An Integrated Conceptual Framework of Decision-Making in Soccer Refereeing

Roy David Samuel, Gershon Tenenbaum and Yair Galily

ABSTRACT

Decision-making (DM) is a critical aspect of soccer refereeing. Referees make numerous repetitive decisions, many of which are dependent on appropriate field location and effective interaction with their assistants. Considering the complexity of the sequential DM process, it is not surprising that referees exhibit high decision error rates. Most extant research has focused on the underlying mechanism of the DM process as well as on various influential factors. Less attention has been given to models that consider the range of influential factors which affect the DM process. This article, therefore, presents a new conceptual framework of sequential DM for soccer referees that relies on Tenenbaum’s (2003). Expert athletes: An integrated approach to decision making. In J. Starkes, & A. Ericsson (Eds.), Expert performance in sports: Advances in research on sport expertise (pp. 192–218). Human Kinetics) conceptual framework and the existing literature. We start by reviewing research on DM in soccer refereeing and then discuss conceptual considerations. We then present the new conceptual framework. This model incorporates a range of factors, including physical fitness and fatigue, field positioning, visual attention, contextual factors, game management, expertise, psychological factors, and team factors. Data obtained from a sample of 20 elite referees and assistants demonstrated the face validity of the new model. Finally, we introduce implications for training referees’ DM skills and highlight avenues for future research and conceptual developments.

[...] They were not better than us. They found the net early and the second goal came from a penalty they didn’t award … The officiating was crazy. There were clear penalties, on Nicolas Otamendi, on Sergio Agüero. He [the referee] was on their side. In every divided ball or dispute, he inclined the pitch their way. It’s not an excuse, but the truth is that, in this Copa, they kept on blowing for stupid things, for handballs, penalties. But today, they didn’t even go to VAR when there were clear plays that should have been looked at. Argentina’s Lionel Messi complains during the July 3rd, 2019 Copa América semi-final vs Brazil (The Guardian, 2019).

CONTACT Roy David Samuel roydsamuel@gmail.com Interdisciplinary Center (IDC) Herzliya, Baruch Ivcher School of Psychology, 8 Ha’universuta St., Herzliya, 4610101, Israel Kibbutzim College of Education Technology and the Arts, Department of Physical Education and Movement, Tel-Aviv, Israel

© 2020 International Society of Sport Psychology
Kick-Off

The referee, an expert individual, is accountable for ruling over the game from an impartial point of view, making split-seconds decisions and enforcing the rules of the game over the players. Yet, his (or her) errors may have negative consequences from an economical, psychological and social perspective for clubs and fans, and also for athletes and teams (Guillén & Feltz, 2011). Indeed, the referee’s job is, before and after, to maintain the fairness and increase the integrity of many sports including the much-loved global game of soccer (Russell et al., 2019). Thus, it is not surprising that The International Football Association Board (The IFAB) unanimously approved the use of the video assistant referee (VAR) system on 3 March 2018. The meeting, chaired by FIFA President Gianni Infantino, “represents a new era for football with video assistance for referees helping to increase integrity and fairness in the game.” The aim of the VAR, according to Infantino, is to correct “clear and obvious errors” and deal with “serious missed incidents” in defined match-changing situations (goal, penalty/no penalty, direct red card and mistaken identity for disciplinary sanctions, IFAB, 2018).

Indeed, VAR might reduce the number of fouls and yellow cards, as the players become more cautious of their actions (e.g., Lago-Peñas et al., 2019), yet much still depends on the referee’s DM. Lessons from similar sports and devices raise major concerns. For example, at a cost of 300,000 USD per match in cricket, video review only resulted in a 2.7% increase in umpiring accuracy, or 100,000 USD per 1% of improvement (Borooah, 2016). One might argue that at such return-on-investment, it might be worth investing more in referees’ decision-making (DM) practice rather, or side, to such technological assistance. Nevertheless, several conceptual and practical challenges still remain for advancing refereeing DM performance. Hence, the present article is aimed at establishing a solid basis for future research and practice, incorporating into a single framework the vast range of factors which have the potential to influence referees’ DM within real match contexts.

DM is considered a fundamental aspect of soccer refereeing (Helsen & Bultynck, 2004; Pina et al., 2018). Moreover, referees perceive DM skills as the most imperative characteristic of their performance (MacMahon et al., 2007; Schnyder & Hossner, 2016; Slack et al., 2013; Voight, 2009). At the elite level, referees interact, process, and make decisions under considerable pressure (Dawson, 2012; Samuel, 2015; Schnyder & Hossner, 2016). Research on referees’ DM mainly focused on the underlying mechanism (e.g., Lane et al., 2006; Russell et al., 2019) and the factors that influence it (e.g., Jones et al., 2002; Mallo et al., 2012; Picazo-Tadeo et al., 2017; Plessner & Betsch, 2001; Spitz et al., 2018b; van Quaquebeke & Giessner, 2010). The complexity of the DM process explains why soccer referees maintain high error rates. The data on foul decisions, for example, indicate accuracy rates of 50–93.1 percent (e.g., Catteeuw et al., 2009; Hossner et al., 2019; MacMahon et al., 2007; Mallo et al., 2012; Mascarenhas et al., 2009; Schweizer et al., 2011; Spitz et al., 2016; Spitz et al., 2018a, 2018b). This means that, in some studies, the accuracy rates were comparable to mere chance.

Less attention was given to developing DM models that consider a range of influential factors. While there are certain conceptual publications on refereeing DM (Mascarenhas et al., 2005; Plessner et al., 2009; Plessner & Haar, 2006), a comprehensive framework is still missing. Developing comprehensive models is important for two reasons. First, to provide some benchmark for this research area; where we are and where we need to
develop. Second, to integrate the various lines of research on soccer refereeing into a model that can guide practitioners and refereeing coaches how to facilitate referees’ DM performance. Therefore, in this article we present a new conceptual framework of sequential DM for soccer referees. We start by reviewing research on DM in soccer refereeing (with a focus on head referees rather than assistants). Our focus centres on conceptualising the refereeing DM performance. Thus, we do not present a systematic review of the general area of soccer refereeing (e.g., referees’ motivation or career development), as provided by Pina et al. (2018), but rather focus on factors and processes which are imperative for capturing referees’ DM (i.e., a narrative review; see Ferrari, 2015). We then present a new conceptual framework, including the findings of a pilot study. Consequently, future research avenues and implications for training are suggested.

