The role of domain-generic and domain-specific perceptual-cognitive skills in association football referees

Jochim Spitz, Koen Put, Johan Wagemans, A. Mark Williams, Werner F. Helsen

PII: S1469-0292(17)30068-7
DOI: 10.1016/j.psychsport.2017.09.010
Reference: PSYSPO 1270

To appear in: Psychology of Sport & Exercise

Received Date: 26 January 2017
Revised Date: 20 September 2017
Accepted Date: 20 September 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
The role of domain-generic and domain-specific perceptual-cognitive skills in association football referees

Jochim Spitz\textsuperscript{a}, Koen Put\textsuperscript{a}, Johan Wagemans\textsuperscript{b}, A. Mark Williams\textsuperscript{c}, Werner F. Helsen\textsuperscript{a}

\textsuperscript{a} Department of Kinesiology, Laboratory of Perception and Performance, Movement Control and Neuroplasticity Research Group, University of Leuven (KU Leuven), Leuven, Belgium.
- jochim.spitz@kuleuven.be (+32 16 37 79 60) (Corresponding author)
- koen.put@kuleuven.be (+32 16 37 95 99)
- werner.helsen@kuleuven.be (+32 16 32 90 68)

\textsuperscript{b} Department of Brain & Cognition, Laboratory of Experimental Psychology, University of Leuven (KU Leuven), Leuven, Belgium.
- johan.wagemans@kuleuven.be (+32 16 32 59 69)

\textsuperscript{c} Department of Health, Kinesiology, and Recreation, College of Health, The University of Utah, Salt Lake City, USA
- mark.williams@health.utah.edu (+1 435 901 2447)
The role of domain-generic and domain-specific perceptual-cognitive skills in association football referees
Abstract

Objectives: We examined the importance of perceptual-cognitive skills in association football referees.

Design and method: Using a novel approach, elite (n = 22) and sub-elite referees (n = 21) completed an extended test battery of perceptual-cognitive measures. Participants were assessed using both domain-generic (e.g., sustained attention; working memory) and domain-specific (e.g., pattern recognition) perceptual-cognitive measures within the domain of refereeing.

Results: A multivariate analysis of variance revealed significant differences between groups on the following domain-specific perceptual-cognitive skill measures, with elite referees performing significantly better than their sub-elite counterparts: overall decision-making performance; anticipation; and recall capacity. No significant differences were reported between the elite and sub-elite referees on domain-generic perceptual-cognitive skill measures. A stepwise discriminant analysis highlighted that a combination of five predictor variables, including four domain-specific and one domain-generic perceptual-cognitive skill measure, significantly discriminated elite and sub-elite referees, with 90.7% of referees being classified correctly.

Conclusions: Our findings shed light on the skills underpinning perceptual-cognitive expertise in referees and are consistent with the existing findings on expert outfield players. Implications are discussed with regard to testing and training programs for referees and across other professional domains.

Key words: decision making; refereeing; expert performance; perceptual-cognitive skills.
Expertise has been defined as a unique knowledge and mastery of performance in a particular area. It has been investigated in different domains, including law enforcement, medicine, military, and sports (Ward, Suss, & Basevitch, 2009). A complex recipe of hereditary, environmental, and training factors is needed before the highest level of expertise is attained. Expert athletes may embody the most visible characteristics of all expert performers, showing exceptional physical, technical, emotional, and perceptual-cognitive skills (Ericsson, 2014a). Researchers have been challenged to unravel the mechanisms underlying expert performance and sport scientists have argued that perceptual-cognitive skills play a crucial role in sports performance, particularly at the highest levels (Williams & Ericsson, 2005; Williams, Ford, Eccles, & Ward, 2011). Perceptual-cognitive expertise refers to the ability of performers to identify and process (environmental) information for integration with existing knowledge to facilitate the selection of appropriate responses under time pressure (Marteniuk, 1976). Perceptual-cognitive superiority in the sports domain can be assessed either in a sport-specific context representing the requirements of a competitive and realistic setting (domain-specific skills) or by use of more generic tests with no direct link to the performance setting (domain-generic skills).

