet to be answered by the literature concerns the impact that enhanced physiological loads have on the decision-making performances of soccer officials. Chapter 6 therefore sought to shed further light on this association through the utility of the newly developed SRS. To this end, 13 soccer FR completed the SRS, with the accuracy of their decisions subsequently assessed in relation to the internal and external loads recorded

in the 10- and 60-s preceding each decision. Utilising this approach, it is shown for the first time that the accuracy of the decisions made by soccer FR may be compromised when these decisions are made during (or shortly after) periods of extremely high workloads. Specifically, soccer FR were found to be more likely to make a decisional error when: 1) HR ≥90% of HR_{max}; 2) RR ≥80% of RR_{peak}; 3) RPE-B was rated as *very strong* to *maximal* (75–100 au); and 4) running speeds ≥250 $m \cdot min^{-1}$ were performed. Such information may be used to inform the design and delivery of future training programmes aimed at preparing soccer FR for the periods of match play that prove most problematic to their decision-making.

7.3 GENERAL DISCUSSION

Soccer officials are tasked with ensuring that match play is contested in accordance with the laws of the game (Helsen & Bultynck, 2004). Whilst such objectives appear to be relatively straightforward, the highly pressurised and physically demanding conditions that characterise match play render the officials' task an immensely difficult one. In considering the implications that such demands may have on the decision-making process, researchers have explored the factors that facilitate (or inhibit) accurate decision-making, thereby offering important insights into the multidimensional nature of soccer officiating. This notwithstanding, previous literature has largely explored the physiological (Castagna et al., 2007; Weston et al., 2012), psychological (Samuel, 2015), and perceptual-cognitive (Helsen et al., 2019) demands in isolation. As contended throughout this thesis, however, such approaches do not reflect the realities of competitive match play whereby these demands are not independent but are instead intricately linked. It therefore follows that each should be considered in unison when evaluating the training and performances of soccer officials. In this regard, the purpose of the present thesis was to: 1) develop valid and reliable measurement tools that could assist in the multidisciplinary testing and evaluation of soccer officials; and 2) to use these newly developed

tools to provide important insights into their current training habits and the factors influencing their decision-making performances. Knowledge regarding current training practices and the factors impeding their accurate decision-making may then be used to help guide and inform the training processes of soccer officials.

7.3.1. Development and Validation of Multidisciplinary Measurement Tools

As revealed through the systematic review of the literature presented within Chapter 2, the decision-making performances of soccer officials are influenced by a myriad of factors relating not only to the individual competencies of the official, but to the infinite number of environmental and contextual variables that confound match decisions. To cope with the breadth of demands encountered during match play, soccer officials therefore require a multifaceted skill profile that comprises well-developed physical, psychological, perceptual-cognitive, and technical attributes (Castagna et al., 2007; Slack et al., 2013; Spitz et al., 2018b). Central to the development of the requisite skills and expertise lies the training process whereby attention should be given to developing each of the cornerstones of performance simultaneously (Weston et al., 2012). In failing to do so, soccer officials may be unable to fulfil their duties on the field of play. Thus, with the intention of identifying potential shortcomings and areas to address in the subsequent chapters, a key objective of the present thesis was to the training practices currently engaged in by Scottish soccer officials. As highlighted by the applied research model put forward by Bishop (2008), descriptive studies represent a fundamental part of the research process, allowing researchers to firstly define the problem and also understand the contextual backdrop against which their research is to be conducted. That is, by assessing the practices and perceptions of those currently working within the field, researchers can ensure that their research efforts are directed towards solving real world problems and improving performance (Bishop, 2008; Drust and Green, 2013). As

outlined within Chapter 3, however, it was necessary to first establish a valid method of collecting information pertaining to the broad range of training activities unique to this cohort. Although physical training activities may be assessed objectively using HR and GPS monitors, there was no standardised method of ascertaining the training activities of soccer officials that are invariably self-led and performed off-field (i.e., decision-making, psychological, and technical skills training). Accordingly, Chapter 3 sought to develop and validate a standardised measurement tool that could be used to monitor the engagement of soccer officials in these crucial yet often overlooked forms of training.

Arising from the work presented in Chapter 3 was the RTAQ – the first self-report questionnaire to be developed and validated to explore the training practices of soccer officials (McEwan et al., 2020). The successful implementation of measurement tools within applied settings is subject to the specificity to the end-users (Bourdon et al., 2017), so direct input was received from soccer officials and other refereeing experts during both the development and validation of the RTAQ. By consulting directly with the end-users and key stakeholders during its initial development, it was possible to identify the training practices that are unique to this cohort, thereby ensuring that the scope of the RTAQ reflected each of the cornerstones of performance identified within Chapter 2. Further, whilst the utility of self-report measures is common within both research and applied settings (Saw et al., 2017), their usefulness depends upon the exhibition of acceptable psychometric properties (Robertson et al., 2014). Amongst the properties that should be established prior to the implementation of such tools into practice are the levels of content, face, and criterion validity (Robertson et al., 2014). In this regard, it is encouraging to note that the consensus received amongst the officials and expert panel provided support for both the content and face validity of the RTAQ. Meanwhile, the RTAQ also satisfied the criteria of criterion validity, with good levels of agreement observed between the RTAQ and 7-day training diaries. Thus, given the previous absence of a valid and

accessible method of exploring the multifaceted training practices of soccer officials, the development of the RTAQ represents an important and welcomed addition to the literature.

In considering the significant physiological demands associated with match play, another key question raised within Chapter 2 relates to the impact that such loads have on the accuracy of the officials' decisions. Intuitively, the necessity to perform complex perceptual-cognitive processes under high levels of physical exertion would appear counterproductive to optimal decision-making (Hüttermann & Memmert, 2014; Smith et al., 2016). It also stands to reason that the dual-task nature of soccer officiating, whereby the physical and decision-making tasks of the role are imposed concurrently, may place an additional strain on the officials' limited attentional resources (Schmidt et al., 2019). Nonetheless, whilst the association between the physical and decision-making performances of team sport officials has been explored, findings have proven equivocal with a clear association yet to be established (Mascarenhas et al., 2009; Elsworthy et al., 2014b; Emmonds et al., 2015). As postured within Chapter 2, however, such equivocality may reflect the challenges of exploring the association between physiological load and decision-making in-situ. For instance, as match play data is susceptible to a host of environmental and contextual variables (Weston et al., 2006; Castillo et al., 2018), delineating the influence of physical load on the in-match decisions of soccer officials is difficult. To circumvent the challenges of conducting research within the actual performance setting, simulation protocols that mimic match play are commonly employed within controlled laboratory conditions, with several protocols already existing for soccer players (Russell et al., 2011; Aldous et al., 2014; Page et al., 2015). Whilst negating the influence of the contextual variables that confound match data, such protocols also ensure the standardisation of the physical stimulus and facilitate the more accessible monitoring of physiological and perceptual responses. Prior to the development of the SRS within Chapter 5, however, a valid and reliable

laboratory-based model that replicated both the physiological and decision-making demands of match play did not exist for soccer officials.

Considering first the fidelity of the external stimulus, the SRS was modelled to mimic the stochastic and non-uniform nature of an elite soccer FRs activity profile during match play, with the frequency, duration, and speed of each discrete locomotive phase reflecting previously published data (Krustrup et al., 2009; Barbero-Alvarez et al., 2012). In contrast to a recent simulation that involved alternating periods of moderate-intensity running (Samuel et al., 2019), the SRS incorporated the frequent bouts of both low- and high-speed running that characterise match play. Whilst the absence of changes of direction and utility movements have cast doubt on the ecological validity of treadmill-based simulations (Aldous et al., 2014), incorporating frequent activity changes of high intensities into treadmill protocols can elicit biomechanical responses comparable to match play (Page et al., 2015). The modelling of the SRS to incorporate these brief but frequent bouts of high-speed running was therefore critically important and ensured that the SRS provided a valid representation of the running demands elicited by elite soccer FR during competitive match play.

Results from Chapter 5 also revealed the physiological and perceptual responses elicited during the SRS to be both valid and reliable. In relation to the selected HR measures, no differences were detected between the SRS and match play, although it is acknowledged that the matches examined for comparative purposes were from the second and third tier of the Scottish Professional Football League, and not the Premier division. This notwithstanding, the physiological and perceptual responses elicited during the SRS remained largely consistent with those observed amongst FR during elite-level matches (D'Ottavio & Castagna, 2002; Weston et al., 2006; Castillo et al., 2016b). Meanwhile, such responses were also found to be highly reproducible. When examined across repeated trials, no systematic differences were detected for any of the measured outcome variables ($P \geq 0.288$). Excellent levels of relative

reliability were also demonstrated for all physiological and perceptual responses (ICC≥0.891), with repeat measures of mean HR, peak HR, and $\dot{V}O_2$ exhibiting good levels of absolute reliability (CV≤3.1%). Such findings were key in considering the potential for the SRS to be used within Chapter 6, and indeed within future research more generally. As previously noted, the physical match performances of soccer officials are subject to large levels of match-to-match variation, with such variability often precluding the detection of real systematic changes in outcome variables when explored *in-situ* (Gregson et al., 2010; Weston et al., 2011b). The development of the SRS therefore constitutes a significant contribution to the literature and now provides researchers with a valid and reliable protocol in which to simulate the physical demands imposed upon elite FR during match play.

In keeping with the overarching theme of the current programme of research, it was important that the SRS reflected not only the physical demands of match play, but the requirement to make match-related decisions under varying levels of physiological stress. With this in mind, a unique feature that differentiates the SRS from existing team sport simulations concerns the inclusion of video-based decision-making. Video-based testing has been used extensively within both research and off-field training settings and has been found to possess high levels of construct validity and reliability (Larkin et al., 2014a; Spitz et al., 2017; Kittel et al., 2019c). Concerns have however been raised over the fidelity and ecological validity of existing methods as they often use match broadcast footage that is presented from a third-person or exocentric perspective (Kittel et al., 2019c). As such perspectives do not replicate those experienced *in-situ*, simulation protocols that have utilised match broadcast footage have been deemed only moderately representative of match play (Samuel et al., 2019). Alternatively, it has been suggested that emerging virtual reality (VR) technology may offer a potential solution to the aforementioned limitations (Craig, 2013). Whilst presenting footage from a first-person perspective, VR enables viewers to immerse themselves within and interact

with the environment, thus enhancing the representativeness of the decision-making task (Bird, 2019). Recent studies exploring the potential application of VR amongst amateur Australian Football umpires have found 360° VR to be rated as significantly more game-like than traditional match broadcast footage (Kittel et al., 2019b; Kittel et al., 2020). Nonetheless, the fidelity of 360° VR was not supported by elite umpires, with the levels of perceived concentration and effort associated with this method proving similar to traditional match broadcast footage (Kittel et al., 2019b; Kittel et al., 2020). Thus, despite the promise of VR to enhance the visual correspondence of the decision-making task, the ecological validity of current methods remains limited. As the aforementioned investigations were conducted within resting conditions, it is possible that the reduced levels of fidelity may be an artefact of the decision-making stimuli being applied in isolation from the physical challenges that are present during match play. When considering the officials' perceptions within Chapter 5, all participants suggested the SRS to be game-like, with 63% of participants stating that the SRS required higher levels of concentration in comparison to traditional match broadcast footage. Although direct comparisons are difficult to make between these findings and those of previous literature, it does appear that the dual-task nature of the SRS had a positive effect upon its levels of ecological validity. Nonetheless, additional research is required to further explore the fidelity of the SRS in relation to match play and other training methodologies.

7.3.2. Insights for the Training and Development of Soccer Officials

Expanding on the work presented in Chapters 3 and 5, the following chapters utilised the newly developed tools to foster a greater understanding of the training and performances of soccer officials. It was anticipated that such information could help inform the training prescription of soccer officials, as well as guide the direction of future research and practice (Bishop, 2008). In Chapter 4, for instance, data collected following the administration of the

RTAQ provided an extensive and up-to-date account of the types and volumes of training engaged in by Scottish soccer officials. In confirmation of previous reports (MacMahon et al., 2007; Catteeuw et al., 2009b), the training and development of respondents was largely focused on physical conditioning, with relatively little time apportioned to decision-making training. Though high levels of physical fitness are indeed necessary to cope with the growing intensity of the modern game, it is the officials' ability to make consistently fair and accurate decisions that underpins their success. It therefore stands to reason that greater attention should be afforded to developing the requisite decision-making and technical skills. With respect to the specific decision-making and technical training engaged in by respondents, these appeared to be comprised almost exclusively of observational learning activities such as video clip and game analyses, with practical decision-making experience being limited to official matches. Considering the finite number of matches available across an official's career, the implications of these findings are significant and likely explain why soccer officials reach elite status at a later stage in life as compared to their playing counterparts (Catteeuw et al., 2009b; Weston et al., 2010). That is, whereas soccer players accumulate large volumes of match-specific practice from an early age (Ford et al., 2020), the officials are presented with limited opportunities to engage in match-related decision-making, therefore prolonging and delaying their development. If researchers and practitioners are to improve the performances of soccer officials and accelerate their development, efforts should be made to create additional opportunities for officials to engage in match-related decision-making training.

Another unique insight gleaned from Chapter 4 concerns the relative dissociation that was present between the officials' perceptions and their actual training practices. That is, in spite of the low volumes of decision-making training accrued by the officials, such skills were acknowledged as being *"very important"* to *"extremely important"* to successful performance. Although the exact reasons for these findings are not clear, these observations could indicate

potential barriers to training such as a lack of time or the absence of effective and accessible training methods. As identified within Chapter 2, a key challenge facing soccer officials is the requirement to develop expertise within several areas of performance simultaneously. Naturalistic training methodologies that allow decision-making abilities to be developed in tandem with the officials' physical and psychological capacities are therefore warranted. Not only would such methods be more time-efficient, but they would also provide a closer reflection of the multidimensional nature of match play. Small-sided games represent a pertinent example of such training and are commonly employed amongst players to develop their technical, tactical, and physical abilities (Light et al., 2014). In contrast, holistic training methods are currently lacking for the officials. With this in mind, the SRS developed in Chapter 5 may be considered as a potential training tool that could enhance the exposure of soccer officials to match-related decision-making under game-like conditions. In addition to reflecting the intensity of elite matches, the decision-making stimuli of the SRS remained consistent with the number of fouls that occur during match play (Weston et al., 2011b; Mallo et al., 2012). Meanwhile, the SRS was appraised positively by all participants and was deemed to be a relevant and enjoyable protocol that necessitated heightened levels of concentration – characteristics that are considered central to deliberate practice (Ericsson et al., 1993). It is worth noting however that the decision-making load associated with the SRS is lower than that of previous video-based training interventions that have proven successful in enhancing the decision-making abilities of team sport officials (Schweizer et al., 2011; Larkin et al., 2017; Kittel et al., 2020). Although the SRS represents an important first step towards developing a naturalistic dual-task protocol for soccer officials, additional research is required to examine the efficacy of the SRS as a training tool for improving their decision-making skills.