**Research on decision-making in soccer refereeing**

Helsen and Bultynck (2004) examined the DM of referees at the Euro 2000 tournament and found that, on average, they performed at 85% ± 5% of their maximal heart rate. The mean number of observable decisions referees were making per match was 137, and that the total number of decisions (including decisions not to interfere with the course of play) was 200, most of which were based on communication with their assistants. Decisions were made in relation to various match events and infringements, including fouls and penalties, out of play and corners, approval or disapproval of goals, offside situations, and disciplinary actions (i.e., yellow and red cards). More recently, Castillo et al. (2019b) analysed the decision-making of referees and assistant referees in 16 matches played in an international tournament and found that 3605 decisions were made in total, pertaining to various match events and infringements; 225.31 decisions per match.

The referee’s positioning and distance from the various match events is an important factor which attracted research attention. For example, Mallo et al.’s (2012) analysis of foul incidents from the 2009 FIFA Confederations Cup showed that the distance of the referee to the match event influenced the accuracy of the decision; thus, in the central area of the field, the highest accuracy rates (83%) were found when referees viewed the match events from distances of 11–15 m. Those authors suggested that being too far away from the match event increased the risk of missing important details, whereas being too close debilitated a comprehensive view of the entire sequence. A similar effect was not evident in the lateral areas of the field.

Hossner et al. (2019) suggested that referees’ physical fitness could indirectly influence their DM performance, through debilitating an optimal position in relation to crucial specifics of the current situation. These authors video metrically analysed the referees’ position relative to 2290 foul-play match infringements in all 64 matches of the 2014 FIFA World Cup. They found that the total error rate was relatively low – 93.1%. Furthermore, on average, the referees were positioned at a distance of 13.80 m (SD = 6.27 m) from the infringement. A position of 10–15 m from the incident was associated with higher chances for making a whistle error (i.e., 2.58 times higher than the overall viewing-distance effect). A close position of 0–5 m from the critical incident was associated with higher chances of making a non-whistle error (i.e., 5.51 times more likely to commit a non-whistle error). The authors concluded that there is no “ideal distance” for making accurate decisions, and that elite referees are able to effectively position themselves in respect to an...
anticipated match event. In cases which they fail to do so, the risk of making errors substantially increases. Castillo et al.’s (2019b) analysis of FIFA top-level referees officiating in an international tournament further indicated lack of differences in DM accuracy levels among 10 field zones.

The referee’s ability to make effective positioning decisions might also be influenced by his/her fitness level and momentary physiological fatigue (Castillo et al., 2016a; Krustrup & Bangsbo, 2001; Mallo et al., 2012; Schenk et al., 2018). For example, Mallo et al. (2012) found that in the central area of the field, accuracy rates were lower in the last 15 min of the matches than in any other 15-minute period. This might have resulted from the referees’ physical or mental fatigue, or due to teams’ tactical modifications. Research findings indicated that referees’ fitness levels might also change throughout the competitive season, being lower at the beginning of the season (Castillo et al., 2017) and higher at the end (Castillo et al., 2019a), with potential influence on decision-making quality. It is advised that referees regulate their hydration status to prevent high body core temperature and thirst which might debilitate DM accuracy (Houssein et al., 2016). Furthermore, as referees age (i.e., specifically above 35 years) their sprint capacity, change of direction ability and endurance capacity might decrease, potentially affecting their field positioning and decision-making (Castillo et al., 2016b). Weston et al. (2011) demonstrated how referees might adapt their training regimen to account for such decrements.

In addition to the effects of referees’ positioning and fitness, several investigations focused on how or why soccer referees make decisions. For example, Lane et al. (2006) interviewed five professionally experienced referees. Content analysis yielded 13 subthemes, which were categorised into three main themes: individual factors (e.g., concentration/avoidance, control), experience factors (e.g., personality, experience), and situational factors (e.g., crowd factors, player’s reaction), which contributed to a higher theme of ideal decision-making (e.g., accuracy/errors, regulations). In a similar fashion, Russell et al. (2019) interviewed nine national-level English referees concerning the reasons referees make decisions, and the perceptions underlying their DM judgments. The content analysis indicated that DM actions were used strategically to achieve two goals: (1) to maintain control, and (2) to preserve the integrity of the game. The referees’ DM satisfied these goals through four “pillars” of the game – safety, fairness, accuracy, and entertainment – which varied in importance depending on the state and context of the game.

**Bias vs. game-management in DM**

One of the conceptual contentions examined in refereeing DM is whether they are biased in their decisions (for a review see Dohmen & Sauermann, 2016; Pina et al., 2018) or rather apply an intentional game-management approach (e.g., Mascarenhas et al., 2002; Mascarenhas et al., 2006; Unkelbach & Memmert, 2008). In support of the first notion, Plessner and Betsch (2001) examined sequential effects in refereeing DM using a laboratory setting. They showed that referees do not tend to award two successive penalty decisions in favour of the same team. In contrast, a positive link was reported between successive penalty decisions involving one team and then the opposing team. However, such an effect was not evident in relation to free-kick situations. The authors interpreted these findings as evidence that the referees were biased by their own previous match decisions.
Commenting on this study, however, Mascarenhas et al. (2002) attributed these findings to referees’ conscious *game management*: they felt the game intensified and, to restrain this heightened aggressiveness, they decided to award the penalty. These authors suggested that referees tend to make decisions that are game-appropriate, allowing the game to flow, and using the whistle only when the match events demand it.