A number of domain-specific perceptual-cognitive skills have been identified, which are related to a superior performance. It is well reported that anticipation is an important domain-specific attribute of expertise and it relates to the performer’s ability to use situational probabilities and advance cues emanating from the postural movements of players (Mann, Williams, Ward, & Janelle, 2007; Williams, Huys, Cañal-Bruland, & Hagemann, 2009). Athletes have an idea of what is likely to happen, creating more time to execute an appropriate response. Previously, researchers have shown that anticipation is strongly associated with expertise in various sports, including badminton (Wright, Bishop, Jackson, & Abernethy, 2011), tennis (Tenenbaum, Sar-El, & Bar-Eli, 2000), and soccer (Van der
Kamp, 2011; Williams, 2000). It is often concluded that experts can recognize patterns, or chunks of information, within the game structure to a deeper level than novices, allowing them to subsequently recall the position of players more accurately or to better distinguish previously seen from novel action sequences of play. This ability to recall and recognize sport-specific patterns of play has been shown to be important in different sports where there is considerable time pressure, requiring individuals to selectively attend to the most relevant sources of information (North, Ward, Ericsson, & Williams, 2011). Scientists have also shown clear expertise effects in referee-specific decision-making tasks across various team sports, such as rugby (Mascarenhas, Collins, Mortimer, & Morris, 2005), association football (Catteeuw, Helsen, Gilis, Van Roie, & Wagemans, 2009; Gilis, Helsen, Catteeuw, & Wagemans, 2008; Put, Wagemans, Spitz, Armenteros Gallardo, Williams, & Helsen, 2014), Australian football (Larkin, Berry, Dawson, & Lay, 2011), and ice hockey (Hancock & Ste-Marie, 2013). Moreover, Gilis et al (2008) showed that international assistant referees were significantly more accurate in recalling the spatial positions of football players in complex offside situations than national assistant referees. Ste-Marie (1999) showed that expert judges in gymnastics were significantly better at anticipating upcoming gymnastic elements from advance information. It is suggested that experts only differ in processing skills directly related to their field of expertise and that those skills do not translate to other domains (Ericsson, Charness, Feltovich, & Hoffman, 2006; Feltovich, Prietula, & Ericsson, 2006). However, there can be transfer of processing skill during role transitions, for example when a player becomes a referee (MacMahon, Starkes, & Deaking, 2009).

A second approach towards expertise acquisition is the cognitive component skills or domain-general approach, which examines the relationship between sport expertise and domain-generic perceptual-cognitive skills (Furley & Memmert, 2010; Hill & Schneider, 2006; Nougier, Stein, & Bonnel, 1991). Domain-generic perceptual-cognitive skills, such as
working memory capacity, processing speed, motor inhibition, and attentional abilities, are
determined largely by the processing efficiency of the central nervous system and are tested
outside the sport-specific domain. Research on the cognitive component skills approach has
yielded contrasting findings regarding the relationship between levels of performance on
domain-generic perceptual-cognitive skills and expertise in sport or other domains, where
perceptual-cognitive skills are of utmost importance (Voss, Kramer, Basak, Prakash, &
Roberts, 2010).

In a meta-analysis of 20 studies, Voss et al. (2010) showed that expert athletes
perform better compared to novices on measures of processing speed and several attentional
paradigms. Moreover, elite and sub-elite team sport players’ scores on tasks measuring
creativity, inhibition, cognitive flexibility, and executive control have been compared (Alves
et al., 2013; Verburgh, Scherder, van Lange, & Oosterlaan, 2014; Vestberg, Gustafson,
Maurex, Ingvar, & Petrovic, 2012). These authors found significant variation between the
two levels of athletes, with better scores for the elite players. Additionally, Vestberg et al.
(2012) observed a significant relationship between the performances on domain-generic
perceptual-cognitive skill measures and the number of goals and assists association football
players had scored two seasons later. Similarly, Ghasemi, Momeni, Jafarzadehpur, Rezaee,
and Taheri (2011) reported differences between successful and unsuccessful Iranian referees
on several general visual skills such as facility of accommodation, peripheral vision,
recognition speed, visual memory, and saccadic eye movements. The link between expertise
and generic skill performance has also been demonstrated in medicine and other domains.
For example, Harenberg et al. (2016) showed that the multiple objects tracking score
predicted laparoscopic surgical skill in medical students.