To ensure that soccer players are adequately prepared for the peak physical demands of competition, efforts have been made within the literature to identify the most demanding

sequences of play (Oliva-Lozano et al., 2021). In doing so, it is suggested that practitioners may be able to design bespoke training drills that replicate and prepare players for these "worst-case scenarios" (WCS). Such research has not yet been conducted with soccer officials and would indeed yield valuable information from which to guide the prescription of physical conditioning programmes. In translating the concept of WCS to soccer officials, however, a more pertinent objective would be to identify the physical conditions under which the officials' decision-making is compromised most. The purpose of Chapter 6 was therefore to investigate the impact that heightened levels of physiological stress have on the decision-making performances of soccer FR. However, in contrast to previous studies whereby the physical and decision-making performances of team sport officials have been explored across fixed 10–20-minute epochs (Mascarenhas et al., 2009; Elsworthy et al., 2014b; Emmonds et al., 2015), a more sensitive analysis was performed with the correctness of each decision assessed in relation to the physiological loads recorded in the moments immediately preceding that decision. Utilising this approach, the current findings demonstrate for the first time that high levels of cardiorespiratory stress may impede the accurate decision-making of soccer FR. Specifically, officials were 5.4 times more likely to make a decisional error when their HR was ≥90% of HR_{max} in the 10-s preceding the decision, and 3.3 times more likely to make a mistake when RR was ≥80% of RR_{peak}. A disproportionate number of errors were also made when RPE-B was rated as *very strong* to *maximal* (75–100 au). Conversely, the accuracy of the officials' decisions was unrelated to the cardiorespiratory loads recorded in the 60-s preceding the decision. The relationship mediating the physiological and decision-making aspects of soccer officiating therefore appears to be a transient one, whereby high internal loads experienced at the time of the decision may compromise the correctness of the decision.

Consistent with the findings of Elsworthy et al (2014a), higher relative running speeds performed prior to a decision were found to increase the likelihood of the official making an

incorrect decision. Specifically, officials were 3.8 times more likely to make an error when running speeds of ≥250 m·min^{-1} were performed in the 10-s prior to the decision. As intense activities such as high-speed running frequently precede decisive moments during match play (Faude et al., 2012), these findings are likely to have important consequences for the in-situ decision-making performances of soccer officials. It is therefore important that officials not only possess the physical qualities necessary to keep up with play, but that they are accustomed to making decisions during (or shortly after) acute bouts of high-speed running. In this regard, the present data may provide coaches and sport scientists with valuable information from which to design training sessions that prepare officials for the types of situations and contexts that hinder their decision-making most. Although the use of WCS as thresholds for physical conditioning sessions has recently been questioned (Novak et al., 2021), there may still be value in training decision-making under the conditions that prove most problematic. The premise being that by developing decision-making skills in tandem with high physical loads, officials may become accustomed to making decisions under such conditions and therefore desensitised to the distracting processes that are believed to hinder performance. This notwithstanding, a recent investigation of Australian football umpires found there to be no additional benefit of incorporating decision-making into high-intensity interval training (HIIT) sessions, as compared to the same decision-making training programme performed at rest (Kittel et al., 2019a). There does however exist a number of methodological limitations that warrant consideration when interpreting these findings. Firstly, the intervention period was relatively short and consisted of only one session per week for 8 weeks. It is therefore possible that the duration of the intervention was simply inadequate to promote any improvements in decision-making performance. Meanwhile, the intensity of the HIIT sessions undertaken remains unclear as no measures of physiological load were recorded during this study. As we have observed decision-making accuracy to be affected only when RR and HR exceeded 80%

and 90% of RR_{peak} and HR_{max}, respectively, it stands to reason that such conditions should be replicated during such sessions. Additional research is therefore required to examine the efficacy of incorporating decision-making into the physical conditioning sessions of team sport officials.

7.4 LIMITATIONS AND FUTURE DIRECTIONS

The findings of the work discussed in this thesis have made an important contribution to the existing literature surrounding the training and decision-making performances of soccer officials. This notwithstanding, there exist a number of limitations that warrant consideration. As the limitations associated with each of the four empirical studies have been outlined within their respective chapters, the following discussion will focus on those limitations that are applicable across the programme of research as a whole. In considering such limitations, recommendations for future research are suggested.

7.4.1. Sample Sizes

The limitations of the present thesis should be considered with respect to the challenges encountered when conducting research of this nature. Firstly, as a consequence of the limited availability and access to elite athletes, studies undertaken within the high-performance sport setting are often subject to small sample sizes. Such challenges were also encountered during the current body of work as there existed a finite pool from which to recruit elite soccer officials. As noted within Chapter 3, for example, a relatively small number of participants were consulted during each stage of the development and validation of the RTAQ. Not only did these sample sizes reflect the naturally limited number of potential participants to begin with, but the requirement to involve different officials during each distinct stage in order to avoid participant fatigue or bias. Similarly, as Chapter 4 sought to document the training practices

of Scottish officials at Category 3 and above, the number of potential responses was limited to the officials listed within these categories. Nonetheless, the response rate (and indeed the total number of responses received) was in line with similar surveys that have explored the perceptions and practices of those working within elite soccer (Towlson et al., 2013; Akenhead & Nassis, 2016; Harper et al., 2016a; McCunn et al., 2018). It is also important to acknowledge that opportunities to conduct empirical research during the in-season are often limited as athletes (and their coaches) may be reluctant to participate in investigations that interfere with their training and performances. A relatively small number of officials were therefore available to participate in Chapters 5 and 6 as a result. Thus, although the sample sizes employed during each study are consistent with those of previous literature and reflect the applied nature of the work, the findings of the present thesis warrant confirmation within larger scale studies that possess greater levels of statistical power.

7.4.2. Generalisability of Findings

Another limitation of this thesis relates to the ability to extend or generalise the findings to other contexts. Indeed, whilst Chapter 4 offers a comprehensive overview of the training practices employed by Scottish officials, readers should remain mindful that these findings reflect the practices and perceptions of match officials from a single football association. Whether these observations reflect the practices of soccer officials from other countries with different sporting cultures remains unclear and warrants further investigation. Meanwhile, as a result of our limited access to soccer officials and the current landscape of elite officiating within Scotland, the cohorts utilised within each of the four investigations were exclusively male. While the training and performance demands placed upon male and female officials are perhaps comparable (Mallo et al., 2010), additional research is required to replicate the findings of this thesis with female officials. It must also be acknowledged that as the SRS has been

designed to replicate the physical and physiological loads imposed upon elite FR during match play, extending the findings of Chapters 5 and 6 to AR would be inappropriate. Lastly, the findings of Chapter 6 should be interpreted in relation to the sub-elite population from which they were derived. Future research may therefore wish to build upon these findings by confirming whether similar observations are observed amongst elite soccer FR.

7.5 PRACTICAL IMPLICATIONS

In addition to the development of novel measurement tools that may permit a more multidisciplinary approach to be taken to the testing and evaluation of soccer officials, the present thesis has offered novel insights into their current training practices and the physiological conditions that may impede their accurate decision-making. Thus, the work undertaken in the present thesis possess several practical implications that may be important for future research and practice.

1. The synthesis of the literature presented within Chapter 2 provides a comprehensive overview of the multifaceted aspects of soccer officiating and offers important insights into how each of these interact to impact the decision-making performances of soccer officials. This work therefore provides researchers and practitioners with a useful framework to consider when conceptualising future research studies and training interventions.
2. As evidenced within Chapter 3, the RTAQ has been found to be a valid and accessible method of monitoring the engagement of soccer officials in the wide range of training activities pertinent to performance. Coaches and sport scientists working with soccer officials may therefore wish to implement the RTAQ as part of their existing

monitoring strategies. In doing so, practitioners will be better placed to evaluate current practices and to intervene where necessary.

3. As per the findings of Chapter 4, soccer officials currently engage in low volumes of decision-making training, with their exposure to match-related decision-making being largely limited to competitive matches. Thus, if researchers and practitioners are to positively impact the development and performances of soccer officials, efforts should be made to develop naturalistic training methodologies that allow officials to develop their decision-making abilities under game-like conditions.
4. The SRS developed within Chapter 5 has been shown to produce physiological and perceptual responses that are commensurate with those of competitive match play. Several opportunities therefore exist for the SRS to be implemented as part of applied practice. For instance, the SRS may be implemented as part of talent identification or development programmes and used to assess the physical and decision-making abilities of prospective soccer officials. Meanwhile, the SRS could also be implemented as part of the "return-to-play" process where it could be used to assess the match readiness of those officials returning from injury.
5. The reproducibility of the physiological responses produced during the SRS also lends support to its use within research settings where factors such as environmental conditions may be manipulated to evaluate their influence on the physical and decision-making performances of soccer officials. The SRS may also be used to assess the efficacy of training or nutritional interventions.
6. The SRS may also hold promise as a potential training tool to develop the decision-making and physical abilities of soccer officials simultaneously. Though further research is required to assess the efficacy of the SRS for improving decision-making,

the SRS could be used to enhance the exposure of officials to the requirement to make decisions in combination with elevated levels of physiological stress.

7. In considering the findings of Chapter 6, it is suggested that future research avoids the use of extended fixed-time epochs when assessing the physical and decision-making performances of soccer officials, and instead considers the use of rolling windows whereby the accuracy of each individual decision is considered in relation to the internal and external loads imposed in the preceding moments.
8. Lastly, the data presented in Chapter 6 may be used to inform the design of combined physical and decision-making training sessions that aim to prepare soccer officials for the periods of match play that present the greatest challenge to their accurate decision-making.

REFERENCES

Abt, G., & Lovell, R. (2009). The use of individualized speed and intensity thresholds for determining the distance run at high-intensity in professional soccer. *Journal of Sports Sciences*, 27(9), 893–898. http://doi.org/10.1080/02640410902998239

Agresti, A. (2007). An introduction to categorical data analysis (2nd ed.). New York, NY: John Wiley & Sons.

Akenhead, R., & Nassis, G. P. (2016). Training load and player monitoring in high-level football: current practice and perceptions. *International Journal of Sports Physiology and Performance,* 11(5), 587–593. http://doi.org/10.1123/ijspp.2015-0331

Alder, D. B., Ford, P. R., Causer, J., & Williams, A. M. (2018). The effect of anxiety on anticipation, allocation of attentional resources, and visual search behaviours. *Human Movement Science*, 61, 81–89. http://doi.org/10.1016/j.humov.2018.07.002

Aldous, J. W. F., Akubat, I., Chrismas, B. C. R., Watkins, S. L., Mauger, A. R., Midgley, A. W., Abt, G., & Taylor, L. (2014). The reliability and validity of a soccer-specific nonmotorised treadmill simulation (Intermittent Soccer Performance Test). *Journal of Strength and Conditioning Research,* 28(7), 1971–1980. https://doi.org/10.1519/JSC.0000000000000310

Allen-Collinson, J. (2009). Sporting embodiment: sports studies and the (continuing) promise of phenomenology. *Qualitative Research in Sport, Exercise, and Health*, 1(3), 279–296. https://doi.org/10.1080/19398440903192340

Aloui, K., Abedelmalek, S., Chtourou, H., Wong, D. P., Boussetta, N., & Souissi, N. (2017). Effects of time-of-day on oxidative stress, cardiovascular parameters, biochemical markers, and hormonal response following level-1 Yo-Yo intermittent recovery test. *Physiology International,* 104(1), 77–90. https://doi.org/10.1556/2060.104.2017.1.6

Anderson, K. J., & Pierce, D. A. (2009). Officiating bias: the effect of foul differential on foul calls in NCAA basketball. *Journal of Sports Sciences*, 27(7), 687–694. https://doi.org/10.1080/02640410902729733

Anderson, L., Orme, P., Di Michele, R., Close, G. L., Morgans, R., Drust, B., & Morton, J. P. (2016). Quantification of training load during one-, two- and three-game week schedules in professional soccer players from the english premier league: Implications for carbohydrate periodization. *Journal of Sports Sciences*, *34*(13), 1250–1259. https://doi.org/10.1080/02640414.2015.1106574

Aragão-Pina, J., Passos, A., Araújo, D., & Maynard, M. T. (2018). Football refereeing: an integrative review. *Psychology of Sport and Exercise*, 35, 10–26. https://doi.org/10.1016/j.psychsport.2017.10.006

Arnold, R., Fletcher, D., & Daniels, K. (2013). Development and validation of the organisational stressor indicator for sport performers. *Journal of Sport and Exercise Psychology*, *35*(2), 180–196. https://doi.org/10.1123/jsep.35.2.180

Artino, A. R., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE guide no. 87. *Medical Teacher*, *36*(6), 463–474. https://doi.org/10.3109/0142159X.2014.889814

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine,* 26(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Baker, J., & Young, B. W. (2014). 20 years later: Deliberate practice and the development of expertise in sport. *International Review of Sport and Exercise Psychology*, *7*(1), 135–157. https://doi.org/10.1080/1750984X.2014.896024

Balmer, N. J., Nevill, A., Lane, A., Ward, P., Williams, A., & Fairclough, S. (2007). Influence of crowd noise on soccer refereeing consistency in soccer. *Journal of Sport Behaviour*, 30(2), 130–145. Available at: https://hdl.handle.net/2436/35332

Banister, E. W. (1991). Modelling athletic performance. In H. J. Green, J. D. McDougal, & H. Wenger (Eds.), Physiological testing of elite athletes (pp. 403–424). Champaign, IL: Human Kinetics.

Barbero-Álvarez, J., Boullosa, D. A., Nakamura, F. Y., Andrín, G., & Castagna, C. (2012). Physical and physiological demands of field and assistant soccer referees during America's cup. *Journal of Strength and Conditioning Research*. 26(5), 1383–1388. https://doi.org/10.1519/JSC.0b013e31825183c5

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

Barte, J. C. M., & Oudejans, R. R. D. (2012). The effects of additional lines on a football field on assistant referees' positioning and offside judgements. *International Journal of Sports Science and Coaching*, 7(3), 481–492. https://doi.org/10.1260/1747-9541.7.3.481

Bishop, D. (2008). An applied research model for the sport sciences. *Sports Medicine,* 38(3), 253–263. https://doi.org/10.2165/00007256-200838030-00005

Bland, J. M., & Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics*, *17*(4), 571–582. https://doi.org/10.1080/10543400701329422

Borg, E., Borg, G., Larsson, K., Letzter, M., & Sundblad, B. M. (2010). An index for breathlessness and leg fatigue. *Scandinavian Journal of Medicine and Science in Sports*, 20(4), 644–650. https://doi.org/10.1111/j.1600-0838.2009.00985.x

Borresen, J., & Lambert, M. (2006). Validity of self-reported training duration. *International Journal of Sports Science & Coaching*, *1*(4), 353–359. https://doi.org/10.1260/174795406779367666

Borresen, J., & Lambert, M. (2009). The quantification of training load, the training response and the effect on performance. *Sports Medicine*, *39*(9), 779–795. https://doi.org/10.2165/11317780-000000000-00000

Bourdon, P. C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M. C., Gabbett, T. J., Coutts, A. J., Burgess, D. J., Gregson, W., & Cable, N. T. (2017). Monitoring training loads: Consensus statement. *International Journal of Sports Physiology and Performance,* 12(S2), S2161–S2170. https://doi.org/10.1123/IJSPP.2017-0208

Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. *Journal of Sports Sciences*, 25(11), 1185–1194. https://doi.org/10.1080/02640410601038576

Brand, R., Schmidt, G., & Schneeloch, Y. (2006). Sequential effects in elite basketball referees' foul decisions: an experimental study on the concept of game management. *Journal of Sport and Exercise Psychology*, 28(1), 93–99. https://doi.org/10.1123/jsep.28.1.93

Brink, M., Kuyvenhoven, J., Toering, T., Jordet, G., & Frencken, W. (2018). What do football coaches want from sport science? *Kinesiology,* 50(S1), 150–154. Available at: https://hrcak.srce.hr/ojs/index.php/kinesiology/article/view/6695