The notion of game management was further demonstrated by Unkelbach and Memmert’s (2008) analysis of the yellow-card administration patterns of German referees in a two-experiment laboratory setting. They initially found that referees who judged scenes in the context of a game awarded fewer yellow cards than referees who viewed the same scenes in a random order. Secondly, they found that soccer referees used both *deliberate game management* (i.e., purposely postponing cards early in the match) and *calibration of the fouls’ scale* (i.e., setting criteria for issuing a yellow card). Continuing this line of research, Schwarz (2011) analysed a large sample of German Bundesliga matches and revealed that the number of two-penalty matches is larger than expected by chance, with considerably more matches in which each team is awarded one penalty than would be expected based on independent penalty kick decisions. Additional analyses suggested that awarding the first penalty to one team raised the referee’s penalty evidence criterion for the same team, and lowered the corresponding criterion for the other team. The author interpreted these findings as a “balancing out” bias, supporting Plessner and Betsch’s (2001) notion of refereeing bias. However, these findings could also be understood in line with Unkelbach and Memmert’s (2008) notion of a calibration process, and do not necessarily indicate a deliberate referee bias. Yet, both explanations require additional research.

**Contextual factors influencing DM**

Various studies focused on examining specific match factors that might influence referees’ DM. These included, among others, home advantage, as manifested by crowd noise (Goumas, 2014; Lovell et al., 2014; Nevill et al., 2017; Picazo-Tadeo et al., 2017), a team’s reputation for aggressiveness (Jones et al., 2002), the referee’s positioning in the field of play (Mallo et al., 2012), players’ vocalisation following foul incidents (Lex et al., 2015), the players’ uniform colour in judging tackle incidents from behind (Krenn, 2014), foul offenders’ height (van Quaquebeke & Giessner, 2010), and extreme weather conditions (Gaoua et al., 2017). In addition, the opponent level (i.e., ranking home team – ranking visiting team) and the score might influence the referee’s decision concerning the addition of extra time (Lago-Peñas & Gómez-López, 2016). Furthermore, Coleclough (2013) examined the perceptions of coaches and referees about decisions made by the UEFA Euro 2012 match referees in 106 tackle incidents. The participants commented that the match referee’s decisions were influenced by contextual factors, such as pitch location, a player’s previous fouls, and time of play.

Another potential contextual factor is the match teams’ ranking. Castillo et al. (2018) examined the external and internal match loads of referees attending teams of different ranking (i.e., top 10, bottom 10, mixed) in championship matches. The results showed that referees who officiated matches involving top 10 teams, covered more distance at a low walking speed (< 3.6 km·h) and performed a higher percentage of high-intensity accelerations and decelerations than referees who officiated lower ranked teams’
matches. These findings suggest that the ranking of the involved teams might influence referees’ running patterns, thereby affect their positioning and DM.

In contrast to the above findings, Castillo et al. (2019b) analysis of an international tournament showed that other contextual factors (i.e., period of the match, tournament phase, match score) did not significantly influence the referees’ decision-making. However, this finding could be attributed to the relatively high decision accuracy levels (i.e., above 90%) which resulted from the high-level international referees involved in this analysis. Also, it is possible that contextual factors pose different influence in international tournaments than in league matches.

**Personal characteristics influencing DM**

Within this domain, several studies aimed at referees’ personal characteristics, including visual skills, perceptual skills, executive attention, stress, expertise level, self-control, referee’s height, and self-efficacy. For example, there is evidence that much like other sport performers, elite referees maintain superior general (Ghasemi et al., 2009) and domain-specific (Spitz et al., 2016) visual skills as well as high visual experience (Pizzera & Raab, 2012). Ste-Marie (2003) claimed that expert sport officials use a visual pivot strategy: focusing on a central location and picking up other information using peripheral vision. Much like expert athletes, elite referees’ selective attention (i.e., attending to relevant information and ignoring irrelevant information) enables a more efficient DM process. Expert referees were found to use fewer gaze fixations for longer durations, resulting in more accurate DM than sub-elite referees (Spitz et al., 2016). Also, for potential foul situations, elite referees fixated longer on the most informative area of the attacking player’s body and less time on the body parts that were not involved in the infringement.

Spitz et al. (2018b) reported that DM in refereeing was attributed to domain-specific (i.e., video-based foul DM), but not domain-generic (e.g., processing speed) perceptual skills. Specifically, elite referees were better than sub-elite referees at making foul decisions and in anticipating upcoming events when the video clip was occluded just before the moment of contact. Likewise, Pietraszewski et al. (2014) found that executive attention played a significant role in referees’ professional level. Specifically, the executive attention (i.e., using the Toulouse-Pieron test) of Polish referees from three professional levels (international FIFA, Extraclass – Premier League, first division) were examined. The Extraclass referees maintained better executive attention compared with the lower level referees. However, there were no significant differences between the international referees and the lower level referees. Age and experience also correlated positively with executive attention. It was suggested that older, more experienced referees, compensate for reductions in fitness and speed levels by being more attentive to important match cues.

Moreover, stress has also been considered as a factor that has received research attention in soccer refereeing. Much of the existing research focused on the high school, college, and amateur levels of soccer refereeing, indicating low to moderate stress levels (e.g., Folkesson et al., 2002; Voight, 2009; Wolfson & Neave, 2007). Performance demands are higher at the elite level, and referees are under the scrutiny of professional factors (e.g., the match observer, the Referee Union Professional Committee), the sporting community (e.g., coaches, players, fans), and the media (Dawson, 2012; Page & Page, 2010; Samuel, 2015; Schnyder & Hossner, 2016; Slack et al., 2013).
The role of experience/expertise in soccer refereeing performance was also evident in additional studies (e.g., Catteeuw et al., 2009; Dawson, 2012; Gilis et al., 2008; MacMahon et al., 2007; Pizzera & Raab, 2012). In addition, Samuel et al. (2018) found that soccer referees possess higher trait self-control than elite players and the general population. Moreover, they reported that referees’ state self-control depletion during a match was a relatively a common phenomenon (i.e., apparent in 49% of the matches). Matches in which referees experienced noticeable ego depletion indicated lower self-rated match performance mean ($p < .05$).