In contrast, other researchers have reported no improved performances on cognitive
measures in favor of (elite) athletes and no consistent relations between domain-generic
perceptual-cognitive skills and domain-specific skills (for a review, see Ericsson, 2014b). For instance, no differences were observed between experts in team handball, expert track athletes and novices on a visual attention test battery (Memmert, Simons, & Grimme, 2009). Furley and Memmert (2010) reported that basketball players and non-athlete college students did not differ in spatial working memory as measured with the ‘Corsi Block-tapping task’. Similarly, generic measures of cognition are not considered predictors of skill level in outfield players in association football (Belling, Suss, & Ward, 2015). In sum, these latter studies claim that experts only differ in processing abilities directly tied to their domain of expertise, in contrast with studies supporting the cognitive component skills approach.

Only a few researchers have examined both domain-generic and domain-specific perceptual-cognitive skill measures to evaluate their relative contribution to expert performance. Helsen and Starkes (1999) adopted a multidimensional approach to predict performance between expert, intermediate, and novice association football players and found that 84% of variance was accounted for by sport-specific skills. The only generic visual component that contributed a small amount (3%) was peripheral horizontal visual range. Very similar results were reported by Ward and Williams (2003), examining the relative contribution of visual, perceptual, and cognitive skills to the development of expert performance in association football players throughout their development. Elite and sub-elite association football players were not discriminated based on their visual function throughout late childhood, adolescence or early adulthood. In contrast, elite players developed superior sport-specific perceptual-cognitive skills allowing them to perform more successfully in each of the respective age groups. These studies mainly used generic visual and optometric parameters, such as static visual acuity and stereoscopic depth sensitivity. These visual ‘hardware’ variables should not be confused with generic measures of enhanced cognitive processing as proposed within the cognitive component skills approach.
In the current study, we extend understanding of the role of domain-generic and domain-specific perceptual-cognitive skills in expert decision making. We examine which components account for the largest proportion of the variance in performance levels using a battery of domain-generic and domain-specific tests. The focus in the current study is on referees in association football, an under-researched target group for whom perceptual-cognitive skills are of utmost importance. In similar vein to the demands on outfield players, referees need to be attentive and to quickly and efficiently make appropriate decisions in dynamic, ever-changing situations. This decision-making process requires an efficient information processing system (MacMahon et al., 2015). Therefore, we included a test battery of four domain-specific and six domain-generic perceptual-cognitive skill measures.

The selected domain-specific skills (decision-making performance; anticipation; recognition/recall capacity) are generally accepted as key performance indicators within sports literature and were adapted for the domain of refereeing (North et al., 2011; Williams et al., 2009). With respect to these domain-specific perceptual-cognitive skills, we hypothesized that elite referees would outperform sub-elite referees because they have acquired specific and more elaborate knowledge from their respective performance environment. Previously, researchers have shown clear expertise effects in decision-making accuracy for referees across various team sports (Catteeuw et al., 2009; Gilis et al., 2008; Put et al, 2014; Hancock & Ste-Marie, 2013; Larkin et al., 2011; Mascarenhas et al., 2005). We also predicted that elite referees would be more accurate than sub-elite referees on the referee-specific anticipation test because their superior indexing and storage of information results in a better perception of relevant information. As previously shown in gymnastic judges, future outcomes can be identified even though not all information is present (Ste-Marie, 1999). Gilis et al (2008) showed that assistant referees in association football were significantly more accurate in recalling the spatial positions of football players in complex
offside situations. Similarly, in keeping with these previous published reports, we predicted that elite referees would demonstrate superior recognition and recall of previously seen foul situations. Although sub-elite referees were expected to encode some important information, they were predicted to have less elaborate representations in long-term memory to help interpret and recognize/recall situations. This more superficial knowledge base would be reflected in a lower decision-making performance, a lower anticipation accuracy score and a decreased recognition/recall capacity compared to elite referees.