Brown, D. J., & Fletcher, D. (2017). Effects of psychological and psychosocial interventions on sport performance: A meta-analysis. *Sports Medicine*, 47(1), 77–99. https://doi.org/10.1007/s40279-016-0552-7

Buchheit, M. (2016). The numbers will love you back in return – I promise. *International Journal of Sports Physiology and Performance,* 11(4), 551–554. https://doi.org//10.1123/IJSPP.2016-0214

Buchheit, M., Al Haddad, H., Millet, G. P., Lepretre, P. M., Newton, M., & Ahmaidi, S. (2009). Cardiorespiratory and cardiac autonomic responses to 30-15 intermittent fitness test in team sport players. *Journal of Strength and Conditioning Research*, 23(1), 93–100. https://doi.org/10.1519/JSC.0b013e31818b9721

Bush, M., Barnes, C., Archer, D. T., Hogg, B., & Bradley, P. S. (2015). Evolution of match performance parameters for various playing positions in the English Premier League. *Human Movement Science*, 39, 1–11. https://doi.org/10.1016/j.humov.2014.10.003

Byrne, K. A., Silasi-Mansat, C. D., & Worthy, D. A. (2015). Who chokes under pressure? The Big Five personality traits and decision-making under pressure. *Personality and Individual Differences,* 74, 22–28. https://doi.org/10.1016/j.paid.2014.10.009

Carling, C., & Dupont, G. (2011). Are declines in physical performance associated with a reduction in skill-related performance during professional soccer match-play? *Journal of Sports Sciences*, 29(1), 63–71. https://doi.org/10.1080/02640414.2010.521945

Casajus, J. A., & Castagna, C. (2007). Aerobic fitness and field test performance in elite Spanish soccer referees of different ages. *Journal of Science and Medicine in Sport,* 10(6), 382–389. https://doi.org/10.1016/j.jsams.2006.08.004

Castagna, C., Abt, G., & D'Ottavio, S. (2007). Physiological aspects of soccer refereeing performance and training. *Sports Medicine*, 37(7), 625–646. https://doi.org/10.2165/00007256-200737070-00006

Castagna, C., Abt, G., D'Ottavio, S., & Weston, M. (2005). Age-related effects on fitness performance in elite-level soccer referees. *Journal of Strength and Conditioning Research*, 19(4), 785–790. https://doi.org/10.1519/R-14984.1

Castagna, C., Bendiksen, M., Impellizzeri, F. M., & Krustrup, P. (2012). Reliability, sensitivity and validity of the assistant referee intermittent endurance test (ARIET) – a modified Yo-Yo IE2 test for elite soccer assistant referees. *Journal of Sports Sciences*, 30(8), 767–775. https://doi.org/10.1080/02640414.2012.668705

Castagna, C., Bizzini, M., Araújo Póvoas, S. C., Schenk, K., Büsser, G., & D'Ottavio, S. (2019). Aerobic fitness in top-class soccer referees. *Journal of Strength and Conditioning Research,* 33(11), 3098–3104. https://doi.org/10.1519/JSC.0000000000002264

Castillo, D., Cámara, J., Castellano, J., & Yanci, Y. (2016a). Football match officials do not attain maximal sprinting speed during matches. *Kinesiology,* 48(2), 207–212. https://doi.org/10.26582/k.48.2.10

Castillo, D., Castagna, C., Cámara, J., Iturricastillo, A., & Yanci, J. (2018). Influence of team's rank on soccer referees' external and internal match loads during official matches. *Journal of Strength and Conditioning Research,* 32(6), 1715-1722. https://doi.org/10.1519/JSC.0000000000002040

Castillo, D., Raya-González, J., Usabiaga, O., Cámara-Tobalina, J., Castellano-Paulis, J., & Yanci-Irigoyen, J. (2019). Analysis of the success in soccer match officials' decisions during an international tournament according to contextual factors. *Revista Internacional de Ciencias del Deporte*, 57, 225–234. https://doi.org/10.5232/ricyde

Castillo, D., Weston, M., McLaren, S. J., Cámara, J., & Yanci, J. (2017). Relationships between internal and external match-load indicators in soccer match officials. *International Journal of Sports Physiology and Performance*, 12(7), 922-927. https://doi.org/10.1123/ijspp.2016-0392

Castillo, D., Yanci, J., Cámara, J., & Weston, M. (2016b). The influence of soccer match play on physiological and physical performance measures in soccer referees and assistant referees. *Journal of Sports Sciences*, 34(6), 557–63. https://doi.org/10.1080/02640414.2015.1101646

Catteeuw, P., Gilis, B., García-Aranda, J. M., Tresaco, F., Wagemans, J., & Helsen, W. (2010a). Offside decision making in the 2002 and 2006 FIFA World Cups. *Journal of Sports Sciences*, 28(10), 1027–1032. https://doi.org/10.1080/02640414.2010.491084

Catteeuw, P., Gilis, B., Jaspers, A., Wagemans, J., & Helsen, W. (2010b). Training of perceptual-cognitive skills in offside decision making. *Journal of Sport and Exercise Psychology,* 32(6), 845–861. https://doi.org/10.1123/jsep.32.6.845

Catteeuw, P., Gilis, B., Wagemans, J., & Helsen, W. (2010c). Offside decision making of assistant referees in the English Premier League: Impact of physical and perceptual-cognitive factors on match performance. *Journal of Sports Sciences*, 28(5), 471-481. https://doi.org/10.1080/02640410903518184

Catteeuw, P., Gilis, B., Wagemans, J., & Helsen, W. (2010d). Perceptual-cognitive skills in offside decision making: expertise and training effects. *Journal of Sport and Exercise Psychology*, 32(6), 828–844. https://doi.org/10.1123/jsep.32.6.828

Catteeuw, P., Helsen, W., Gilis, B., Van Roie, E., & Wagemans, J. (2009a). Visual scan patterns and decision-making skills of expert assistant referees in offside situations. *Journal of Sport and Exercise Psychology*, 31(6), 786–797. https://doi.org/10.1123/jsep.31.6.786

Catteeuw, P., Helsen, W., Gilis, B., & Wagemans, J. (2009b). Decision-making skills, role specificity, and deliberate practice in association football refereeing. *Journal of Sports Sciences*, 27(11), 1125–1136. https://doi.org/10.1080/02640410903079179

Červený, J., Van Ours, J. C., & Van Tuijl, M. A. (2018). Effects of a red card on goal-scoring in World Cup football matches. *Empirical Economics*, 55, 1287–1285. https://doi.org/10.1007/s00181-017-1287-5

Cohen, J. (1992). A power primer. *Psychological Bulletin,* 112(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Constantinou, A. C., Fenton, N. E., Jackson, L., & Pollock, H. (2014). Bayesian networks for unbiased assessment of referee bias in association football. *Psychology of Sport and Exercise*, 15(5), 538-547. https://doi.org/10.1016/j.psychsport.2014.05.009

Cooke, C. (2009). Maximal oxygen uptake, economy and efficiency. In R. Eston, & T. Reilly (3rd Ed.), Kinanthropometry and exercise physiology laboratory manual: tests, procedures and data (pp. 174–212). London: Routledge.

Corrigan, S. L., Dwyer, D. B., Harvey, B., & Gastin, P. B. (2019). The influence of match characteristics and experience on decision-making performance in AFL umpires. *Journal of Science and Medicine in Sport*, 22(1), 112–116. https://doi.org/10.1016/j.jsams.2018.06.005

Coutts, A. J., & Duffield, R. (2010). Validity and reliability of GPS devices for measuring movement demands of team sports. *Journal of Science and Medicine in Sport*, 13(1), 133–135. https://doi.org/10.1016/j.jsams.2008.09.015

Craig, C. (2013). Understanding perception and action in sport: how can virtual reality technology help? *Sports Technology*, 6(4), 161–169. https://doi.org/10.1080/19346182.2013.855224

Dawson, P., & Dobson, S. (2010). The influence of social pressure and nationality on individual decisions: evidence from the behaviour of referees. *Journal of Economic Psychology*, 31(2), 181–191. https://doi.org/10.1016/j.joep.2009.06.001

Delaney, J.A., Thornton, H.R., Rowell, A.E., Dascombe, B.J., Aughey, R.J., & Duthie, G.M. (2018). Modelling the decrement in running intensity within professional soccer players. *Science and Medicine in Football,* 2(2), 86–92. https://doi.org/10.1080/24733938.2017.1383623

Del Campo, V. L., Fariñas, A. C., Márquez, F. J. D., & Martín, J. M. (2018). The influence of refereeing experiences judging offside actions in football. *Psychology of Sport and Exercise*, 37, 139–145. https://doi.org/10.1016/j.psychsport.2018.03.004

De Oliveira, M. C., Orbetelli, R., & De Barros Neto, T. L. (2011). Call accuracy and distance from the play: a study with Brazilian soccer referees. *International Journal of Exercise Science*, 4(1), 30–38. Available at: https://digitalcommons.wku.edu/ijes/vol4/iss1/5

Dias, C., Cruz, J. F., Fonseca, A. M. (2012). The relationship between multidimensional competitive anxiety, cognitive threat appraisal, and coping strategies: A multi-sport study. *International Journal of Sport and Exercise Psychology*, 10(1), 52–65. https://doi.org/10.1080/1612197X.2012.645131

Doncaster, G., Page, R., White, P., Svenson, R., & Twist, C. (2019). Analysis of physical demands during youth soccer match-play: considerations of sampling method and epoch length. *Research Quarterly for Exercise and Sport*, 91(2), 1–9. https://doi.org/10.1080/02701367.2019.1669766

D'Ottavio, S., & Castagna, C. (2002). Physiological aspects of soccer refereeing. In: W. Spinks, T. Reilly, & A. Murphy (Eds.), Science and football IV (pp. 144–150). London: Routledge.

Downward, P., & Jones, M. (2007). Effects of crowd size on referee decisions: analysis of the FA cup. *Journal of Sports Sciences*, 24(14), 1541–1545. https://doi.org/10.1080/02640410701275193

Drust, B., & Green, M. (2013). Science and football: evaluating the influence of science on performance. *Journal of Sports Sciences,* 31(13), 1377–1382. https://doi.org/10.1080/02640414.2013.828544

Dunn, J. G. H., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale-construction research: Issues and recommendations. Measurement in Physical Education and Exercise Science, 3(1), 15–36. https://doi.org/10.1207/s15327841mpee0301_2

Eccles, D. W., & Arsal, G. (2017). The think aloud method: what is it and how do I use it? *Qualitative Research in Sport, Exercise, and Health*, 9(4), 514–531. https://doi.org/10.1080/2159676X.2017.1331501

Edwards, S. (1993). High performance training and racing. In S. Edwards (Ed.), The heart rate monitor book (pp. 113–123). Sacramento, CA: Feet Fleet Press.

Elliott, S., Whitehead, A., & Magias, T. (2020). Thought processes during set shot goalkicking in Australian Rules football: An analysis of youth and semi-professional footballers using Think Aloud. *Psychology of Sport and Exercise*, 48, 101659. https://doi.org/10.1016/j.psychsport.2020.101659

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, K. (2014). Qualitative content analysis: A focus on trustworthiness: *SAGE Open*, 4(1), 1–10. https://doi.org/10.1177/2158244014522633

Elsworthy, N., Burke, D., & Dascombe, B. J. (2014a). Factors relating to the decision-making performance of Australian football officials. *International Journal of Performance Analysis in Sport*, 14(2), 401–410. https://doi.org/10.1080/24748668.2014.11868730

Elsworthy, N., Burke, D., Scott, B. R., Stevens, C. J., & Dascombe, B. J. (2014b). Physical and decision-making demands of Australian football umpires during competitive matches. *Journal of Strength and Conditioning Research*, 28(12), 3502–3507. https://doi.org/10.1519/JSC.0000000000000567

Emmonds, S., O'Hara, J., Till, K., Jones, B., Brightmore, A., & Cooke, C. (2015). Physiological and movement demands of Rugby League referees: influence on penalty accuracy. *Journal of Strength and Conditioning Research*, 29(12), 3367–3374. https://doi.org/10.1519/JSC.0000000000001002

Ericsson, K., Krampe, R. T., & Tesch-Roemer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, 7(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Farrow, D., McCrae, J., Gross, J., & Abernethy, B. (2010). Revisiting the relationship between pattern recall and anticipatory skill. *International Journal of Sport Psychology*, 41(1), 91–106. Available at: https://psycnet.apa.org/record/2010-10759-006

Faude, O., Koch, T., & Meyer, T. (2012). Straight sprinting is the most frequent action in goal situations in professional football. *Journal of Sports Sciences*, 30(7), 625–631. https://doi.org/10.1080/02640414.2012.665940

Faull, O. K., Cox, P. J., & Pattinson, K. T. S. (2016). Psychophysical differences in ventilator awareness and breathlessness between athletes and sedentary individuals. *Frontiers in Physiology*, 7:231. https://doi.org/10.3389/fphys.2016.00231

Fédération Internationale de Football Association. (2016). Laws of the game. Zurich:FIFA.

Fédération Internationale de Football Association. (2018). Laws of the game. Zurich:FIFA.

Fereday, K., Hills, S. P., Russell, M., Smith, J., Cunningham, D. J., Shearer, D., McNarry, M., & Kilduff, L. P. (2020). A comparison of rolling averages versus discrete time epochs for assessing the worst-case scenario locomotor demands of professional soccer match-play. *Journal of Science and Medicine in Sport,* 23(8), 764–769. https://doi.org/10.1016/j.jsams.2020.01.002

Fitzpatrick, J. F., Hicks, K. M., & Hayes, P, R. (2018). Dose-response relationship between training load and changes in aerobic fitness in professional youth soccer players. *International Journal of Sports Physiology and Performance*, 13(10), 1365–1370. https://doi.org/10.1123/ijspp.2017-0843

Ford, J. L., Ildefonso, K., Jones, M. L., & Arvinen-Barrow, M. (2017). Sport-related anxiety: current insights. *Open Access Journal of Sports Medicine,* 8, 205–212. https://doi.org/10.2147/OAJSM.S125845

Ford, P. R., Delgado Bordonau, J. L., Bonanno, D., Tavares, J., Groenendijk, C., Fink, C., Gualtieri, D., Gregson, W., Varley, M. C., Weston, M., Lolli, L., Platt, D., & Di Salvo, V. (2020). A survey of talent identification and development processes in the youth academies of professional soccer clubs from around the world. *Journal of Sports Sciences,* 38(11-12), 1269–1278. https://doi.org/10.1080/02640414.2020.1752440

Ford, P. R., Low, J., McRobert, A. P., & Williams, A. M. (2010). Developmental activities that contribute to high or low performance by elite cricket batters when recognizing type of delivery from bowlers' advanced postural cues. Journal of Sport and Exercise Psychology, 32(5), 638–654. https://doi.org/10.1123/jsep.32.5.638

Franchini, M., Ferraresi, I., Modena, R., Schena, F., Coutts, A. J., & Impellizzeri, F. M. (2016). Use of the CR100 scale for session rating of perceived exertion in soccer and its interchangeability with the CR10. *International Journal of Sports Physiology and Performance*, 11(3), 388–392. https://doi.org/10.1123/ijspp.2015-0273

Gabbett, T. J. (2016). The training-injury prevention paradox: should athletes be training smarter and harder? *British Journal of Sports Medicine*,. 50(5), 273–280. https://doi.org/10.1136/bjsports-2015-095788

Gaoua, N., de Oliveira, R. F., & Hunter, S. (2017). Perception, action, and cognition of football referees in extreme temperatures: impact on decision performance. *Frontiers in Psychology,* 8, 1479. https://doi.org/10.3389/fpsyg.2017.01479