Recently, the relationships between referees’ height and punitive behaviours (i.e., yellow and red cards, penalties) given during an entire season were analysed for 61 male soccer referees from four professional leagues in England (McCarrick et al., 2020). Their underlying notion was that physical height may underpin a referee’s ability to control the match and deliver accurate decisions. The findings indicated no effect of referee height on fouls awarded. Shorter referees, however, tended to issue more yellow cards ($r = -.28, p < .05$). A similar effect was also evident for red card ($r = -.34, p < .05$), yet moderated by league; in the lower leagues, more red cards were awarded by relatively shorter referees, whereas in the higher leagues more red cards were awarded by relatively taller referees. A similar league level effect was also evident for penalty decisions. These findings suggest that context plays an important role in red card and penalty decision-making.

Finally, Guillén and Feltz (2011) developed a conceptual model of efficacy for sports referees (termed refficacy), using a focus group of soccer referees. The model outlines the sources, dimensions, and outcomes of refficacy. Having high DM skills – which translates as making critical match decisions, demonstrating accurate judgments, and being firm in one’s decisions – was an essential dimension of refficacy. Other essential dimensions of referees’ self-efficacy were focusing attention, staying cool under pressure, and coping with mistakes and adverse situations.

The refereeing team and DM

Referees might also rely and incorporate with their assistants to make decisions (Helsen & Bultynck, 2004; Mallo et al., 2012). Mallo et al. found that decisions in the lateral areas of the field of play (i.e., involving the referee and assistant referee) were characterised by lower accuracy rates than decisions made in the central areas (13% vs. 17%, respectively). In this context, Samuel (2015) suggested that the “referee team” must communicate effectively during the DM process to establish shared mental operations. Shared mental models (SMMs) are an important factor in team DM, facilitating implicit coordination through a shared understanding of (a) the problem definition (e.g., what constitutes a reckless foul), (b) the strategies for solving the problem (e.g., how to manage players and use proactive approach) (c) the interpretation of cues and information (e.g., assessing where and when a reckless foul might be committed), and (d) the roles and responsibilities of the team members (e.g., what is expected from the referee team when a reckless foul is committed and identified – delivering the message effectively; Mascarenhas et al., 2005).

Within this context, Schnyder and Hossner (2016) identified within-team communication as one of the main factors underlying refereeing performance. Within-team communication, including with the VAR (Sánchez Cid & García García, 2020), especially
during critical match incidents, seems highly significant for effective DM, but there is hardly any research concerning this aspect of refereeing.

To conclude, various factors have the potential to affect the soccer refereeing decision-making process. These can be associated with the referee’s characteristics, the match’s characteristics, and performance related aspects (including the referee team). These factors are specified in Table 1, with the related references.

**Conceptual perspectives on refereeing DM**

Plessner and Haar (2006) applied a social cognition perspective (Bless et al., 2004) in modelling the DM sequence in soccer refereeing. According to this perspective, social cognition researchers study the biases involved in social judgments. Using Bless et al.’s framework, Plessner and Haar suggested that soccer referees’ DM follows a sequence of (1) stimulus events (2) perception, (3) categorisation, (4) memory processes, (5) information integration, and (6) behavioural response. It is suggested that, as part of this process, the encoded episode is stored in long-term memory and may influence future decisions, just as retrieved episodic memories influence current processing. Therefore, DM errors can result from smaller errors or incorrect information from different steps of information processing.

Plessner and colleagues (Plessner et al., 2009; Schweizer et al., 2011) further applied the multiple-cue learning approach in describing the categorisation stage within referees’ DM process. According to this perspective, in the case of foul situations, soccer referees make decisions based on a perceptual categorisation task in which the referee must categorise a diverse set of features into two discrete classes (i.e., foul or no foul). Specifically, referees must judge a distal criterion (e.g., foul or no foul) by taking visible cues into account. The accuracy of the decision is determined by how closely a referee’s cue utilisation mirrors ecological validity (i.e., taking only the relevant cues into account). It is also suggested that, due to time pressure, referees’ responses are more likely to result from automatic (intuitive) than from deliberative processes. Therefore, multiple features of a DM situation (e.g., cues) may be processed simultaneously and quickly, using an automatic (i.e., without conscious awareness) and effortless processing. However, Schweizer et al. (2011) did acknowledge that other decisional tasks in soccer refereeing (e.g., decisions pertaining to managing players’ dissent) may require deliberative, and not intuitive, processing.