The task of a referee requires good perceptual skills, specifically, motion perception, local and global perception, and the fast switching between these modes when needed. In addition, referees must distribute their attention, hold all relevant information from players in working memory, and quickly process this information to make a proper decision. Based on these basic features of refereeing and on face validity, we selected six measures of domain-generic (sustained attention; working memory; processing speed; global-local attention switching; local information processing; global motion perception) perceptual-cognitive skills (Chamberlain, Van der Hallen, Huygelier, Van de Cruys, & Wagemans, 2017; Chamberlain & Wagemans, 2015; de-Wit, Huygelier, Van der Hallen, Chamberlain, & Wagemans, 2017; Langenecker, Zubieta, Young, Akil, & Nielson, 2007; Van Biesen, Mactavish, McCulloch, Lenaerts, & Vanlandewijck, 2016). Furthermore, the selection of the domain-generic perceptual-cognitive skill measures was based on previous publications that examined individual differences in cognitive processing speed and attentional paradigms within sports and other representative domains (Ericsson, 2014b; Voss et al., 2010). With respect to these domain-generic perceptual-cognitive skills, it seems reasonable to assume that elite referees are characterized by more general perceptual-cognitive adaptations since they must constantly switch attention, recognize and appropriately classify situations, and
make quick decisions to act accordingly (Mascarenhas et al., 2005). The measured domain-
generic perceptual-cognitive skills are representative of the task requirements and potentially
important for refereeing performance. If the cognitive component skills approach is valid,
then elite referees should perform better on these measures. **However, in light of the**
diverse findings in the literature, it could equally be argued that, when combining both
domain-generic and specific perceptual-cognitive skills in one analysis, expertise-
related differences would mainly appear for processing measures directly tied to their
field of expertise and to a lesser extent for the domain-generic perceptual-cognitive skill
measures.

**Method**

**Participants**

Two groups of association football referees in Belgium were recruited according to
their competitive level. The group of elite referees (n = 22, mean age 33.1 years, SE = 1.4,
range 24-44) was active in the highest divisions of professional football and had 16 years of
experience as a referee. According to Swann, Moran, and Piggott’s (2015) classification
system of elite athletes, these referees were classified as successful elite. A second group
consisted of sub-elite referees (n = 21, mean age 33.4 years, SE = 1.8, range 23-50), all
actively involved in Belgian amateur levels without any refereeing experience at the
professional level. The sub-elite referees had on average 12 years of experience as a referee.
Participants provided written informed consent and the study was approved by the local
University ethics committee (G-201504218).

**Materials and Procedure**

**Altogether,** four separate tests to measure domain-specific perceptual-cognitive
skills and six measures of domain-generic perceptual-cognitive skills **were administered.** At
the start of each test, participants were seated at a distance of approximately 60 cm from the
computer screen and were given standardized instructions. A break of 2-5 minutes was provided on completion of each test. To minimize potential order effects, participants completed the tests in a counterbalanced manner. The total duration of the test session was approximately 90 minutes.

Domain-specific perceptual-cognitive skills.

Video clips of foul play situations were captured using a high-definition video camera (Sony, PMW-F55 4K, digital cinema) from the first-person perspective of the additional assistant referee (i.e., from the position next to the left goal post). These situations were simulated by trained players and represented key situations normally encountered in a regular game. They consisted of both open play and corner kick situations. During the corner kick situations, multiple attacking and defending players were involved in front of the goal and two players interacted for a possible infringement (for a more detailed overview of the test footage, see Spitz, Put, Wagemans, Williams, & Helsen, 2016). Three independent and experienced ex-international referees, still involved as referee match observers for the Union des Associations Européennes de Football (UEFA), were involved in the final selection of the video clips and the determination of the reference decisions. Altogether, three separate sets of 20 game situations were selected, including video clips for which all three experts reached a consensus decision. The three sets of clips were used to measure four domain-specific perceptual-cognitive skills: 1) decision-making performance; 2) anticipation; 3) recognition capacity; and 4) recall capacity. The video clips were edited within Final Cut Pro 6.0.6 (Apple, Inc., USA) and were presented on a 17-inch computer monitor. One video clip was always used to help referees familiarize themselves with the test. Participants received no feedback about their performance. A previous study showed excellent levels of test-retest reliability for the specific decision-making task in association football referees, with intra-
class correlation coefficients ranging from 0.76-0.82 (Spitz, Put, Wagemans, Williams, & Helsen, 2017).

**Decision-making performance.**

To measure overall decision-making performance, referees had to assess whether a foul occurred in each of the 20 video clips (10 open play situations; 10 corner kick situations). Immediately after each clip, they were instructed to make a technical (no foul; indirect free kick; direct free kick; or penalty kick) and a disciplinary decision (no card; yellow card or red card) according to Law 12: Fouls and Misconduct (FIFA, 2015).

Decision-making accuracy scores were calculated as the percentage of correct decisions, that is, decisions that were in correspondence with the reference decision.

**Anticipation.**

The temporal occlusion paradigm was used to test the referee’s ability to anticipate and predict the outcome of a situation. Twenty video clips, 10 open play, and 10 corner kick situations, were occluded just before the moment of defender-attacker contact and participants had to anticipate what would happen next (Fig. 1). At the moment of occlusion, the screen turned black and participants were asked to make the appropriate technical and disciplinary decision according to Law 12: Fouls and Misconduct (FIFA, 2015). Anticipation accuracy scores were calculated as the percentage of correct decisions.