Gencay, O. A. & Aydin, E. E. (2017). The effects of teams, individual players and spectators on basketball officials' decision-making processes. *The Anthropologist*, 20(3), 503–509. https://doi.org/10.1080/09720073.2015.11891754

Ghasemi, A., Momeni, M., Jafarzadehpur, E., Rezaee, M., & Taheri, H. (2011). Visual skills involved in decision making by expert referees. *Perceptual and Motor Skills*, 112(1), 161–171. https://doi.org/10.2466/05.22.24.27.PMS.112.1.161-171

Gilis, B., Helsen, W., Catteeuw, P., & Wagemans, J. (2008). Offside decisions by expert assistant referees in association football: perception and recall of spatial positions in complex dynamic events. *Journal of Experimental Psychology Applied*, 14(1), 21– 35. https://doi.org/10.1037/1076-898X.14.1.21

Gilis, B., Weston, M., Helsen, W. F., Junge, A., & Dvorak, J. (2006). Interpretation and application of the laws of the game in football incidents leading to player injuries. *International Journal of Sports Psychology*, 37(2-3), 121–138. Available at: https://psycnet.apa.org/record/2006-22320-004

Gómez-Carmona, C. D., & Pino-Ortega, J. (2016). Kinematic and physiological analysis of the performance of the football referee and its relationship with decision making. Journal of Human Sport and Exercise, 11(4), 397–414. https://doi.org/10.14198/jhse.2016.114.01

Goodall, S., Thomas, K, Harper, L. D., Hunter, R., Parker, P., Stevenson, E., West, D., Russell, M., & Howatson, G. (2017). The assessment of neuromuscular fatigue during 120 min of simulated soccer exercise. *European Journal of Applied Physiology,* 117(4), 687–697. https://doi.org/10.1007/s00421-017-3561-9

Goumas, C. (2013). Referee expertise and home-team bias in European football. *International Journal of Sport Psychology*, 44(5), 429–440. https://doi.org/10.7352/IJSP.2013.44.429

Goumas, C. (2014). Home advantage and referee bias in European football. *European Journal of Sport Science*, 14(Suppl1), S243–S249. https://doi.org/10.1080/17461391.2012.686062

Gregson, W., Drust, B., Atkinson, G., & Salvo, V. D. (2010). Match-to-match variability of high-speed activities in premier league soccer. *International Journal of Sports Medicine,* 31(4), 237–242. https://doi.org/10.1055/s-0030-1247546

Guillén, F., & Feltz, D. L. (2011). A conceptual model of referee efficacy. *Frontiers in Psychology,* 2, 25. https://doi.org/10.3389/fpsyg.2011.00025

Gulec, U., & Yilmaz, M. (2016). A serious game for improving the decision making skills and knowledge levels of Turkish football referees according to the laws of the game. *Springerplus*, 5, 622. https://doi.org/10.1186/s40064-016-2227-0

Hacicaferoğlu, S., & Gündoğdu, C. (2014). Surveying the exposure level of intimidation (mobbing) behaviours of the football referees. *Journal of Physical Education and Sport*, 14(1), 120–126. Available at: https://doi.org/10.7752/jpes.2014.01019

Hader, K., Mendez-Villanueva, A., Ahmaidi, S., Williams, B. K., & Buchheit, M. (2014). Changes of direction during high-intensity intermittent runs: neuromuscular and metabolic responses. *BMC Sports Science, Medicine and Rehabilitation*, 6(1), 2. https://doi.org/10.1186/2052-1847-6-2

Halson, S. L. (2014). Monitoring training load to understand fatigue in athletes. *Sports Medicine*, 44(Suppl2), S139–S147. https://doi.org/10.1007/s40279-014-0253-z

Hancock, D. J., & Ste-Marie, D. M. (2013). Gaze behaviours and decision making accuracy of higher- and lower-level ice hockey referees. *Psychology of Sport and Exercise*, 14(1), 66–71. https://doi.org/10.1016/j.psychsport.2012.08.002

Harper, L. D., Fothergill, M., West, D. J., Stevenson, E., & Russell, M. (2016a). Practitioners' perceptions of the soccer extra-time period: implications for future research. *PLoS One*, 11(7), e0157687. https://doi.org/10.1371/journal.pone.0157687

Harper, L. D., Hunter, R., Parker, P., Goodall, S., Thomas, K., Howatson, G., West, D. J., Stevenson, E., & Russell, M. (2016b). Test-retest reliability of physiological and performance responses to 120 minutes of simulated soccer match play. *Journal of Strength and Conditioning Research,* 30(11), 3178–3186. https://doi.org/10.1519/JSC.0000000000001400

Harper, L. D., & McCunn, R. (2017). "Hand in glove": using qualitative methods to connect research and practice. *International Journal of Sports Physiology and Performance*, 12(7), 990–993. https://doi.org/10.1123/ijspp.2017-0081

Hawley, J. A. (2008). Specificity of training adaptation: Time for a rethink? *Journal of Physiology*, 586(1), 1–2. https://doi.org/10.1113/jphysiol.2007.147397

Heikkilä, M., Valve, R., Lehtovirta, M., & Fogelholm, M. (2017). Development of a nutrition knowledge questionnaire for young endurance athletes and their coaches. *Scandinavian Journal of Medicine and Science in Sports*, 28(3), 873–880. https://doi.org/10.1111/sms.12987

Helsen, W., & Bultynck, J. B. (2004). Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, 22(2), 179–89. https://doi.org/10.1080/02640410310001641502

Helsen, W., Gilis, B., & Weston, M. (2006). Errors in judging "offside" in association football: test of the optical error versus the perceptual flash-lag hypothesis. *Journal of Sports Sciences*, 24(5), 521–528. https://doi.org/10.1080/02640410500298065

Helsen, W., MacMahon, C. and Spitz, J. (2019). Decision making in match officials and judges. In A. M. Williams, & R. C. Jackson (Eds.), *Anticipation and Decision Making in Sport* (pp. 250–266). London: Routledge.

Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. *Journal of Sport and Exercise Psychology*, 20(1), 12–34. https://doi.org/10.1123/jsep.20.1.12

Hill, D.M., Matthews, N., & Senior, R. (2016). The psychological characteristics of performance under pressure in professional rugby union referees. *The Sport Psychologist*, 30(4), 376–387. https://doi.org/10.1123/tsp.2015-0109

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. Medicine and Science in Sports and Exercise, 41(1), 3–13. https://doi.org/10.1249/MSS.0b013e31818cb278

Hopkins, W. G. (2010). Linear models and effect magnitudes for research, clinical and practical applications. Sportscience. 14, 49–57. Available at: https://sportsci.org/2010/wghlinmod.htm

Hoseini, S. H., Aslankhani, M.A., Abdoli, B., & Mohammadi, F. (2011). The relationship between the number of crowds with anxiety and the function of the soccer premier league referees. *Procedia - Social and Behavioural Sciences*, 30, 2374–2378. https://doi.org/10.1016/j.sbspro.2011.10.463

Hossner, E. J., Schnyder, U., Schmid, J., & Kredel, R. (2019). The role of viewing distance and viewing angle on referees' decision-making performance during the FIFA World Cup 2014. *Journal of Sports Sciences*, 37(13), 1481–1489. https://doi.org/10.1080/02640414.2019.1570898

Hüttermann, S., Helsen, W. F., Put, K., & Memmert, D. (2018). Does visual attention impact on decision making in complex dynamic events? *Journal of Sport and Exercise Psychology*, 40(3), 163–166. https://doi.org/10.1123/jsep.2017-0169

Hüttermann, S., & Memmert, D. (2014). Does the inverted-U function disappear in expert athletes? An analysis of the attentional behaviour under physical exercise of athletes and non-athletes. *Physiology and Behaviour*, 131, 87–92. https://doi.org/10.1016/j.physbeh.2014.04.020

Hüttermann, S., Noël, B., & Memmert, D. (2017). Evaluating erroneous offside calls in soccer. *PLoS One*, 12(3), e0174358. https://doi.org/10.1371/journal.pone.0174358

Jeacocke, N. A., & Burke, L. M. (2010). Methods to standardize dietary intake before performance testing. *International Journal of Sports Nutrition and Exercise Metabolism,* 20(2), 87–103. https://doi.org/10.1123/ijsnem.20.2.87

Jennings, D., Cormack, S., Coutts, A. J., Boyd, L., & Aughey, R. J. (2010). The validity and reliability of GPS units for measuring distance in team sport specific running patterns. *International Journal of Sports Physiology and Performance*, 5(3), 328–341. https://doi.org/10.1123/ijspp.5.3.328

Johansen, B. T., & Haugen, T. (2013). Anxiety level and decision-making among Norwegian top-class soccer referees. *International Journal of Sport and Exercise Psychology*, 11(2), 215–226. https://doi.org/10.1080/1612197X.2013.773665

Johnson, U., Andersson, K., Fallby, J. (2011). Sport psychology consulting among Swedish premier soccer coaches. *International Journal of Sport and Exercise Psychology,* 9(4), 308–322. https://doi.org/10.1080/1612197X.2011.623455

Johnstone, R. (2008). On referee bias, crowd size, and home advantage in the English soccer Premiership. *Journal of Sports Sciences*, 26(6), 563–568. https://doi.org/10.1080/02640410701736780

Joo, C. H., & Jee, H. (2019). Activity profiles of top-class players and referees and accuracy in foul decision-making during Korean National League games. *Journal of Strength and Conditioning Research*, 33(9), 2530–2540. https://doi.org/10.1519/JSC.0000000000003083

Kaiseler, M., Polman, R. C. J., & Nicholls, A. R. (2013). Gender differences in stress, appraisal, and coping during golf putting. *International Journal of Sport and Exercise Psychology,* 11(3), 258–272. https://doi.org/10.1080/1612197X.2013.749004

Kelly, D. M., Strudwick, A. J., Atkinson, G., Drust, B., & Gregson, W. (2020). Quantification of training and match-load distribution across a season in elite English Premier League soccer players. *Science and Medicine in Football*, 4(1), 59–67. https://doi.org/10.1080/24733938.2019.1651934

Kittel, A., Elsworthy, N., & Spittle, M. (2019a). Incorporating perceptual decision-making training into high-intensity interval training for Australian football umpires. *Journal of Sports Sciences,* 37(1), 29-35. https://doi.org/10.1080/02640414.2018.1480257

Kittel, A., Larkin, P., Elsworthy, N., Lindsay, R., & Spittle, M. (2020). Effectiveness of 360° virtual reality and match broadcast video to improve decision-making skill. *Science and Medicine in Football,* 4(4), 255–262. https://doi.org/10.1080/24733938.2020.1754449

Kittel, A., Larkin, P., Elsworthy, N., & Spittle, M. (2019b). Using 360° virtual reality as a decision-making assessment tool in sport. *Journal of Science and Medicine in Sport*, 22(9), 1049–1053. https://doi.org/10.1016/j.jsams.2019.03.012

Kittel, A., Larkin, P., Elsworthy, N., & Spittle M. (2019c). Video-based testing in sporting officials: a systematic review. *Psychology of Sport and Exercise*, 43, 261–270. https://doi.org/10.1016/j.psychsport.2019.03.013

Kittel, A., Larkin, P., Elsworthy, N., & Spittle M. (2021). Transfer of 360° virtual reality and match broadcast video-based tests to on-field decision-making. *Science and Medicine in Football,* 5(1), 79–86. https://doi.org/10.1080/24733938.2020.1802506

Knaier, R., Infanger, D., Niemeyer, M., Cajochen, C., & Schmidt-Trucksäss, A. (2019). In athletes, the diurnal variations in maximum oxygen uptake are more than twice as large as the day-to-day variations. *Frontiers in Physiology*, 10, 219. https://doi.org/10.3389/fphys.2019.00219

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. *Handbook of Survey Research*, *2*(3), 263–314.

Krustrup, P., Helsen, W., Randers, M. B., Christensen, J. F., MacDonald, C., Rebelo, A. N., & Bangsbo, J. (2009). Activity profile and physical demands of football referees and assistant referees in international games. *Journal of Sports Sciences*, *27*(11), 1167–1176. https://doi.org/10.1080/02640410903220310

Lago-Peñas, C., & Gómez-López, M. (2016). The influence of referee bias on extra time in elite soccer matches. *Perceptual and Motor Skills*, 122(2), 666–677. https://doi.org/10.1177/0031512516633342

Lane, A. M., Nevill, A. M., Ahmad, N. S., & Balmer, N. (2006). Soccer referee decision-making: "shall I blow the whistle?". *Journal of Sports Science and Medicine*, 5(2), 243–53. Available at: https://www.jssm.org/jssm-05-243.xml%3EFulltext

Larkin, P., Mesagno, C., Berry, J., & Spittle, M. (2014a). Development of a valid and reliable video-based decision-making test for Australian football umpires. *Journal of Science and Medicine in Sport*, 17(5), 552–555. https://doi.org/10.1016/j.jsams.2013.08.001

Larkin, P., Mesagno, C., Berry, J., & Spittle, M. (2018a). Exploration of the perceptual-cognitive processes that contribute to in-game decision-making of Australian football umpires. *International Journal of Sport and Exercise Psychology*, 16(2), 112–124. https://doi.org/10.1080/1612197X.2016.1167760

Larkin, P., Mesagno, C., Berry, J., Spittle, M., & Harvey, M. (2018b). Video-based training to improve perceptual-cognitive decision-making performance of Australian football umpires. *Journal of Sports Sciences*, 36(3), 239–246. https://doi.org/10.1080/02640414.2017.1298827

Larkin, P., O'Brien, B., Mesagno, C., Berry, J., Harvey, J., & Spittle, M. (2014b). Assessment of decision-making performance and in-game physical exertion of Australian football umpires. *Journal of Sports Sciences*, 32(15), 1446–1453. https://doi.org/10.1080/02640414.2014.896998

Levi, H. R., & Jackson, R. C. (2018). Contextual factors influencing decision making: Perceptions of professional soccer players. Psychology of Sport and Exercise, 37, 19–25. https://doi.org/10.1016/j.psychsport.2018.04.001

Lex, H., Pizzera, A., Kurtes, M., & Schack, T. (2015). Influence of players' vocalisations on soccer referees' decisions. *European Journal of Sport Science*, 15(5), 424–428. https://doi.org/10.1080/17461391.2014.962620

Light, R. L., Harvey, S., & Mouchet, A. (2014). Improving "at-action" decision-making in team sports through a holistic coaching approach. *Sport, Education and Society,* 19(3), 258–275. https://doi.org/10.1080/13573322.2012.665803

Los Arcos, A., Mendez-Villanueva, A., & Martínez-Santos, R. (2017). In-season training periodization of professional soccer players. *Biology of Sport*, 34(2), 149–155. https://doi.org/10.5114/biolsport.2017.64588

Lovell R, A., Halley S., Siegler J., Wignell T., Coutts A. J., & Massard, T. (2020). Use of numerically blinded ratings of perceived exertion in soccer: Assessing concurrent and construct validity. *International Journal of Sports Physiology and Performance*, 15(10), 1430–1436. https://doi.org/10.1123/ijspp.2019-0740

MacMahon, C., Helsen, W. F., & Starkes, J. L., & Weston, M. (2007). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, 25(1), 65–78. https://doi.org/10.1080/02640410600718640

MacMahon, C., & Plessner, H. (2008). The sport official in research and practice. In D. Farrow, J. Baker, & C. MacMahon (Eds.), Developing sport expertise (pp. 172–192). New York, NY: Routledge.