Mascarenhas et al. (2005) adopted a somewhat different perspective, examining DM among rugby union referees using the naturalistic decision-making approach (NDM; Orasanu & Connolly, 1993). This perspective suggests that expert referees’ decisions are governed by their superior organisation of knowledge into knowledge structures. Specifically, they referred to mental models, including shared mental models within a referee team, and independent DM. This perspective assumes that, in addition to evaluating referees’ decisions, their underpinning reasons for making these decisions must also be examined, as reflected by their shared mental model. These conceptual perspectives present a solid basis for exploring soccer referees’ DM processes. However, these frameworks were not specifically designed for elaborating soccer refereeing DM, and therefore might not include the entire range of factors (e.g., referee’s physiological state, referee’s stress response, contextual information, referee team communication) that can potentially influence the DM process. Moreover, soccer referees do not engage in an isolated DM
Table 1. Factors associated with soccer refereeing decision-making.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Definition</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field positioning</td>
<td>Covering the field of play according to the ‘diagonal system of control’ (IFAB, 2019) in order to obtain the most effective position for identifying and reacting to match events.</td>
<td>Castillo et al. (2019b), Hossner et al. (2019), MacMahon et al. (2014), Mallo et al. (2012), Schnyder and Hossser (2016), Slack et al. (2013), Voight (2009)</td>
</tr>
<tr>
<td>Visual attention and search strategy</td>
<td>Gazing towards relevant information (i.e., areas of the field with potential actions) while ignoring irrelevant information. Using fewer gaze locations for a longer duration. Maintaining an external wide – narrow attentional focus.</td>
<td>Ghasemi et al. (2009), Hüttermann et al. (2018), Spitz et al. (2016), Ste-Marie (2003)</td>
</tr>
<tr>
<td>Fitness level and fatigue</td>
<td>Fitness: Referee’s ability to meet the physical demands of the match in terms of running, sprinting, and walking in order to assume appropriate field positioning for effective decision-making and player management. Measured in-situ as well as by field tests. Fatigue: Referee’s objective (e.g., lactate levels, reduction in speed) and subjective (e.g. perceived exertion) experience of physiological fatigue during a match.</td>
<td>Castillo et al. (2016a, 2016b, 2017, 2019a), Casajus and Castagna (2007), Helsen and Bultynck (2004), Krustrup and Bangsbo (2001), Mallo et al. (2012), Mascarenhas et al. (2009), Pina et al. (2019), Schenk et al. (2018, 2013), Weston (2015), Weston et al. (2011, 2012)</td>
</tr>
<tr>
<td>Experience and expertise</td>
<td>Time spent in developing refereeing expertise through deliberate practice or non-directed activities. Years of experience in refereeing Level of officiating (e.g., amateur, semi-professional or professional soccer, elite vs. sub-elite).</td>
<td>Catteeuw et al. (2009), Dawson (2012), Gilis et al. (2008), MacMahon et al. (2007), Pietraszewski et al. (2014), Pizzera and Raab (2012), Samuel (2017), Spitz et al. (2018b)</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>Referee’s self-beliefs about one’s ability to successfully officiate the match in terms of physical fitness, knowledge of the game, decision-making skills, psychological skills and readiness, game management and control, and team-work.</td>
<td>Guillén and Feltz (2011), Johansen et al. (2018), Samuel (2015), Schnyder and Hoosser (2016), Slack et al. (2013)</td>
</tr>
<tr>
<td>Stress and coping</td>
<td>Referee’s feelings of apprehension, nervousness, and arousal because of internal stressors such as apprehension of match demands, fear of failure, fear of disappointing oneself and others, fear of choking under pressure, and external stressors such as aggression and dissent from players, coaches, fans, and the threat of being intervened by the VAR. Applying problem-focused (e.g., increasing effort, setting performance goals, tactical planning) and emotion-focused (e.g., relaxation, confidence building, communication with assistants) coping strategies.</td>
<td>Dawson (2012), Folkesson et al. (2002), Page and Page (2010), Samuel (2015), Schnyder and Hoosser (2016), Slack et al. (2013), Voight (2009), Wolfson and Neave (2007)</td>
</tr>
<tr>
<td>Self-control</td>
<td>Referee’s ability to volitionally apply attention regulation, make consecutive and difficult decisions, cope with internal and social pressure, control one’s impulses, and suppress a tendency to reduce running pace due to fatigue.</td>
<td>Lane et al. (2006), Samuel (2015), Samuel et al. (2018)</td>
</tr>
<tr>
<td>Passion for refereeing</td>
<td>Referee’s strong inclination or desire toward refereeing. Harmonious passion involves a flexible engagement in the refereeing activity, which is conducive to a willingness to</td>
<td>Johansen (2015), Philippe et al. (2009)</td>
</tr>
</tbody>
</table>

(Continued)
task, but rather make multiple successive (sequential) decisions within a match (e.g., Helsen & Bultynck, 2004; Mallo et al., 2012). Therefore, as part of viewing referees as performers on their own merit, specific models must be developed to elaborate on their performance.

A model of sequential DM for soccer referees

To account for the complexity of the DM process in soccer refereeing, we developed a descriptive model that represents the sequential DM in referees. The model is presented in Table 2 (the model) and Figure 1 (illustrations of decisions – upper section; putative input-output relations among the concepts and influencing factors – lower section). The new model consists of Tenenbaum’s (2003) sequential DM model and additional refereeing-specific aspects that were reviewed. According to Tenenbaum, response selection in dynamic sport contexts refers to adaptive behaviour based upon the performer’s capacity to solve problems. The information-processing perspective suggests that motor behaviours in complex sport situations consist of encoding the relevant environmental cues through the utilisation of attention strategies, information processing involving an

<table>
<thead>
<tr>
<th>Factor</th>
<th>Definition</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game management</td>
<td>Making decisions that are game appropriate and reflect match-specific accuracy and fairness. Allow match flow. Managing players through effective communication skills.</td>
<td>Coleclough (2013), Mascarenhas et al. (2002), Mascarenhas et al. (2006), Mellick et al. (2005), Pina et al. (2019), Russell et al. (2019), Samuel (2015), Schnyder and Hossner (2016), Slack et al. (2013), Unkelbach and Memmert (2008)</td>
</tr>
<tr>
<td>Calibration processes</td>
<td>Referee’s observed judgments within a match that are assessed according to an internal classification scale or category system (e.g., what constitutes a reckless foul that results in a yellow card, within a specific match context).</td>
<td>Unkelbach and Memmert (2008)</td>
</tr>
<tr>
<td>Referee team shared mental model and communication</td>
<td>A shared understanding of the match tactical approach (‘speaking the same language’). A unified calibration of various match decisions. Effective communication within the referee-team during the decision-making process. A shared approach towards player management.</td>
<td>Helsen and Bultynck (2004), Mallo et al. (2012), Pina et al. (2019), Samuel (2015), Sánchez-Cid and García-García (2020), Schnyder and Hossner (2016), Slack et al. (2013)</td>
</tr>
</tbody>
</table>
ongoing interaction between working memory and long-term working memory (LTWM; Ericsson & Kintsch, 1995), action-related decision-making, and action execution while maintaining optional modifications. The DM sequence includes several consecutive decisions beginning with visual and attention selection and ending with action evaluation.