****Figure 1 near here****

**Recognition/recall capacity.**

Participants were presented with 20 corner kick situations, of which 10 had been presented 15 minutes before in the decision-making performance test and 10 clips were entirely new situations. The video clips were presented in a random order. For each video
clip, participants first had to indicate whether it had been shown earlier, or not (binary decision: yes or no). Recognition accuracy was calculated as the percentage of correct decisions. In 13 out of 20 situations, a foul occurred and referees subsequently performed a recall task. The recall task required referees to recall, as accurately as possible, the location of the contact between the attacking and defending player by clicking onto a blank football field showing the same viewing perspective as the original video clip (Fig. 2). The coordinates of the locations indicated by the referee were compared to the actual locations of the contact. Mean recall deficit scores were calculated and expressed in computer screen pixels in vertical direction (Y-axis) and horizontal direction (X-axis).

****Figure 2 near here****

Domain-generic perceptual-cognitive skills.
As discussed in the introduction, six measures of domain-generic perceptual-cognitive skills were selected based on a referee’s task requirements as well as on previous publications that examined skill differences in cognitive processing speed and attentional paradigms (Ericsson, 2014b; Voss et al., 2010). An overview of the different tests is given in Table 1. The reliability of these tests has previously been reported. The domain-generic tests were presented on a 15.6-inch computer monitor.

****Table 1 near here****

Sustained attention, inhibitory control.
A parametric Go/No-Go Test with three levels of increasing difficulty (for a detailed description of the task, see Langenecker et al., 2007) was used to assess referee’s sustained
attention and inhibitory control. A serial stream of letters, 500 ms for each letter, was presented and participants were required to respond (pressing the space bar of a keyboard) as quickly as possible whenever a target letter ‘x’, ‘y’ or ‘z’ appeared within the serial stream. Referees had to withhold their responses when the target letters were shown in a repeating order. An efficiency score for level 3 of this test \[(((5 \times \text{percentage correct targets} + \text{percentage correct inhibition trials}) / 6) / \text{reaction time}) \times 100\] was measured in accordance with previous research by Votruba and Langenecker (2013).

Working memory.

The Corsi Block-Tapping Task (see Van Biesen et al., 2016), an adaptation of the original Corsi task (Corsi, 1972), was used to assess spatial working memory. Fifteen random sequences of lightening blocks were presented and referees had to reproduce these block sequences by tapping on a touch screen in the same order as the blocks light up. The length of a sequence started with two blocks and increased or decreased depending on the correctness of the participant’s response. The working memory score was calculated by taking the best average score of five subsequent trials.

Processing speed.

A Simple Reaction Time Test was used to assess processing speed (see Van Biesen et al., 2016). On 12 subsequent trials, participants were asked to respond (pressing the space bar of a keyboard) quickly and accurately whenever a white target circle appeared on the computer screen. To prevent participants from anticipating, trials were shown with randomized time intervals. Reaction time was measured as the mean score in ms over 10 trials (i.e., after removal of the best and worst time).

Global-local attention switching.

The Navon Level-Switching Task was used to measure referee’s ability to selectively attend to and switch between global and local levels of hierarchical visual stimuli (for a
detailed description of the task, see Chamberlain & Wagemans, 2015). Thirty-two trial pairs were presented in a random order. On each trial, a large shape (global level) made up of 18 smaller shapes (local level) appeared. Referees were asked to indicate quickly and accurately whether a square (pressing the F key) or circle (pressing the J key) was present, which could be either at the global or local level (Fig. 3A). Mean Global Local and Local Global reaction time costs were computed based on the subtractions of local from global (GL pair) and global from local (LG pair) trial reaction times.

**Figure 3 near here**

Global motion perception.

Global motion perception was assessed using the Coherent Motion Test (for a detailed description of the task, see Chamberlain & Wagemans, 2015). During this test, 600 moving dots were presented on the computer display for 500 ms (Fig. 3B). A proportion of the dots in every trial, 200 trials in total, moved in the same direction (i.e., the ‘global motion’). Participants were instructed to indicate the global motion by pressing the matching arrow key (i.e., up, down, left or right) on the keyboard. An overall accuracy score was calculated as the percentage of correct decisions.