MacNamara, A., & Collins, D. (2011). Development and initial validation of the psychological characteristics of developing excellence questionnaire. *Journal of Sports Sciences*, 29(12), 1273–1286. https://doi.org/10.1080/02640414.2011.589468

Mallo, J, Frutos, P. G., Juárez, D., & Navarro, E. (2012). Effect of positioning on the accuracy of decision-making of association football top-class referees and assistant referees during competitive matches. *Journal of Sports Sciences*, 30(13), 1437–1445. https://doi.org/10.1080/02640414.2012.711485

Mallo, J., Veiga, S., López de Subijana, C., & Navarro, E. (2010). Activity profile of top-class female soccer refereeing in relation to the position of the ball. *Journal of Science and Medicine in Sport*, 13(1), 129–132. https://doi.org/10.1016/j.jsams.2008.09.006

Malone, J. J., Di Michelle, R., Morgans, R., Burgess, D., Morton, J. P., & Drust, B. (2015). Seasonal training-load quantification in elite English Premier League soccer players. International Journal of Sports Physiology and Performance, 10(4), 489–497. https://doi.org/10.1123/ijspp.2014-0352

Mann, D. T., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: a meta-analysis. *Journal of Sport and Exercise Psychology*, 29(4), 457–478. https://doi.org/10.1123/jsep.29.4.457

Marteniuk, R. G. (1976). Cognitive information processes in motor short-term memory and movement production. In G. E. Stelmach (Ed.), *Motor control* (pp.175–186). Academic Press.

Martin-Garcia, A., Casamichana, D., Gomez Diaz, A., Cos, F., & Gabbett, T.J. (2018). Positional differences in the most demanding passages of play in football competition. *Journal of Sports Science and Medicine*, 17(4), 563–570.

Mascarenhas, D. R., Button, C., O'Hara, D., & Dicks, M. (2009). Physical performance and decision making in association football referees: A naturalistic study. *Open Sports Science Journal*, 2(9), 1–9. https://doi.org/10.2174/1875399X00902010001

Mascarenhas, D., O'Hare, D., & Plessner, H. (2006). The psychological and performance demands of association football refereeing. *International Journal of Sport Psychology*, 37(2-3), 99–120.

Mathers, J. F., & Brodie, K. (2011). Elite refereeing in professional soccer: A case study of mental skills support: A case study of mental skills support. *Journal of Sport Psychology in Action*, 2(3), 171–182. https://doi.org/10.1080/21520704.2011.609018

McCunn, R., Gibson, N. V., Fullagar, H. H. K., & Harper, L. D. (2018). Professional youth football academy injury data: collection procedures, perceived value, and use. *Science and Medicine in Football*, 2(2), 141–148. https://doi.org/10.1080/24733938.2017.1410564

McDougall, M., Nesti, M., & Richardson, D. (2015). The challenges of sport psychology delivery in elite and professional sport: Reflections from experienced sport psychologists. *The Sport Psychologist,* 29(3), 265–277. https://doi.org/10.1123/tsp.2014-0081

McEwan, G., Arthur, R., Phillips, S. M., Gibson, N. V., & Easton C. (2018). Interval running with self-selected recovery: physiology, performance, and perception. *European Journal of Sport Science*, 18(8), 1058–1067. https://doi.org/10.1080/17461391.2018.1472811

McEwan, G., Unnithan, V., Easton, C., & Arthur, R. (2020). Development and validation of the Referee Training Activity Questionnaire (RTAQ): Towards a better understanding of the training practices of soccer officials. *Journal of Sports Sciences,* 38(24), 2782–2793. https://doi.org/10.1080/02640414.2020.1800371

McLaren, S. J., Taylor, J. M., Macpherson, T. W., Spears, I. R., & Weston, M. (2020). Systematic reductions in differential ratings of perceived exertion across a 2-week repeated-sprint-training intervention that improved players' high-speed-running abilities. *International Journal of Sports Physiology and Performance,* 15(10), 1414–1421. https://doi.org/10.1123/ijspp.2019-0568

Mellalieu, S. D., Hanton, S., & Fletcher, D. (2009). A competitive anxiety review: Recent directions in sport psychology research. In S. Hanton, & S. D. Mellalieu (Eds.) *Literature reviews in sport psychology research* (pp. 1–45). New York, NY: Nova Science Publishers.

Memmert, D. (2009). Pay attention! A review of visual attentional expertise in sport. *International Review of Sport and Exercise Psychology*, 2(2), 119–138. https://doi.org/10.1080/17509840802641372

Midgley, A. W., McNaughton, L. R., Polman, R., & Marchant, D. (2007). Criteria for determination of maximal oxygen uptake: A brief critique and recommendations for future research. *Sports Medicine,* 37(12), 1019–1025. https://doi.org/10.2165/00007256-200737120-00002

Moher, D., Shamsee, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group (2015). Preferred reporting items for systematic review and meta-analysis protocol (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1):1. https://doi.org/10.1186/2046-4053-4-1

Mohr, M., Krustrup, P., & Bangsbo, J. (2005). Fatigue in soccer: a brief review. *Journal of Sports Sciences,* 23(6), 593–599. https://doi.org/10.1080/02640410400021286

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of Clinical Epidemiology, 63(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Moore, L. J., Harris, D. J., Sharpe, B. T., Vine, S. J., & Wilson, M. R. (2019). Perceptual-cognitive expertise when refereeing the scrum in rugby union. *Journal of Sports Sciences*, 37(15), 1778–1786. https://doi.org/10.1080/02640414.2019.1594568

Mujika, I. (2017). Quantification of training and competition loads in endurance sports: Methods and applications. *International Journal of Sports Physiology and Performance,* 12(Suppl2), S29–S217. https://doi.org/10.1123/ijspp.2016-0403

Müller, M. L., Redfern, M. S., & Jennings, J. R. (2007). Postural prioritization defines the interaction between a reaction time task and postural perturbations. *Experimental Brain Research*, 183(4), 447–456. https://doi.org/10.1007/s00221-007-1065-2

Morris, G., & O'Connor, D. (2017). Key attributes of expert NRL referees. *Journal of Sports Sciences*, 35(9), 852–857. https://doi.org/10.1080/02640414.2016.1194524

Nevill, A. M., Atkinson, G., Hughes, M. D., & Cooper, S. M. (2002). Statistical methods for analysing discrete and categorical data recorded in performance analysis. Jo*urnal of Sports Sciences*, 20(10), 829–844. https://doi.org/10.1080/026404102320675666

Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, 3(4), 261–272. https://doi.org/10.1016/S1469-0292(01)00033-4

Nevill, A. M., Hemingway, A., Greaves, R., Dallaway, A., & Devonport, T. J. (2017). Inconsistency of decision-making, the Achilles heel of referees. *Journal of Sports Sciences*, 35(22), 2257–2261. https://doi.org/10.1080/02640414.2016.1265143

Nicholls, A. R., & Polman, R. C. J. (2008). Think aloud: Acute stress and coping strategies during golf performance. *Anxiety Stress Coping*, 21(3), 283–294. https://doi.org/10.1080/10615800701609207

Nicholls, A. R., Polman, R., & Levy, A. R. (2010). Coping self-efficacy, pre-competitive anxiety, and subjective performance among athletes. *European Journal of Sport Science*, 10(2), 97–102. https://doi.org/10.1080/17461390903271592

North, J. S., Hope, E., & Williams, A. M. (2016). The relative importance of different perceptual-cognitive skills during anticipation. *Human Movement Science*, 49, 170–177. https://doi.org/10.1016/j.humov.2016.06.013

North, J. S., Ward, P., Ericsson, A., & Williams, A. M. (2011). Mechanisms underlying skilled anticipation and recognition in a dynamic and temporally constrained domain. *Memory*, 19(2), 155–168. https://doi.org/10.1080/09658211.2010.541466

Oliva-Lozano, J. M., Gómez-Carmona, C. D., Rojas-Valverde, D., Fortes, V., & Pino-Ortega, J. (2021). Effect of training day, match, and length of the microcycle on the worst-case

scenarios in professional soccer players. *Research in Sports Medicine.* https://doi.org/10.1080/15438627.2021.1895786

Oudejans, R. R. D., Bakker, F. C., Verheijen, R., Gerrits, J. C., Steinbrückner, M., & Beek, P. J. (2005). How position and motion of expert assistant referees in soccer relate to the quality of their offside judgements during actual match play. *International Journal of Sport Psychology*, 36, 3–21. Available at: https://psycnet.apa.org/record/2005-06291-002

Oudejans, R. R. D., & Pijpers, J. R. (2009). Training with anxiety has a positive effect on expert perceptual-motor performance under pressure. *Quarterly Journal of Experimental Psychology,* 62(8), 1631–1647. https://doi.org/10.1080/17470210802557702

Oudejans, R. R. D., & Pijpers, J. R. (2010). Training with mild anxiety may prevent choking under higher levels of anxiety. *Psychology of Sport and Exercise,* 11(1), 44–50. https://doi.org/10.1016/j.psychsport.2009.05.002

Page, K., & Page, L. (2010). Alone against the crowd: individual differences in referees' ability to cope under pressure. *Journal of Economic Psychology*, 31(2), 192–199. https://doi.org/10.1016/j.joep.2009.08.007

Page, R. M., Marrin, K., Brogden, C. M., & Greig, M. (2015). Biomechanical and physiological responses to a contemporary soccer match-play simulation. *Journal of Strength and Conditioning Research,* 29(10), 2860–2866. https://doi.org/10.1519/JSC.0000000000000949

Pain, M. A., & Harwood, C. G. (2004). Knowledge and perceptions of sport psychology within English soccer. *Journal of Sports Sciences*, 22(9), 813–826. https://doi.org/10.1080/02640410410001716670

Paradis, K., Larkin, P., & O'Connor, D. (2016). The effects of physical exertion on decision-making performance of Australian football umpires. *Journal of Sports Sciences*, 34(16), 1535–1541. https://doi.org/10.1080/02640414.2015.1122205

Patton, M. Q. (2001). Qualitative research and evaluation methods (3rd ed.). Newbury Park, CA: Sage.

Picazo-Tadeo, A. J., González-Gómez, F., & Guardiola, J. (2017). Does the crowd matter in refereeing decisions? Evidence from Spanish soccer. *International Journal of Sport and Exercise Psychology*, 15(5), 447–459. https://doi.org/10.1080/1612197X.2015.1126852

Pietraszewski, P., Roczniok, R., Maszczyk, A., Grycmann, P., Roleder, T., Stanula, A., Fidos-Czuba, O., & Ponczek, M. (2014). The elements of executive attention in top class soccer referees and assistant referees. *Journal of Human Kinetics*, 40(1), 235–243. https://doi.org/10.2478/hukin-2014-0025

Piggott, B., Müller, S., Chivers, P., Papaluca, C., & Hoyne, G. (2019). Is sports science answering the call for interdisciplinary research? A systematic review. *European Journal of Sport Science*, 19(3), 267–286. https://doi.org/10.1080/17461391.2018.1508506

Pizzera, A., Marrable, J., & Raab, M. (2016). How effective is a video review system in soccer? *Research Quarterly for Exercise and Sport,* 87(S1), S83. https://doi.org/10.1080/02701367.2016.1213610

Plessner, H., & Betsch, T. (2001). Sequential effects in important referee decisions" the case of penalties in soccer. *Journal of Sport and Exercise Psychology,* 23(3), 254–259. https://doi.org/10.1123/jsep.23.3.254

Plessner, H., & Haar, T. (2006). Sports performance judgements from a social cognitive perspective. *Psychology of Sport and Exercise*, 7(6), 555–575. https://doi.org/10.1016/j.psychsport.2006.03.007

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539–569. http://10.1146/annurev-psych-120710-100452

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health*, 29(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing and Health*, 30(4), 459–467. https://doi.org/10.1002/nur.20199

Pugliese, L., Serpiello, F. R., Millet, G. P., & La Torre, A. (2014). Training diaries during altitude training camp in two olympic champions: An observa- tional case study. *Journal of Sports Science and Medicine*, 13(3), 666–672. https://www.jssm.org/hf.php?id=jssm-13-666.xml

Put, K., Wagemans, J., Jaspers, A., & Helsen, W. F. (2013). Web-based training improves on-field offside decision-making performance. *Psychology of Sport and Exercise*, 14(4), 577–585. https://doi.org/10.1016/j.psychsport.2013.03.005

Put, K., Wagemans, J., Spitz, J., Gallardo, M. A., Williams, A. M., & Helsen, W. F. (2014). The use of 2D and 3D information in a perceptual-cognitive judgement task. *Journal of Sports Sciences*, 32(18), 1688–1697. https://doi.org/10.1080/02640414.2014.912760

Put, K., Wagemans, J., Spitz, J., Williams, A. M., & Helsen, W. F. (2016). Using web-based training to enhance perceptual-cognitive skills in complex dynamic offside events. *Journal of Sports Sciences*, 34(2), 181–189. https://doi.org/10.1080/02640414.2015.1045926

Raab, M., Bar-Eli, M., Plessner, H., & Araújo, D. (2018). The past, present and future of research on judgment and decision-making in sport. *Psychology of Sport and Exercise*, 42, 25–32. https://doi.org/10.1016/j.psychsport.2018.10.004

Raglin, J., & Hanin, Y. (2000). Competitive anxiety and athletic performance. In Y. L. Hanin (Eds.), *Emotions in sport* (pp. 93–112). Champaign, IL: Human Kinetics.