The refereeing DM model presents the series of decisions referees might go through during each match event/infringement, and the underlying operational mechanism and potential influencing factors (for operational definitions of the influencing factors see Table 1). While the potential influencing factors were specified for each decision in the sequence, it should be noted that certain factors (e.g., self-efficacy, self-control, fitness) might influence the entire sequence. As shown in Figure 1(a), at each point of the match referees must decide upon the best spatial location to observe the ongoing actions (i.e., the first decision in the process). Referees typically run a diagonal route, covering the field of play from one side of the goal area to the other (known as a diagonal system [IFAB, 2019]). However, skilled referees also anticipate the actions’ locations on the field of play, relying on their professional experience, prior knowledge of the teams’ style, and present match style (Spitz et al., 2018b). In fact, the latest practical guidelines for match officials concerning positioning and movement (IFAB, 2019) recommend that “the best position is one from which the referee can make the correct decision. All recommendations about positioning must be adjusted using specific information about the teams, the players and events in the match” (p. 182). For example, if a referee anticipates or identifies that a team consistently attacks from the right side of the goal area, he/she will change his/her running pattern and spend a longer time in this area of the field. Thus, the referee’s first decision is where to run on the field of play (see Figure 1(a)).
Inappropriate positioning might increase the likelihood of decision error, as the referee might not be able to detect the essential stimuli (e.g., the position of the ball). As

Figure 1. An illustration (upper section) and putative input-output relations among the concepts (lower section) of a referee performing the sequential decision-making process in a tackle situation.
shown above, the referee’s ability to make effective positioning decisions might also be influenced by his/her fitness level and momentary physiological fatigue (Castillo et al., 2016a; Krustrup & Bangsbo, 2001; Mallo et al., 2012; Schenk et al., 2018).

Once the referee is positioned well on the field of play, he/she must pay visual attention to (look at) the appropriate location (i.e., the second decision in the process; see Figure 1(b)). This involves detecting the most relevant visual cues in the field of play, based on experience and prior knowledge concerning the teams and forms of play (Hüttermann et al., 2018; Ste-Marie, 2003). Expert referees use more effective gaze strategies than non-experts (e.g., they gaze toward fewer locations for a longer duration) as they select relevant information and ignore the irrelevant one (Spitz et al., 2016; Ste-Marie, 2003). Experts also use more efficient visual strategies than non-experts, such as peripheral vision, eye saccadic movements and speed of recognition (Ghasemi et al., 2009).

The skilled referee then anticipates potential events and infringements, such as a foul, ball handling, or out of play (i.e., the third decision in the process; see Figure 1(c)). Roca et al. (2011) examined the processes underpinning anticipation and decision-making in a dynamic time-constrained task among skilled and less skilled soccer players. They suggested that expert sport performers have superior cognitive-perceptual skills, such as recognising information from the postural orientation, faster and more accurate recognition of representative patterns, and a superior ability to make accurate predictions. This leads to making more accurate decisions, potentially due to more advanced memory representations enabling the skilled players to easily retrieve task-specific information and make better judgments. In this context, Ericsson and Kintsch (1995) argued that expert performers develop superior LTWM as a result of deliberate practice. This allows experts to efficiently retrieve critical task-relevant information from long-term memory.

Therefore, the anticipation process involves uploading potential events and infringements from the schemas stored in the referee’s LTM (Ericsson & Kintsch, 1995; Tenenbaum, 2003). These options depend on the referee’s field location (e.g., substituting players tend to make aggressive fouls closer to the coach’s area), as well as on contextual factors that tend to change the tempo of the match and the teams’ strategies. These may include the score (e.g., a tied score might mean that more fouls are committed in order to win the match), the match’s playing time (which tends to influence referees’ usage of cards; Unkelbach & Memmert, 2008), the atmosphere in the stands (e.g., Picazo-Tadeo et al., 2017), and specific players (e.g., some players may strive to deceive the referee by simulating in the penalty area; Sabag et al., 2018).

The ability of referees to retrieve information using LTWM could also be affected by their current stress levels and fatigue. Specifically, when the situation is appraised as threatening or risky, the referee may develop extended stress accompanied with a feeling of high exertion. Also, extreme weather conditions may pose distractions in LTWM functionality, which can result in DM efficiency reduction (Gaoua et al., 2017; Samuel, 2015; Tenenbaum, 2003; Weston et al., 2012).

When the event/infringement occurs, the referee identifies and processes this information (see Figure 1(d)). The effectiveness of the information processing might be affected by the referees’ positioning, the obstruction of referee’s view by the players, or the referee’s attentional focus (e.g., narrow attention or internal attention; Hüttermann et al., 2018; Samuel, 2015). Additionally, the identification process is influenced by the
number of generated options. In efficient DM, the **referee anticipates the action and then activates possible actions held for execution**. This process allows the referee to quickly decide upon the appropriate action (Ste-Marie, 2003). When the referee is unable to anticipate the upcoming event and retrieve an appropriate option from memory, the decision is delayed and in cases is less reliable (Helsen & Bultynck, 2004; Mallo et al., 2012).