Local information processing.

Local information processing was assessed using a modified version of the Embedded Figures Test (see de-Wit et al., 2017). On each trial, participants had to detect a local target figure embedded within a more complex pattern. Three response alternatives were presented, only one containing the local target figure (Fig. 3C). Participants were asked to indicate accurately and quickly in which of the complex patterns the target figure could be recognized by using the computer mouse (64 trials in total). The dependent variable for this
task was the mean decision time to identify correctly the local target figure in its embedded form.

Statistical Analysis

Separate one-way MANOVAs were performed on the domain-specific and domain-generic perceptual-cognitive measures with group (elite vs. sub-elite) as the between-participants factor. A preliminary check for MANOVA’s statistical assumptions of homogeneity of variance and covariance matrices using the Box’s M-test indicated no significant effects for both the domain-specific $[F(15, 6734) = .739, p > .05]$ and domain-generic perceptual-cognitive skills $[F(28, 5829) = 1.198, p > 0.05]$. Follow-up pairwise comparisons were used where appropriate. To increase the power and representativeness of the comparison of the domain-specific skills, the video clips on which both groups performed correctly or performed below chance level were deleted from the analysis. Finally, all domain-specific and domain-generic perceptual-cognitive skill variables were analyzed together to determine which combination of variables could reliably predict group membership. A stepwise discriminant function was used in which group (elite vs. sub-elite) was the dependent variable. Correlations of all variables were assessed to determine if multicollinearity ($r > .90$) was present among the variables (Tabachnick & Fidell, 2007). Multicollinearity was not reported. All statistical analyses were carried out using IBM SPSS statistical program version 24. Effect sizes were calculated as partial eta-squared values ($\eta_p^2$). Partial eta-squared values of .01 to .06, .06 to .14 and >0.14 were interpreted as small, medium, or large effects, respectively. A $p$-value of < .05 was considered significant (Field, 2005).

Results

Domain-Specific Perceptual-Cognitive Skills
MANOVA revealed that there was a significant overall difference between the elite and sub-elite referees for the domain-specific perceptual-cognitive skill measures [Wilks’ lambda = .521, $F(5, 37) = 6.798$, $p < .001$, $\eta^2_p = 0.479$]. The mean scores, standard errors and effect sizes are presented in Table 2. For the overall decision-making performance, elite referees ($M = 63.1\%$, $SE = 2.1$) were more accurate than sub-elite referees ($M = 55.4\%$, $SE = 2.1$) [$F(1, 41) = 6.993$, $p = .012$, $\eta^2_p = .146$]. Also, elite referees ($M = 63.9\%$, $SE = 2.0$) were significantly more accurate than their sub-elite counterparts ($M = 57.4\%$, $SE = 2.1$) in anticipating what would happen next when the video clip was occluded just before the moment of contact [$F(1, 41) = 4.910$, $p = .032$, $\eta^2_p = .107$]. Furthermore, elite referees ($M = 68.3$ pixels, $SE = 5.3$) were better able to recall the exact position of the foul in vertical direction on a blank football field when compared with sub-elite participants ($M = 92.7$ pixels, $SE = 6.2$) [$F(1, 41) = 8.970$, $p = 0.005$, $\eta^2_p = 0.180$]. No significant differences were found for recognition capacity or for the recall deficit score in horizontal direction.

**Table 2 near here**

**Domain-Generic Perceptual-Cognitive Skills**

MANOVA showed that there were no significant differences between the elite and sub-elite referees on the domain-generic perceptual-cognitive skill measures [Wilks’ lambda $= 0.873$, $F(7,35) = 0.977$, $p = 0.463$, $\eta^2_p = 0.163$]. The mean scores and standard errors for the different tests are presented in Table 2.