Reilly, T., & Gilbourne, D. (2003). Science and football: a review of applied research in the football codes. *Journal of Sports Sciences,* 21(9), 693-705. https://doi.org/10.1080/0264041031000102105

Reilly, T., Morris, T., & Whyte, G. (2009). The specificity of training prescription and physiological assessment: a review. *Journal of Sports Sciences*, 27(6), 575–589. https://doi.org/10.1080/02640410902729741

Riboli, A., Semeria, M., Coratella, G., & Esposito, F. (2020). Effect of formation, ball in play and ball possession on peak demands in elite soccer. *Biology of Sport,* 38(2), 195–205. https://doi.org/10.5114/biolsport.2020.98450

Riedl, D., Strauss, B., Heuer, A., & Rubner, O. (2015). Finale furioso: referee-biased injury times and their effects on home advantage in football. *Journal of Sports Sciences*, 33(4), 327–336. https://doi.org/10.1080/02640414.2014.944558

Riiser, A, Anderson, V, Sæterbakken, A, Ylvisaker, E, Moe, V. F. (2019). Running performance and position is not related to decision-making accuracy in referees. *Sports Medicine International Open*, 3(2), E66-E71. https://doi.org/10.1055/a-0958-8608

Ritchie, J., Basevitch, I., Rodenberg, R., & Tenenbaum, G. (2017). Situation criticality and basketball officials' stress levels. *Journal of Sports Sciences*, 35(21), 2080–2087. https://doi.org/10.1080/02640414.2016.1255770

Robertson, S. J., Burnett, A. F., & Cochrane, J. (2014). Tests examining skill outcomes in sport: a systematic review of measurement properties and feasibility. *Sports Medicine,* 44(4), 501–518. https://doi.org/10.1007/s40279-013-0131-0

Rohde, M., & Breuer, C. (2016). Europe's elite football: financial growth, sporting success, transfer investment, and private majority investors. *International Journal of Financial Studies*, 4(12), 1–20. https://doi.org/10.3390/ijfs4020012

Rollo, I., Impellizzeri, F. M., Zago, M., & Iaia, F. M. (2014). Effects of 1 versus 2 games a week on physical and subjective scores of subelite soccer players. *International Journal*

of Sports Physiology and Performance, 9(3), 425–431. https://doi.org/10.1123/ijspp.2013-0288

Russell, M., Rees, G., Benton, D., & Kingsley, M. (2011). An exercise protocol that replicates soccer match-play. *International Journal of Sports Medicine,* 32(7), 511–518. https://doi.org/10.1055/s-0031-1273742

Sæther, S. A., & Aspvik, N. P. (2014). Seasonal variation in objectively assessed physical activity among young Norwegian talented soccer players: A description of daily physical activity level. *Journal of Sports Science and Medicine*, 13(4), 964–968. https://www.jssm.org/hf.php?id=jssm-13-964.xml

Samuel, R. D. (2015). A psychological preparation framework for elite soccer referees: a practitioner's perspective. *Journal of Sport Psychology in Action*, 6(3), 170–187. https://doi.org/10.1080/21520704.2015.1065938

Samuel, R. D., Galily, Y., Guy, O., Sharoni, E., & Tenenbaum, G. (2019). A decision-making simulator for soccer referees. *International Journal of Sports Science and Coaching*, 14(4), 480–489. https://doi.org/10.1177/1747954119858696

Saw, A. E., Kellmann, M., Main, L. C., & Gastin, P. B. (2017). Athlete self-report measures in research and practice: Consideration for the discerning reader and fastidious practitioner. *International Journal of Sports Physiology and Performance*, 12(Suppl2), S2127–S2135. https://doi.org/10.1123/ijspp.2016-0395

Saw, A. E., Main, L. C., & Gastin, G. B. (2015). Role of a self-report measure in athlete preparation. *Journal of Strength and Conditioning Research*, 29(3), 685–691. https://doi.org/10.1519/JSC.0000000000000698

Scheadler, C. M., Garver, M. J., & Hanson, N. J. (2017). The gas sampling interval effect on V̇O2peak is independent of exercise protocol. *Medicine and Science in Sports and Exercise,* 49(9), 1911–1916. https://doi.org/10.1249/MSS.0000000000001301

Schmidt, S. L., Schmidt, G. J., Padilla, C. S., Simões, E. N., Tolentino, J. C., Barroso, P. R., Narciso, J. H., Godoy, E. S., & Costa Filho, R. L. (2019). Decrease in attentional resources after repeated bouts of high intensity exercise in Association football referees and assistant referees. *Frontiers in Psychology,* 10. https://doi.org/10.3389/fpsyg.2019.02014

Schnyder, U., Koedijker, J. M., Kredel, R., & Hossner, E. J. (2017). Gaze behaviours in offside decision-making in football: a field study. *German Journal of Exercise and Sport Research*, 47, 103–109. https://doi.org/10.1007/s12662-017-0449-0

Schwarz, W. (2011). Compensating tendencies in penalty kick decisions of referees in professional football: evidence from the German Bundesliga 1963-2006. *Journal of Sports Sciences*, 29(5), 441–447. https://doi.org/10.1080/02640414.2010.538711

Schweizer, G., Plessner, H., Kahler, D., & Brand, R. (2011). A video-based training method for improving soccer referees' intuitive decision-making skills. *Journal of Applied Sport Psychology*, 23, 429–442. https://doi.org/10.1080/10413200.2011.555346

Scoopa, V. (2008). Are subjective evaluations biased by social factors or connections? An econometric analysis of soccer referee decisions. *Empirical Economics*, 35(1), 123–140. https://doi.org/10.1007/s00181-007-0146-1

Shrier, I., Raglin, J. S., Levitan, E. B., Mittleman, M. A., Steele, R. J., & Powell, J. (2014). Procedures for assessing psychological predictors of injuries in circus artists: A pilot prospective study. *BMC Medical Research Methodology*, 14(1), 77. https://doi.org/10.1186/1471-2288-14-77

Sirotic, A. C., & Coutts, A. J. (2007). The reliability of physiological and performance measures during simulated team-sport running on a non-motorised treadmill. *Journal of Science and Medicine in Sport,* 11(5), 500–509. https://doi.org/10.1016/j.jsams.2007.04.008

Slack, L. A., Maynard, I. W., Butt, J., & Olusoga, P. (2013). Factors underpinning football officiating excellence: perceptions of English Premier League referees. Journal of

Applied Sport Psychology, 25(3), 298–315. https://doi.org/10.1080/10413200.2012.726935

Small, K., McNaughton, L., Greig, M., & Lovell, R. (2010). The effects of multidirectional soccer-specific fatigue on markers of hamstring injury risk. *Journal of Science and Medicine in Sport,* 13(1), 120–125. https://doi.org/10.1016/j.jsams.2008.08.005

Smith, M., Tallis, J., Miller, A., Clarke, N. D., Guimarães-Ferreira, L., & Duncan, M. J. (2016). The effect of exercise intensity on cognitive performance during short duration treadmill running. *Journal of Human Kinetics*, 51, 27–35. https://doi.org/10.1515/hukin-2015-0167

Sors, F., Tomé Lourido, D., Parisi, V., Santoro, I., Galmonte, A., Agostini, T., & Murgia, M. (2019). Pressing crowd noise impairs the ability of anxious basketball referees to discriminate fouls. *Frontiers in Psychology*, 10, 2380. https://doi.org/10.3389/fpsyg.2019.02380

Sparkes, A. C., & Smith, B. (2013). Qualitative research methods in sport, exercise and health: From process to product. London: Routledge.

Sperlich, P. F., Holmberg, H. C., Reed, J. L., Zinner, C., Mester, J., & Sperlich, B. (2015). Individual versus standardized running protocols in the determination of VO2max. *Journal of Sports Science and Medicine,* 14(2), 386-393. Available at: https://www.jssm.org/jssm-14-386.xml%3EFulltext

Spitz, J., Moors, P., Wagemans, J., & Helsen, W. F. (2018a). The impact of video speed on the decision-making process of sports officials. *Cognitive Research: Principles and Implications*, 3, 16. https://doi.org/10.1186/s41235-018-0105-8

Spitz, J., Put, K., Wagemans, J., Williams, A. M., & Helsen, W. F. (2016). Visual search behaviours of association football referees during assessment of foul play situations. *Cognitive Research: Principles and Implications*, 1(1), 12. https://doi.org/10.1186/s41235-016-0013-8

Spitz, J., Put, K., Wagemans, J., Williams, A. M., & Helsen, W. F. (2017). Does slow motion impact on the perception of foul play in football? *European Journal of Sport Science,* 17(6), 748–756. https://doi.org/10.1080/17461391.2017.1304580

Spitz, J., Put, K., Wagemans, J., Williams, A. M., & Helsen, W. F. (2018b). The role of domain-generic and domain-specific perceptual-cognitive skills in association football referees. *Psychology of Sport and Exercise*, 34, 47–56. https://doi.org/10.1016/j.psychsport.2017.09.010

Spitz, J., Wagemans, J., Memmert, D., Williams, A. M., & Helsen, W. F. (2021). Video assistant referee (VAR): The impact of technology on decision making in association football referees. *Journal of Sports Sciences,* 39(2), 147–153. https://doi.org/10.1080/02640414.2020.1809163

Stevens, T. G. A., de Ruiter, C. J., Twisk, J. W. R., Savelsbergh, G. J. P., & Beek, P. J. (2017). Quantification of in-season training load relative to match load in professional Duth Eredivisie football players. *Science and Medicine in Football*, 1(2), 117–125. https://doi.org/10.1080/24733938.2017.1282163

Taylor, J. M., MacPherson, T., Spears, I. R., & Weston, M. (2015). The effects of repeated-sprint training on field-based fitness measures: a meta-analysis of controlled and non-controlled trials. *Sports Medicine*, 45(6), 881–891. https://doi.org/10.1007/s40279-015-0324-9

Thatcher, R., & Batterham, A. M. (2004). Development and validation of a sport-specific exercise protocol for elite youth soccer players. *The Journal of Sports Medicine and Physical Fitness,* 44(1), 15–22. Available at: https://pubmed.ncbi.nlm.nih.gov/15181385/

Towlson, C., Midgley, A. W., & Lovell, R. (2013). Warm-up strategies of professional soccer players: practitioners' perspectives. *Journal of Sports Sciences*, 31(13), 1393–1401. https://doi.org/10.1080/02640414.2013.792946

Unkelbach, C., & Memmert, D. (2008). Game management, context effects, and calibration: the case of yellow cards in soccer. *Journal of Sport and Exercise Psychology*, 30(1), 95–109. https://doi.org/10.1123/jsep.30.1.95

Unkelbach, C., & Memmert, D. (2010). Crowd noise as a cue in referee decisions contributes to the home advantage. *Journal of Sport and Exercise Psychology*, 32(4), 483–498. https://doi.org/10.1123/jsep.32.4.483

Van Biemen, T., Koedijker, J., Renden, P. G., & Mann, D. L. (2018). The effect of blurred perceptual training on the decision making of skilled football referees. *Frontiers in Psychology*, 9, 1803. https://doi.org/10.3389/fpsyg.2018.01803

Varley, M. C., Elias, G. P., & Aughey, R. J. (2012). Current match-analysis techniques' underestimation of intense periods of high-velocity running. *International Journal of Sports Physiology and Performance,* 7(2), 183–185. https://doi.org/10.1123/ijspp.7.2.183

Vater, C., Roca, A., & Williams, A. M. (2016). Effects of anxiety on anticipation and visual search in dynamic, time-constrained situations. *Sport, Exercise, and Performance Psychology*. 5(3), 179–192. https://doi.org/10.1037/spy0000056

Vilar, L., Araújo, D., Davids, K., & Renshaw, I. (2012). The need for "representative task design" in evaluating efficacy of skills tests in sport: A comment on Russell, Benton and Kingsley (2010). *Journal of Sports Sciences*, 30(16), 1727–1730. https://doi.org/10.1080/02640414.2012.679674

Voight, M. (2009). Sources of stress and coping strategies of US soccer officials. *Stress and Health*, 25(1), 91–101. https://doi.org/10.1002/smi.1231

Wade, M. V. (2006). Likert-type scale response anchors. Clemson International Institute for Tourism & Research Development, department of parks. Recreation and Tourism Management, Clemson University.

Wallace, J. L., & Norton, K. I. (2014). Evolution of World Cup soccer final games 1966-2010: game structure, speed and play patterns. *Journal of Science and Medicine in Sport*, 17(2), 223–228. https://doi.org/10.1016/j.jsams.2013.03.016

Webb, T., Dicks, M., Thelwell, R, van der Kamp, J., & Rix-Lievre, G. (2020). An analysis of soccer referee experiences in France and the Netherlands: abuse, conflict, and level of support. *Sport Management Review*, 23(1), 52–65. https://doi.org/10.1016/j.smr.2019.03.003

Weston, M. (2015). Match performances of soccer referees: The role of sports science. *Movement & Sports Science*, 87, 113–117. https://doi.org/10.1051/sm/2014011

Weston, M. (2018). Training load monitoring in elite English soccer: a comparison of practices and perceptions between coaches and practitioners. *Science and Medicine in Football,* 2(3), 216–224. https://doi.org/10.1080/24733938.2018.1427883

Weston, M., Batterham, A. M., Castagna, C., Portas, M. D., Barnes, C., Harley, J., & Lovell, R. J. (2011a). Reduction in physical match performance at the start of the second half in elite soccer. *International Journal of Sports Physiology and Performance*, 6(2), 174–182. https://doi.org/10.1123/ijspp.6.2.174

Weston, M., Bird, S., Helsen, W., Nevill, A., & Castagna, C. (2006). The effect of match standard and referee experience on the objective and subjective match workload of English Premier League referees. *Journal of Science and Medicine in Sport*, 9(3), 256–262. https://doi.org/10.1016/j.jsams.2006.03.022

Weston, M., Castagna, C., Helsen, W. F., & Impellizzeri, F. M. (2009). Relationships among field-test measures and physical match performance in elite-standard soccer referees. *Journal of Sports Sciences*, 27(11), 1177–1184. https://doi.org/10.1080/02640410903110982

Weston, M., Castagna, C., Impellizzeri, F. M., Bizzini, M., Williams, M., & Gregson, W. (2012). Science and medicine applied to soccer refereeing: an update. *Sports Medicine*, 42(7), 615–631. https://doi.org/10.2165/11632360-000000000-00000

Weston, M., Castagna, C., Impellizzeri, F. M., Rampinini, E., & Abt, G. (2007). Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates. *Journal of Science and Medicine in Sport,* 10(6), 390–397. https://doi.org/10.1016/j.jsams.2006.09.001

Weston, M., Castagna, C., Impellizzeri, F. M., Rampinini, E., Breivik, S. (2010). Ageing and physical match performance in English Premier League soccer referees. *Journal of Science and Medicine in Sport*, 13(1), 96–100. https://doi.org/10.1016/j.jsams.2008.07.009

Weston, M., Drust, B., Atkinson, G., & Gregson, W. (2011b). Variability of soccer referees' match performances. *International Journal of Sports Medicine*, 32(3), 190–194. https://doi.org/10.1055/s-0030-1269843

Weston, M., Drust, B., & Gregson, W. (2011c). Intensities of exercise during match-play in FA Premier League referees and players. Journal of Sports Sciences, 29(5), 527–532. https://doi.org/10.1080/02640414.2010.543914

Weston, M., Gregson, W., Castagna, C., Breivik, S., Impellizzeri, F. M., & Lovell, R. J. (2011d). Changes in a top-level soccer referee's training, match activities, and psychology over an 8-year period: a case study. *International Journal of Sports Physiology and Performance*, 6(2), 281–286. https://doi.org/10.1123/ijspp.6.2.281

Weston, M., Siegler, J., Bahnert, A., McBrien, J., & Lovell, R. (2015). The application of differential ratings of perceived exertion to Australian Football League matches. *Journal of Science and Medicine in Sport,* 18(6), 704–708. https://doi.org/10.1016/j.jsams.2014.09.001

Williams, A. M., Ford, P. R., Eccles, D. W., & Ward, P. (2011). Perceptual-cognitive expertise in sport and its acquisition: implications for applied cognitive psychology. *Applied Cognitive Psychology*, 25(3), 432–42. https://doi.org/10.1002/acp.1710

Williams, J. D., Abt, G., & Kilding, A. E. (2010). Ball-sport endurance and sprint test (BEAST90): validity and reliability of a 90-minute soccer performance test. *Journal of*

Strength and Conditioning Research, 24(12), 3209–3218. https://doi.org/10.1519/JSC.0b013e3181bac356

Wilson, M. R., Wood, G., & Vine, S. J. (2009). Anxiety, attentional control, and performance impairment in penalty kicks. *Journal of Sport and Exercise Psychology*, 31(6), 761–775. https://doi.org/10.1123/jsep.31.6.761

APPENDICES

Appendix A: Participant Information Sheet (Example)

University of the West of Scotland
Division of Sport and Exercise
Participant Information Sheet

"Reliability, validity, and feasibility of a referee-specific treadmill protocol"

You are be ng nv ted to take part n a research study sponsored by the Scott sh Footba Assoc at on (SFA). Before you dec de whether or not to part c pate, t s mportant that you understand why the research s be ng undertaken and what t w nvo ve. P ease take t me to read the fo ow ng nformat on carefu y and fee free to d scuss t w th fr ends and re at ves f you w sh. If there s anyth ng you are uncerta n about or f you wou d ke further nformat on, then p ease do not hes tate to d scuss th s w th us. P ease take your t me to dec de whether or not you w sh to take part. Thank you for your nterest.

What is the purpose of the study?