To make an accurate decision (i.e., the fourth decision in the process), the referee is required to match between the identified action and the criteria specified for each infringement in the Laws of the Game (IFAB, 2019; see Figure 1(e)). However, the decision should also be in line with the match contextual factors (e.g., the match score, the playing time, previous decisions, the varying levels of aggressiveness) to be perceived as fair and correct (Russell et al., 2019). Thus, the referee not only uses his/her knowledge, but also applies game management (Mascarenhas et al., 2002). For example, when a referee identifies that a match flows quietly and fairly, he/she will not tend to call many fouls. On the other hand, if a referee expects (i.e., as part of his/her tactical preparation for the match: see Samuel, 2015) or concurrently identifies (i.e., during the match) that the match is highly aggressive and unfair, then he/she will enforce each infringement more carefully. A referee might also develop a “story” with a specific player during a match, which can be affected by the player’s reputation (e.g., a player who tends to dive in the penalty area) and his/her previous match behaviour (e.g., conducting several consecutive fouls, dissenting towards the referee). Referees at the elite level are expected to present the teams and the crowd with a clear story; their decisions must be in line with the match social context, so they are understood and accepted (Coleclough, 2013; Russell et al., 2019; Samuel, 2015).

When the referee indicates his/her decision by whistling or hand signals, he or she then receives internal and external feedback concerning the decision. External feedback may come from the assistants via the communication system (Samuel, 2015; Schnyder & Hossner, 2016), from the players (Lex et al., 2015), from the fans (Picazo-Tadeo et al., 2017), or from the VAR. Internal feedback takes place when the referee reruns the event in his/her mind and considers whether the original decision was accurate. Typically, the external feedback matches the internal feedback for the referee to be certain of the decision. Match events in which the referee receives contradicting information from his/her assistants while still assessing the situation are particularly challenging, resulting in competing DM processes, which might lead to an error (e.g., Mallo et al., 2012). Furthermore, when the internal feedback is unclear (e.g., the referee is uncertain that he/she identified and evaluated the event correctly), the referee will typically ask for external information. Also, in critical match decisions that have a high influence on the match outcomes (e.g., penalties, red cards), referees might also experience aggression from players and mobbing (Folkesson et al., 2002), which make these match instances particularly stressful. Therefore, the method with which referees manage decision-related feedback is influenced by their efficacy (Guillén & Feltz, 2011) as well as their self-control strength (Samuel et al., 2018).

Referees who make an unpopular decision must remain calm and under control, continuing with the DM process. This psychological demand is augmented in cases of VAR intervention, particularly when the referee is challenged to change his/her original decision. In this context, Philippe et al. (2009) found that following an important decision...
error, referees with obsessive passion for refereeing engaged in maladaptive functioning, including subsequent poor decision-making.

When the referee receives feedback about his/her decision, a fifth decision emerges: whether to keep the original decision or alternate it (see Figure 1(f)). The insertion of VAR makes this decision particularly viable in cases of major decisions, such as penalties and red cards (IFAB, 2017). Then, the referee executes the decision by using the appropriate signalling and applying match protocols (e.g., free kick, yellow card). Decisions that involve long protocols, such as management of free kicks, also put a mental load on the referee and might affect subsequent DM. Finally, the model suggests that since referees are required to make numerous decisions in each match (Helsen & Bultynck, 2004), a circular process is followed, so that future decisions are affected by present ones.

### Preliminary evaluation of the model

The evaluation of the model requires considerable research effort. Thus, to conduct a preliminary evaluation of the face validity of the model, a survey of experts in the domain was conducted. Specifically, the model, including the illustrations (i.e., Table 2 and Figure 1-upper section), was circulated using Google Forms among Israeli referees in the first and second divisions and assistant referees in the first division. Overall, 13 referees and seven assistant referees responded to the survey. These were all semi-professional referees or assistants, with an average of 15.85 years (SD = 6.20) of refereeing experience. Seven of the participants had an international badge. The referees were instructed to read the description of the model, go over the table and figure, and respond to three questions about the model using a seven-point Likert type scale ranging from 1 (not at all) through 4 (moderately) to 7 (very much). The questions were: (1) To what degree does the model represent the DM process of soccer referees? (2) To what degree does the model include the critical components of the DM process of soccer referees? and (3) To what degree did you find the model clear and concise? Furthermore, the participants were given an option to openly indicate any other important factors that were not specified in the model. The referees felt that the model accurately represented their DM process ($M = 5.75, SD = 0.85$), included the critical components ($M = 5.70, SD = 0.98$), and was clear and concise ($M = 5.65, SD = 1.04$). They did not identify unspecified factors.

### Conclusions and implications for training

We further contribute to the area of soccer refereeing DM by conceptually integrating the existing knowledge in the referring domain. As such, we elaborate on the work of previous authors (Mascarenhas et al., 2005; Plessner et al., 2009; Plessner & Haar, 2006; Schweizer et al., 2011) by zooming out and looking at the entire refereeing task rather than at specific decisions or parts of the DM process.

As various factors might influence the sequential DM process of soccer referees, it seems reasonable that no single study design can account for all factors. Our preliminary evaluation of the model’s face validity among high-level referees was positive; yet should be addressed with cautious due to the limited sample size and the nature of evaluation (i.e., an online inventory rather than in-depth interviews). Still, considerable research effort is required to examine how the various factors affect DM separately and jointly.
Also, what the critical influencing factors are. Presently, the above framework does not differentiate between the various influencing factors as to their significance within the DM sequence. We, therefore, propose that research on soccer referees’ DM should develop in three main ways. The first is by making a methodological shift from focusing on refereeing bias to focusing on how to facilitate successful refereeing performance, as suggested by Schweizer et al. (2011). This is reflected in a DM process which yields accurate and appropriate (i.e., fair, entertaining) decisions. Morgulev et al. (2018) found that, in basketball offensive foul decisions, common biases (e.g., home advantage, star players) were not very robust as well as context sensitive.

The second area for development is adopting research designs which better reflect the DM task, involving physical and mental load as well as contextual information. It is advised that refereeing coaches should first establish clear criteria for which contextual information should be applied within the various match situations (e.g., match score, time of play) and which must be ignored (e.g., reactions of the crowd, players’ shirt colour). The third area of research development is to find novel ways to examine the mutual effects of various professional, physical, psychological, and social factors on referees’ DM performance. For example, to adequately perform the sequential DM process, referees must be self-efficacious and exhibit high self-control strength. There is no current research on how referees’ efficacy and self-control strength fluctuate during a match, and how this might affect their DM process. So far, researchers have mostly used strict qualitative or quantitative designs. However, using mixed-methods designs, such as measuring performance indices in-situ and using retrospective self-analysis (e.g., Cottyn et al., 2006) might lead to an in-depth understanding of the underlying mechanisms.