**Combination of Domain-Specific and Domain-Generic Perceptual-Cognitive Skills**

The results of a stepwise discriminant analysis are illustrated in Table 3. The analysis resulted in a combination of five predictor variables that significantly discriminated the elite and sub-elite referees ($\chi^2 = 36.7$, Wilks’ $\lambda = .386$, $p < .001$). The discriminant function
classified 90.7% of the referees correctly. According to the model, the remaining 9.3% of cases included one sub-elite referee falsely classified as an elite referee and three elite referees falsely classified as sub-elite referees. The five measures that successfully discriminated between groups were the recall capacity in vertical direction, overall decision-making performance, the efficiency level of the Go/No-Go test, the recognition capacity, and anticipation performance. The large standardized coefficients reflect the importance of these variables. The eigenvalue for the function (1.592) suggests that the discriminating power is high, with a canonical correlation of .784, indicating that approximately 61% of variability is accounted for by these five variables. Standardized and canonical correlation coefficients are provided in Table 3. Final conclusions can be made when interpreting the signs of the discriminant function. Positive signs denote significantly higher mean referrals for elite referees, while negative signs indicate significantly higher referrals for sub-elite referees. As such, higher scores on the overall decision-making test and the anticipation test contribute more to the elite centroid, whereas better performances on the Go/No-Go test and the recognition capacity test contribute more to the sub-elite centroid. Attention needs to be paid regarding the contribution of the recall deficit score in vertical direction. Participants with higher scores for recall deficit actually performed worse and this explains why these scores contribute more to the sub-elite centroid.

****Table 3 near here****

Discussion

Scientists have been challenged to unravel the perceptual-cognitive skills underlying expert anticipation and decision making. Perceptual-cognitive skills can be classified as domain-generic or domain-specific depending on their relatedness with the actual
performance setting. Only a few researchers have examined both domain-generic and
domain-specific perceptual-cognitive skill measures simultaneously to evaluate their relative
contribution. Moreover, existing published reports have mainly focused on visual ‘hardware’
variables, with generic measures of enhanced cognitive processing being excluded (Helsen
& Starkes, 1999; Ward & Williams, 2003). We examined whether domain-generic and/or
domain-specific perceptual-cognitive skill variables discriminate between elite and sub-elite
referees. We expected that expertise-related differences would mainly be evident on
processing measures directly tied to the field of expertise and would not be evident on the
domain-generic perceptual-cognitive skill measures.

As predicted, significant differences were reported between elite and sub-elite
referees on the domain-specific perceptual-cognitive skill variables. These findings are in
line with a well-established body of research showing that experts and novices differ in
information processing capabilities directly related to the domain of expertise (Ericsson, et
al., 2006; Feltovich, et al., 2006). The elite referees showed superior performance in
assessing, anticipating, and recalling foul play situations. First, overall decision-making
performance was assessed using video clips of foul play situations. This task is
representative and closely aligned with the demands of referees’ everyday performance
context. It was therefore not surprising that elite referees outperformed those from a lower
level. Similar results have previously been obtained in a group of Australian football referees
(Larkin et al., 2011) and assistant referees in association football (Catteeuw et al., 2009;
Gulis et al., 2008; Put et al., 2014).

The ability to pick-up information arising early in an action or game sequence is
another important perceptual-cognitive skill in expert athletes (Williams et al., 2009). It was
the first time that this domain-specific perceptual-cognitive skill measure was applied to a
group of association football referees using a temporal occlusion paradigm. In line with
previous results in gymnastic judges (Ste-Marie, 1999), the elite group appeared better able to utilize advance cues from the foul play situations, which enabled them to anticipate possible future infringements. Moreover, elite referees were more accurate than sub-elite referees in recalling the exact position of the foul in the vertical direction. As a result of their previous experience of refereeing at elite level, referees show superior information processing and performance. A previous published report has shown that elite referees fixate on specific body parts of the players that are involved in the infringement such as the upper or lower body part, and this might explain why the observed differences in recall performance were only present in the vertical direction and not in horizontal direction (Spitz et al., 2016). The recognition accuracy scores were just above chance level and elite referees did not differentiate themselves in the ability to recognize whether foul play situations in complex corner kick situations had been shown earlier. As distinct from players, referees seem to be focused upon actions and movements at the location around the ball and the contact and are less likely to identify structure and patterns forming within the play (Larkin et al., 2011; Spitz et al., 2016).

A selection of generic skill tests was used to measure cognitive and perceptual functions outside the specific domain of refereeing. There were no significant differences between the elite and sub-elite referees on these domain-generic perceptual-cognitive skill measures. It is suggested that elite referees, through engagement in referee-related activities at a higher level, are able to extend the limits imposed by basic information processing capacities (Ericsson, 2014a). The accumulation of deliberate practice allows elite referees to acquire domain-specific knowledge and elaborate representations in long-term memory. These higher-order representations and domain-specific retrieval structures facilitate information processing and performance within the domain of expertise, not only to make
better decisions but also to anticipate and recall potential fouls in an accurate way (Ericsson & Kintsch, 1995).