The purpose of the present study s to exam ne the re ab ty, va d ty and feas b ty of a nove treadm protoco des gned to rep cate the nterm ttent act v ty of soccer referee ng.

Why have I been chosen?

You have been dent f ed as a potent a part c pant for th s nvest gat on because you have shown an nterest n vo unteer ng and match the nc us on cr ter a of the study. Spec f ca y, you are a moderate y tra ned ma e (aged between 18 – 45 years) who has prev ous footba exper ence (p ay ng, coach ng, referee ng, or spectat ng). Add t ona y, you are n good hea th and present no contrad ct ons to h gh- ntens ty phys ca act v ty.

Do I have to take part?

Part c pat on n the present study s ent re y vo untary and part c pants who dec de to take part are free to w thdraw at any po nt w thout prov d ng a reason.

What will happen to me if I take part?

If you dec de to take part, you w be asked to attend the aboratory on four occas ons w th each v s t ast ng around 60 m nutes. On the f rst v s t, a max ma f tness test w be performed on a motor sed treadm as we as a br ef 15-m nute fam ar sat on of the exper menta protoco .

On arr va to the aboratory for the subsequent three v s ts, you w be asked to comp ete two short quest onna res that seek to assess your eve of mot vat on, se f-conf dence, and anx ety for the upcom ng sess ons. Fo ow ng th s, you w be asked to perform a 30-m nute nterm ttent exerc se protoco compr s ng of per ods of stand ng, wa k ng, jogg ng, cru s ng, and h gh- ntens ty runn ng. At var ous t me-po nts throughout the exerc se protoco , you w be asked to perform a br ef footba -re ated dec s on-mak ng task, wh st var ous outcomes

measures (heart rate, oxygen consumpt on, b ood actate, perce ved exert on, and se ect ve attent on) w be co ected. Add t ona y, you w be nv ted to part c pate n a short nterv ew fo ow ng the f na exper menta tr a whereby you w be asked to appra se the sess on.

What are the possible disadvantages and risks of taking part?

As w th a forms of exerc se (max ma or sub-max ma), there are hea th r sks nvo ved such as muscu oske eta njury; however, these are m n ma w th n hea thy adu t popu at ons and the appropr ate steps w be taken to ensure the safety and comfort of a part c pants.

What are the possible benefits of taking part?

A part c pants w rece ve extens ve feedback concern ng the r resu ts. Spec f ca y, part c pants w have the r max ma oxygen uptake ($\dot{V}O_{2max}$) and max ma heart rate (HR_{max}) determ ned wh ch can be used to nd v dua se and mprove the eff c ency of subsequent tra n ng sess ons. Such tests are typ ca y very cost y w th a standard max ma test cost ng upwards of £150.

What if something goes wrong?

If you are harmed by tak ng part n th s research project, there are no spec a compensat on arrangements. If you are harmed due to someone's negligence then you may have grounds for ega act on but you may have to pay for t. The pr nc pa nvest gator(s) w prov de bas c fe support nc ud ng chest compress ons and vent at on unt emergency med ca staff are on hand. You may want to consu t your GP f you are exper enc ng any s de effects from tak ng part n the study and shou d a so nform the pr nc pa nvest gator.

Will my participation in the study be kept confidential?

A nformat on re at ng to you that s co ected dur ng the course of th s research w be treated as str ct y conf dent a amongst the research team. Data re at ng to you w be encrypted and stored on a password protected computer. At no stage w your name appear on any data or n any subsequent pub cat on; you w be dent f ab e on y by a un que part c pant dent f cat on code.

What should I do if I have a complaint to make about any aspect of the study?

If you are unhappy w th any aspect of the study then you shou d address th s w th the research team. A contact deta s are prov ded at the end of th s part c pant nformat on sheet.

What will happen to the results of the research study?

Resu ts w be stored on a password protected computer. Group resu ts w be subm tted for pub cat on w th n a peer-rev ewed sc ent f c journa . If successfu ,

you w be sent a copy of the fu art c e. You w not be dent f ed w th n the pub cat ons.

Whom may I contact if I require further information regarding this study?

If you w sh to f nd out more about th s study, p ease do not hes tate to contact:

Gary McEwan
PhD Cand date
Un vers ty of the West of Scot and
Ema : [redacted]

OR

Dr Ros e Arthur
Lecturer n Sport and Exerc se Psycho ogy
Un vers ty of the West of Scot and
Ema : [redacted]

OR

Professor Chr s Easton
Professor n Exerc se Phys o ogy
Un vers ty of the West of Scot and
Ema [redacted]

OR

Professor V sh Unn than
Professor of Paed atr c Exerc se Phys o ogy
Un vers ty of the West of Scot and
Ema : [redacted]

Thank you very much for your interest in our research.

Appendix B: Informed Consent Form (Example)

University of the West of Scotland

Division of Sport and Exercise

Participant Consent Form

"Reliability, validity, and feasibility of a referee-specific treadmill protocol"

I .....

g ve my consent to take part n the research study ent t ed ***"Reliability, validity, and feasibility of a referee-specific treadmill protocol"***. I have been made fu y aware of the potent a r sks nvo ved w th the research procedures, a of wh ch are out ned n deta n the separate part c pant nformat on sheet prov ded, and have had the opportun ty to c ar fy any concerns I have. I understand that f I w sh to w thdraw my part c pat on from the study then I may do so at any po nt w thout prov d ng a reason.

Participant Signature ……

Print Name ……

Date ……

Researcher Signature ……

Researcher Name ……

Date ……

Appendix C: Interview guide (Chapter 3)

University of the West of Scotland
Division of Sport and Exercise
Interview Guide

"The Referee Training Activity Questionnaire (RTAQ): Towards a better understanding of the training practices of soccer officials"

Firstly, I'd like to thank you again for your participation in the study. Before we start, I wou d just ke to make you aware that th s nterv ew s be ng recorded to ensure that no mportant nformat on s m ssed. Are you happy for th s to be recorded?

We are conduct ng th s nterv ew n an attempt to ga n a better understand ng of the attr butes and character st cs that nd v dua s nvo ved w th n referee ng deem mportant to successfu referee ng performance. We are a so nterested n bu d ng a comprehens ve st of tra n ng act v t es that referees common y part c pate n. In th s regard, I have a number of quest ons that I wou d ke to ask you. If you have d ff cu ty n reca ng any nformat on or are unsure about how to respond to any of the questions then don't worry but if you can just let me know. Also, please be aware that this isn't an assessment of your ability or su tab ty to referee and that there are no r ght or wrong answers – we are s mp y ook ng to obta n your op n on as you are a great source of exper ence and know edge n th s area. F na y, a responses w be kept comp ete y conf dent a w th n the research team so your honesty s someth ng we wou d be very gratefu for.

Before we get started, do you have any quest ons?

Ice-breakers

1. Can you te me a b t about your ro e as a referee w th the SFA p ease?

Practices and perceptions

1. What have been your key areas of focus w th regards to your tra n ng over the past year?

2. What attr butes or character st cs wou d you say are requ red n order to be a successfu referee? Prompts f requ red.
    a. Phys ca -phys o og ca
    b. Perceptua -cogn t ve (dec s on-mak ng)
    c. Psycho og ca
    d. On the p tch behav our and game management
    e. Know edge of the aws of the game

3. Cou d you te me some more deta regard ng the tra n ng prov ded by the SFA (e.g. week y assoc at on tra n ng sess ons, camps, other tra n ng)? Prompts f requ red.
    a. Phys ca – phys o og ca
        . Focus of the sess ons (object ves)

. Content (what has been done)
. Cond t ons (env ronment – where)
v. Resources (access – fac t es requ red)

b. Perceptua – cogn t ve (dec s on-mak ng)
c. Psycho og ca
d. On the p tch behav our and game management
e. Know edge of the aws of the game

4. Can you te me about any assoc at on tra n ng sess ons whereby a ho st c approach has been taken? For examp e, where dec s on-mak ng tasks have been ncorporated to phys ca cond t on ng sess ons. Prompts f requ red.
   a. How often
   b. Content (what has been done)
   c. Cond t ons (env ronment – where)
   d. Support

5. What other tra n ng have you taken part n (outs de of that prov ded by the SFA)? Prompts f requ red.
   a. Focus of the sess ons (object ves)
   b. Content (what has been done)
   c. Cond t ons (env ronment – where)
   d. Support

6. Look ng back to prev ous years, how wou d you say the tra n ng of referees has changed over the years? Prompts f requ red.
   a. Focus of the sess ons (object ves)
   b. Content (what has been done)
   c. Cond t ons (env ronment – where)
   d. Support

Barriers

1. What factors or barr ers (f any) do you fee restr ct the type of tra n ng you are ab e to part c pate n? *Prompts if required*
   a. Fac t es
   b. Equ pment
   c. Support (coaches or peers)
   d. Know edge

Decision-making

The f na sect on s re ated to your dec s on-mak ng dur ng compet t on and w seek to exam ne the types and contexts where dec s on-mak ng s most prob emat c.

1. Th nk ng back to a part cu ar t me where you have had to make a d ff cu t dec s on, what were the deta s of th s? What made th s dec s on d ff cu t to make? *Prompts if required*
   a. Type of dec s on

b. Context (nterna and externa factors)

2. What factors (f any) do you fee mpact most upon your ab ty to make the correct dec s on dur ng match p ay? *Prompts if required*
   a. Interna
   b. Externa

3. Are you ab e to out ne the type and context of dec s on that you fee s most open to mprovement through de berate pract ce? *Prompts if required*

[END OF INTERVIEW]

Appendix D: Referee Training Activity Questionnaire (RTAQ)

The Referee Training Activity Questionnaire (RTAQ)

McEwan, G. P., Unn than, V., Easton, C., & Arthur, R. (2020). Deve opment and va dat on of the Referee Tra n ng Act v ty Quest onna re (RTAQ): Towards a better understand ng of the tra n ng pract ces of soccer off c a s. *Journal of Sports Sciences,* 38(20), 2782–2793.

1: DEMOGRAPHIC INFORMATION

The fo ow ng sect on focuses on your demograph c nformat on such as persona deta s, career m estones and recent performances.

Note: It must be stressed that a nformat on co ected w be kept comp ete y conf dent a amongst the research team.

PERSONAL DETAILS

What s your date of b rth?

C ck or tap here to enter text.

What s your gender?

C ck or tap here to enter text.

REFEREEING CAREER

What s your pr mary ro e?

C ck or tap here to enter text.

When (rough est mat on of date) d d you f rst become a footba match off c a ?

C ck or tap here to enter text.

What s your pr mary mot vat on to off c ate? *Please be as specific as possible.*

C ck or tap here to enter text.

What referee ng goa s have you set yourse f for the next f ve years? *Please be as specific as possible.*

C ck or tap here to enter text.

Wh ch category are you current y sted to?

C ck or tap here to enter text.

Wh ch referee assoc at on do you be ong?

C ck or tap here to enter text.

What s the h ghest eve at wh ch you have off c ated?

C ck or tap here to enter text.

2: GENERAL TRAINING PRACTICES

The fo ow ng sect on focuses on your genera tra n ng pract ces and seeks nformat on concern ng the tra n ng support you rece ve, the env ronment n wh ch you typ ca y tra n, and the object ves of your tra n ng sess ons.

DOMAIN-SPECIFIC ACTIVITIES

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have engaged n each of the fo ow ng act v t es.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
Match off c at ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Coach ed nd v dua tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Coach ed group tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Se f ed nd v dua tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Se f ed group tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Peer ed p ay	C ck or tap here to enter text.	C ck or tap here to enter text.

TRAINING ENVIRONMENT

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have tra ned n each of the fo ow ng env ronments.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
Gym based	C ck or tap here to enter text.	C ck or tap here to enter text.
Poo based	C ck or tap here to enter text.	C ck or tap here to enter text.
P tch based	C ck or tap here to enter text.	C ck or tap here to enter text.
Track based	C ck or tap here to enter text.	C ck or tap here to enter text.
Road based	C ck or tap here to enter text.	C ck or tap here to enter text.
C assroom based	C ck or tap here to enter text.	C ck or tap here to enter text.

TRAINING OBJECTIVE

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have engaged n each of the fo ow ng types of tra n ng sess on.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
Sk deve opment	C ck or tap here to enter text.	C ck or tap here to enter text.
F tness deve opment	C ck or tap here to enter text.	C ck or tap here to enter text.
Injury prevent on	C ck or tap here to enter text.	C ck or tap here to enter text.
Match preparat on	C ck or tap here to enter text.	C ck or tap here to enter text.
Recovery	C ck or tap here to enter text.	C ck or tap here to enter text.

3: SPECIFIC TRAINING PRACTICES

Th s sect on focuses on your spec f c tra n ng pract ces and seeks nformat on concern ng how often you tra n var ous attr butes of potent a mportance to referee ng performance as we as the amount of t me (frequency and durat on) that you engage n spec f c tra n ng act v t es.

PHYSICAL TRAINING

Us ng the 5 po nt L kert sca e be ow *(less than once a month to more than once a week)*, p ease nd cate how often you engage n tra n ng that s focused on deve op ng each of the fo ow ng phys ca attr butes.

	Less than once a month	Once a month	Once every two weeks	Once a week	More than once a week
Aerob c endurance	☐	☐	☐	☐	☐
Repeated spr nt ab ty	☐	☐	☐	☐	☐
Speed endurance	☐	☐	☐	☐	☐
Speed acce erat on	☐	☐	☐	☐	☐
Max ma spr nt speed	☐	☐	☐	☐	☐
Ag ty	☐	☐	☐	☐	☐
Strength	☐	☐	☐	☐	☐
Power	☐	☐	☐	☐	☐
Muscu ar endurance	☐	☐	☐	☐	☐
F ex b ty	☐	☐	☐	☐	☐

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have engaged n each of the fo ow ng tra n ng act v t es.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
H gh ntens ty nterva tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Cont nuous aerob c tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Speed endurance tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Repeated spr nt ab ty tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Speed and ag ty tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Res stance tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Game based act v t es	C ck or tap here to enter text.	C ck or tap here to enter text.
Cross tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Mob ty tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.

DECISION-MAKING TRAINING

Us ng the 5 po nt L kert sca e be ow *(less than once a month to more than once a week)*, p ease nd cate how often you engage n tra n ng that s focused on deve op ng each of the fo ow ng dec s on mak ng attr butes.

	Less than once a month	Once a month	Once every two weeks	Once a week	More than once a week
Dec s on mak ng when phys ca y t red	☐	☐	☐	☐	☐
Dec s on mak ng under stressfu cond t ons	☐	☐	☐	☐	☐
V sua sk s	☐	☐	☐	☐	☐
Spat a awareness	☐	☐	☐	☐	☐
Ant c pat on	☐	☐	☐	☐	☐
Concentrat on	☐	☐	☐	☐	☐
Mu t task ng sk s	☐	☐	☐	☐	☐
Speed of thought	☐	☐	☐	☐	☐
React on speed	☐	☐	☐	☐	☐

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have engaged n each of the fo ow ng tra n ng act v t es.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
On f e d referee ng s mu at ons	C ck or tap here to enter text.	C ck or tap here to enter text.
V deo c p ana yses	C ck or tap here to enter text.	C ck or tap here to enter text.
Interact ve v deo based dec s on mak ng tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Other	C ck or tap here to enter text.	C ck or tap here to enter text.

PSYCHOLOGICAL SKILLS TRAINING

Us ng the 5 po nt L kert sca e be ow *(less than once a month to more than once a week)*, p ease nd cate how often you engage n tra n ng that s focused on deve op ng each of the fo ow ng psycho og ca attr butes.