The development of effective shared mental models (Mascarenhas et al., 2005) and headset communication (Schnyder & Hossner, 2016) within the referee team is critical for successful DM. This demand has been augmented by the introduction of VAR (IFAB, 2017) and potential intervention in critical match incidents. This system involves transformations between a typical DM process and on-field review (OFR; IFAB, 2017) DM, which is much closer to Plessner et al.’s (2011) idea of categorisation. The referee is required to abruptly modify his/her mental state from including contextual factors to more rigid law-criteria-based decisions. Another interesting aspect is related to the speed (i.e., regular vs. slow motion) in which a referee should watch the match event reply during OFR to obtain the most accurate decision (Spitz et al., 2018a). This complex process deserves conceptual and empirical consideration. Moreover, the inclusion of the VAR raises an important question within the socio-cultural realm of the match; do the players and fans seek for the most accurate decisions (i.e., as might be produced by the VAR intervention), or do they also wish the referee’s decisions to be fair and relevant within the context of the match? This issue raises psychological dilemmas for referees – use the VAR system to “protect” their decisions, or use their knowledge and experience to produce quality refereeing? Perhaps do both, but how exactly? These issues necessitate future examination.

Several researchers have suggested that effective training of soccer referees’ DM remains a challenge (e.g., Catteeuw et al., 2009; MacMahon et al., 2007). To account for this gap, Schweizer et al. (2011) designed a video-based DM training programme of foul situations that was based on Plessner and Haar’s model (2006) and the idea of intuitive processing. This training method indicated some improvements compared to a control
condition, yet Schweizer et al. (2011) concluded that there was no evidence that learning effects obtained during the training transferred to the real world. In addition, Put et al. (2013) examined the extent to which off-field, web-based, offside DM training transferred to real-life offside situations. The study’s findings showed that the training group increased in response accuracy and decreased in the number of flag errors from pre- to post-test, in both an on- and off-field transfer test. These authors suggested that the online learning environment could mimic the perceptual difficulties of real-match situations. Nevertheless, it seems that the offside task was performed in a stationary mode both off- and on-field, allowing the assistant referees to save any early run prior to making the offside decisions and preventing fatigue.

While web-based training tools might have facilitated perceptual skills, a real match typically presents additional demands (e.g., crowd, fatigue, stress). Hence, further research is required to establish training modes that effectively simulate the refereeing DM task in a manner that transfers to real-world performance. Developing effective 2-D and 3-D simulators (e.g., Gulec et al., 2019; Put et al., 2014) provides new opportunities, but must involve the entire range of potential factors, including a demand for sequential DM, to create a realistic experience.

To account for the above gap in training modalities, Samuel et al. (2019) developed a decision-making simulator for soccer referees. Referees ran on a treadmill for 60 min at varying paces while watching on a regular tablet two video sections depicting events from real (unfamiliar) matches, in context and in mixed order, calling their match decisions out loud. The referees felt the simulator was moderately representative of a real match and offered suggestions for improvement. Decision accuracy levels for careless fouls (62.86%) was on par with previous studies, supporting the simulator’s validity. The authors concluded that the findings supported the use of the new simulator for training referees in sequential decision-making.

It should be acknowledged that the suggested DM framework has limitations. First, a systematic review (Ferrari, 2015) of the soccer refereeing literature was not performed, thus it is possible that additional relevant publications were not reported. Furthermore, we must accept that the DM process is highly dynamic and might not be absolutely captured using a sequential process. For example, certain aspects (e.g., anticipation, refficacy, self-control) might play a role throughout the entire sequence and not only in specific decisions, as indicated by the model. It is also unclear, at this point, how the various factors interact to influence DM as each point along the sequence.

In addition, we did not directly compare the framework to other DM perspectives. For example, the ecological dynamics perspective (see Araújo et al., 2019) advocates that behaviours emerge from the continuous interactions in the performer-environment system. In open, dynamic systems there is no “best decision,” as the most functional decision at any moment may compromise future decisions. This perspective might suggest that elite referees are better attuned to the various information exist on the field, not necessarily because they have better schemas located in their LTWM. We believe that soccer refereeing is indeed performed in a dynamic environment, yet the types of decisions expected to be made by the referees are limited in number and in criteria, thereby “the best” decision tends to be the most accurate one. However, we agree that the best decision is also the most appropriate one, and this has to do with being highly attuned to the environment, in anticipating events and in understanding decisions
within match contexts. Thus, the ecological dynamics perspective certainly complements our framework.

Finally, the initial evaluation of the model was limited in sample size and in the method used. It is unclear whether referees have the necessary meta-cognition to evaluate the model, although they serve as experts in this field. We also did not account for potential effects of demand characteristics, social desirability, and confirmation bias. Therefore, the model should be further validated using more robust methodology.

In conclusion, several conceptual and practical challenges still remain for advancing refereeing DM performance. This article aimed to establish a solid basis for future research and practice, incorporating into a single framework the vast range of factors which have the potential to influence referees’ DM, within real match contexts. Still, we acknowledge that no single framework can fully account for such a complex process and therefore call for additional theoretical and empirical endeavours to promote a clearer understanding of this process, especially in light of a “game-change” factor such as VAR. Developing sophisticated simulators can be of high importance for researchers who wish to examine the underlying factors of the DM process presented in this framework, for example, the relationships between physical fitness, fatigue, self-control strength, contextual information, and DM performance.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Roy David Samuel http://orcid.org/0000-0001-9914-3754

Gershon Tenenbaum http://orcid.org/0000-0001-6357-873X

Yair Galily http://orcid.org/0000-0001-5308-5675

References