	Less than once a month	Once a month	Once every two weeks	Once a week	More than once a week
Conf dence	☐	☐	☐	☐	☐
Mot vat on	☐	☐	☐	☐	☐
Emot ona contro	☐	☐	☐	☐	☐
Res ence	☐	☐	☐	☐	☐
Menta toughness	☐	☐	☐	☐	☐

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have engaged n each of the fo ow ng tra n ng act v t es.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
Se f ta k	C ck or tap here to enter text.	C ck or tap here to enter text.
Emot ona contro	C ck or tap here to enter text.	C ck or tap here to enter text.
Automat c ty	C ck or tap here to enter text.	C ck or tap here to enter text.
Goa sett ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Imagery	C ck or tap here to enter text.	C ck or tap here to enter text.

Act vat on	C ck or tap here to enter text.	C ck or tap here to enter text.
Re axat on	C ck or tap here to enter text.	C ck or tap here to enter text.
Negat ve th nk ng	C ck or tap here to enter text.	C ck or tap here to enter text.
D stract b ty	C ck or tap here to enter text.	C ck or tap here to enter text.

TECHNICAL SKILLS TRAINING

Us ng the 5 po nt L kert sca e be ow *(less than once a month to more than once a week)*, p ease nd cate how often you engage n tra n ng that s focused on deve op ng each of the fo ow ng techn ca attr butes.

	Less than once a month	Once a month	Once every two weeks	Once a week	More than once a week
Know edge of the aws of the game	☐	☐	☐	☐	☐
Tact ca awareness	☐	☐	☐	☐	☐
Read ng the game	☐	☐	☐	☐	☐
F ag techn que	☐	☐	☐	☐	☐
Referee spec f c movement	☐	☐	☐	☐	☐
On f e d pos t on ng	☐	☐	☐	☐	☐

P ease nd cate the number of t mes *(within the past two weeks)* and the average durat on per sess on that you have engaged n each of the fo ow ng tra n ng act v t es.

	Frequency (number of sess ons w th n the past two weeks)	Durat on (m nutes per sess on)
Techn ca sk s tra n ng	C ck or tap here to enter text.	C ck or tap here to enter text.
Tact ca research	C ck or tap here to enter text.	C ck or tap here to enter text.
Laws of the Game study	C ck or tap here to enter text.	C ck or tap here to enter text.
Game ana ys s	C ck or tap here to enter text.	C ck or tap here to enter text.

GAME-MANAGEMENT TRAINING

Us ng the 5 po nt L kert sca e be ow *(less than once a month to more than once a week)*, p ease nd cate how often you engage n tra n ng focused on the deve opment of each of the fo ow ng game management attr butes.

	Less than once a month	Once a month	Once every two weeks	Once a week	More than once a week
Body anguage	☐	☐	☐	☐	☐
Effect ve commun cat on sk s	☐	☐	☐	☐	☐
Teamwork	☐	☐	☐	☐	☐
Leadersh p	☐	☐	☐	☐	☐

Appendix E: Health-Screening Questionnaire

Medical Questionnaire for Physiological Testing

Before any physiological tests are carried out on you, we will have to check whether you are in a satisfactory condition to undergo strenuous exercise. We would therefore like you to fill in the following questionnaire about yourself. All information will be treated as strictly confidential.

Name _____ Date of Birth ____/_____/_____

Specialist Sport (if any) _____ Male ☐ Female ☐

PLEASE TICK ONE ANSWER ONLY

1. Have you ever had a heart attack, coronary revascularisation surgery or a stroke?

No □ **Yes** □

2. Has your doctor ever told you that you have heart trouble or vascular disease?

No □ **Yes** □

3. Has your doctor ever told you that you have a heart murmur?

No □ **Yes** □

4. Do you ever suffer from pains in your chest, especially with exercise?

No □ **Yes** □

5. Do you ever get pains in your calves, buttocks or at the back of your legs during exercise which are not due to soreness or stiffness?

No □ **Yes** □

6. Do you ever feel faint or have spells of severe dizziness, particularly with exercise?

No □ **Yes** □

7. Do you experience swelling or accumulation of fluid about the ankles?

No □ **Yes** □

8. Do you ever get the feeling that your heart is suddenly beating faster, racing or skipping beats, either at rest or during exercise?

No □ **Yes** □

9. Do you have chronic obstructive pulmonary disease, interstitial lung disease, or cystic fibrosis?

No □ **Yes** □

10. Have you ever had an attack of shortness of breath that developed when you were not doing anything strenuous, at any time in the last 12 months?

No □ **Yes** □

11. Have you ever had an attack of shortness of breath that developed after you stopped exercising, at any time in the last 12 months?

No □ **Yes** □

12. Have you ever been woken at night by an attack of shortness of breath, at any time in the last 12 months?

No □ **Yes** □

13. Do you have diabetes [Type I (Insulin-dependent diabetes mellitus (IDDM)) or Type II (non-insulin-dependent diabetes mellitus (NIDDM)]? If so, do you have trouble controlling your diabetes?

No □ **Yes** □

14. Do you have any ulcerated wounds or cuts on your feet that do not seem to heal?

No □ **Yes** □

15. Do you have any liver, kidney or thyroid disorders?

No □ **Yes** □

16. Do you experience unusual fatigue or shortness of breath with usual activities?

No □ **Yes** □

17. Is there any other physical reason or medical condition, or are you taking any medication(s) which could prevent you from undertaking an exercise program, or that you are concerned about? If so, please explain.*(see notes)

No □ **Yes** □

18. Have you ever suffered from a viral infection in the last two weeks?

No □ **Yes** □

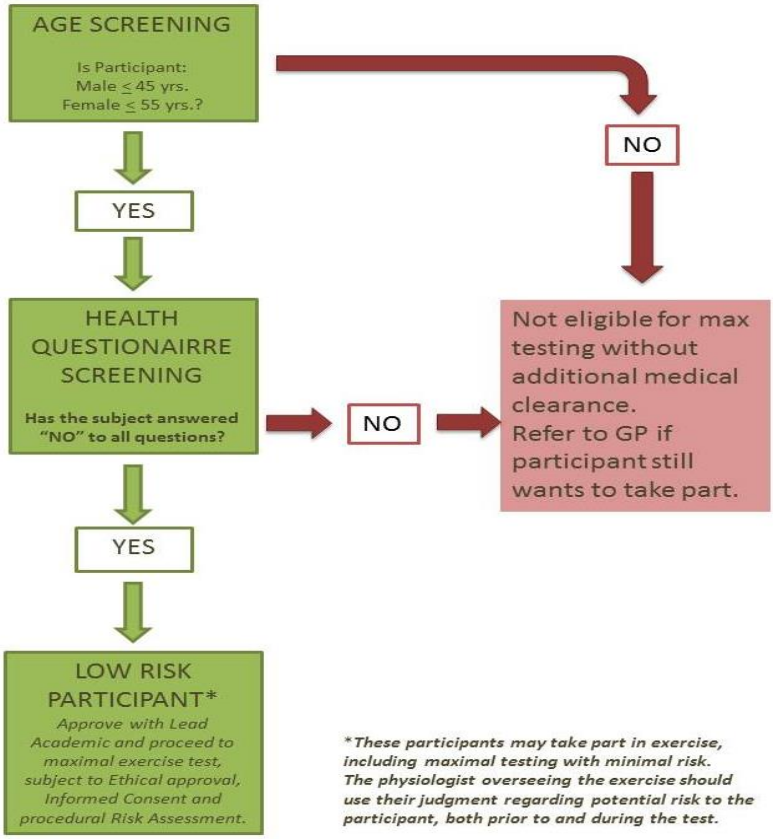
*** NOTES: Some of these conditions might include a history of blood clotting, osteoporosis, bone fractures or serious musculoskeletal disorders, or if they have recently lost a weight without trying to. Other types of conditions might include psychiatric disorders, later-stage pregnancy or those with a history of health problems during pregnancy. Those people taking medication(s) for medical conditions listed may also need medical clearance.**

- **I have read, understood and completed this questionnaire. Any questions I had were answered to my full satisfaction. If there will be a change of status to any of the conditions above I will inform the researcher or the senior academic involved IMMEDIATELY.**

NAME _____SIGNATURE_____DATE____/____/_____

References: (*1) Sports Medicine Australia (SMA) pre-exercise screening system 2005, Australian Government, Department of Health and Ageing (2) American College of Sports Medicine (ACSM) (2000). ACSM's guidelines for exercise testing and prescription (6th ed). New York, Lippincott Williams & Wilkins. (3) Australian Institute of Health and Welfare (AIHW) (2004). Heart, stroke and vascular diseases Australian facts 2004. Canberra, AIHW and National Heart Foundation of Australia.(4) National Heart Foundation (NHF) (2005). Physical activity recommendations for people with cardiovascular disease. Sydney, National Heart Foundation of Australia.(5) Olds, T. and Norton, K. (1999). Pre-exercise health screening guide. Champaign, Ill, Human Kinetics.*

UWS Sport and Exercise Laboratories Screening Flow Chart



To be completed by test supervisor: **MUST BE SIGNED BEFORE THE SUBJECT STARTS EXERCISE**
Checked by_____ **Signed**_____ **Date**____/_____/_____

Appendix F: Confirmation of Ethical Approval (Chapter 5)

UNIVERSITY OF THE WEST of SCOTLAND
UWS

10/05/2019

Dear Gary McEwan,

Your application 8202: Reliability, validity, and feasibility of a referee-specific treadmill protocol, submission reference 6591, has been approved, with conditions, by the Health and Life Sciences SEC **You may proceed with your study provided you meet the conditions outlined below**;

Title	Comment
Please give a full summary of the purpose, justification, design and methodology of the planned study: (word limit 1000 words)	Comment- Consider the recruitment of participants who have completed an entry level referee course?
Please upload a copy of the Questionnaire or Interview Schedule	Comment- Interview Schedule: Do you need to document the subject's confirmation that they agree for interview to be videoed?
	Comment- Consider rephrasing sentence 'Finally, all responses will be kept completely confidential within the research team so your honesty is something we would be very grateful for ' (i e implies they might be dishonest)
Please upload a copy of the Participant Information Sheet(s)	Condition- Age range (17-45) on PI sheet differs from elsewhere in application/poster (18-45), ensure they are the same

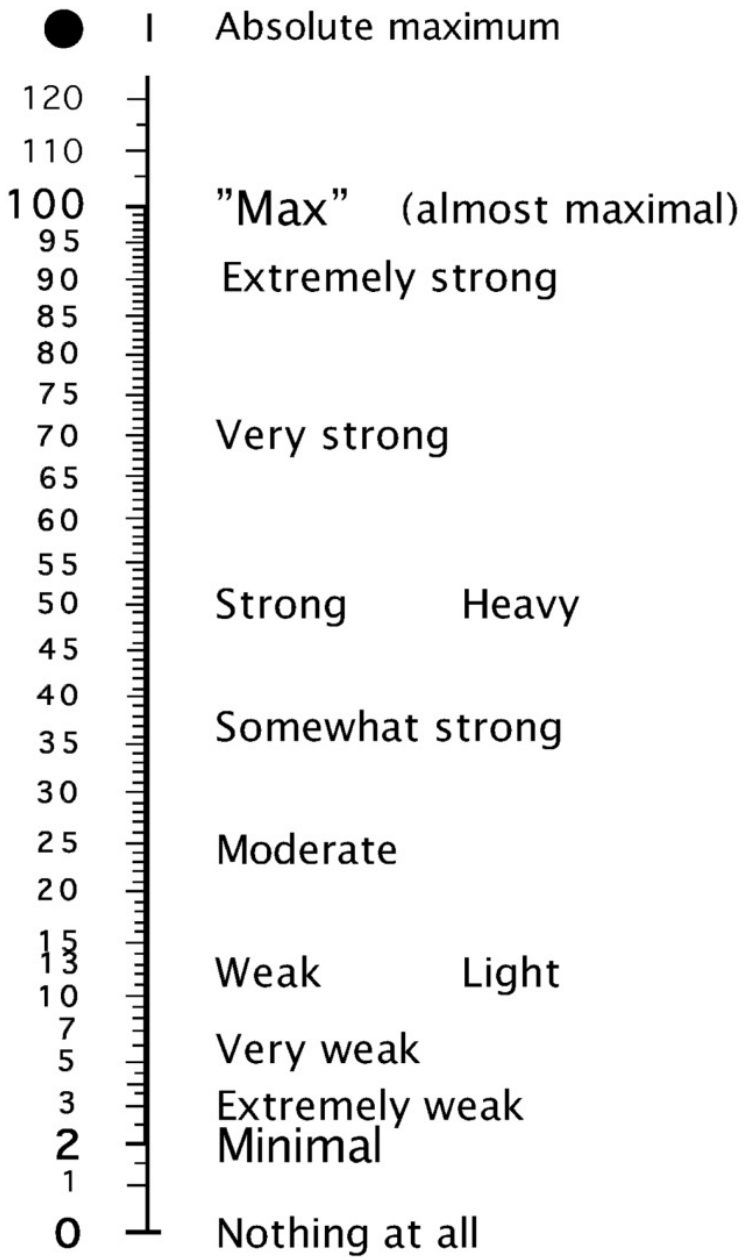
If you wish to make any significant changes to your study you must seek the committee's approval before actioning them

Good luck with your research

[redacted]
Dr Gary Boyd

Page 1 of 1

Appendix G: CR100 Scale



Borg centiMax (CR100) scale

English

Appendix H: Interview Guide (Chapter 5)

University of the West of Scotland
Division of Sport and Exercise

"Reliability, validity, and feasibility of a referee-specific treadmill protocol"

F rst y, I'd ke to thank you aga n for your part c pat on n th s study. Before we start, I wou d just ke to make you aware that th s nterv ew s be ng recorded to ensure that no mportant nformat on s m ssed. Are you happy for th s to be recorded?

We are conduct ng th s nterv ew n order to ga n your appra sa of the sess on you have just comp eted. We are a so nterested n your eva uat on of how the exerc se protoco may (or may not) have mpacted upon your performance dur ng the dec s on mak ng task. In th s regard, I have a number of quest ons that I wou d ke to ask you. If you have d ff cu ty n reca ng any nformat on or are unsure about how to respond to any of the quest ons then don't worry but f you can just et me know. A so, p ease be aware that th s sn't an assessment of your ab ty and that there are no r ght or wrong answers we are s mp y ook ng to obta n your op n on. F na y, a responses w be kept comp ete y conf dent a w th n the research team so your honesty s someth ng we wou d be very gratefu for.

Before we get started, do you have any quest ons?

Ice-breakers

1. How d d you f nd that sess on?

2. Have you ever done anyth ng ke that before?

Exercise protocol

1. What s your assessment of the exerc se protoco ?

    a. What s your assessment of the exerc se protoco n re at on to the frequent changes of act v ty?

2. What s your assessment of the dec s on mak ng task?

    a. What mpact (f any) do you fee the exerc se protoco had on your dec s on mak ng performance?

[END OF INTERVIEW]

Appendix I: Confirmation of ethical approval (Chapter 6)

UNIVERSITY OF THE WEST of SCOTLAND
UWS

20/03/2019

Dear Gary McEwan,

Your application 6623: Decision-making accuracy of elite soccer referees in relation to markers of internal and external physiological load, submission 6512, has been approved by the Health and Life Sciences SEC You may now proceed with your study If you wish to make any significant changes to the study you must seek the committee's approval before actioning them

Good luck with your research

Dr Gary Boyd