The present study is unique since it combines a battery of domain-generic and domain-specific perceptual-cognitive skill measures to explore the processes underpinning elite and sub-elite performance in refereeing. The stepwise discriminant analysis revealed that five variables were significant predictors of group membership, correctly classifying 90.7% of the referees. Decision-making performance, anticipation, and recall capacity were positively associated with elite group membership, indicating that better performance on these domain-specific skill measures contributed to elite group membership. On the contrary, the better the performance on the Go/No-Go test and the recognition capacity test, the more likely participants belonged to the sub-elite group. The ability to recognize structure and patterns forming within the play can thus not be seen as a determining skill for referees. An explanation might be that by paying attention to patterns and players who are not involved in the presented foul situation, referees miss information from the players who actually interact, which is crucial for decision making. The negative association of the Go/No-Go test with elite group membership indicates that high sustained attention and inhibitory control are not crucial skills and are not important for referees belonging to the elite group. Sometimes less can even be more. This finding contradicts the conclusions of the cognitive component skills approach because generic attentional capabilities seem not to be necessary and an intuitive, uninhibited style of decision making might even be better (Plessner, Schweizer, Brand, & O’Hare, 2009; Raab & Gigerenzer, 2015). This issue definitely needs further investigation but overall, our results are consistent with previous published reports and suggest that domain-specific skills are more predictive of elite group membership than domain-generic perceptual-cognitive skill measures (Belling et al., 2015; Helsen & Starkes, 1999; Ward & Williams, 2003).
Our findings may have important implications for the development of training programs and test protocols for referees. Currently, physical performance of referees is assessed using fitness tests, specifically developed to mimic the physical demands of matches. Decision-making skills, however, are tested using a written test of the Laws of the Games. This written test is not representative of the perceptual and cognitive demands of the sport context and the use of realistic video-based test protocols may be an effective supplementary tool in the selection process for elite level referees. Furthermore, a systematic training approach to developing domain-specific perceptual-cognitive skills, including decision making, anticipation, and recall, would seem to have significant merit in refereeing. Previously, researchers have indicated that the perceptual-cognitive skills of referees and players can be improved with specialized coaching and training (Put, Wagemans, Spitz, Williams, & Helsen, 2015; Schweizer, Plessner, Kahlert, & Brand, 2011). In contrast, there is limited evidence to suggest that training programs advertised to improve generic attentional skills transfer onto the field and lead to better performances, beyond the tasks that are actually being trained (Abernethy & Wood, 2001; Owen et al., 2010; Romeas, Guldner, & Faubert, 2016). It can be argued that generic perceptual-cognitive training increases performance in the field because of an expanded capability to perceive and process movement patterns from players. Generic perceptual-cognitive skills might improve the foundation upon which expert decision making is based and free up resources for other attentional demands (Faubert & Sidebottom, 2012). These explanations are based on the premise that domain-generic skills make an important contribution to expertise; a premise not confirmed by our results. Further research is needed to examine issues related to the specificity of training and how perceptual-cognitive skills can be developed in sports and refereeing.
Some limitations of the present study need to be acknowledged. First, the small sample size was due to the fact that the number of referees involved at the top level is limited. **In future studies, researchers should consider to use a larger sample size and an equal number of domain-generic and domain-specific tests.** Furthermore, the relative importance and influence of domain-generic and domain-specific perceptual-cognitive skills may vary according to the nature of each domain. For example, in a study from Hüttermann, Memmert, and Simons (2014), it was concluded that athletes with different demands in their sport showed differences in a generic measure of attention breadth (primarily horizontal or vertical attention). As such, we have to consider the sport-specific context and caution is warranted not to generalize specific findings from one context to another. Although the results may be context- or task-specific to some extent, the approach might have practical utility in a variety of professional settings in which time-constraint decision making in a complex environment is key (e.g. drivers, medical doctors involved in reading medical imaging outcomes, pilots, police officers). Understanding the entire process of skill acquisition requires the exploration of the intermediate steps and changes that performers go through in their development. Therefore, a cross-sectional comparison of experts and novices is not sufficient and longitudinal study designs and training interventions are required. Greater collaboration between national sporting organizations and academic institutions could facilitate data collection using longitudinal and quasi-longitudinal designs. Moreover, cooperation of coaches and scientists could provide deeper insights in the area of talent identification and the skills that are needed at each level of a referee’s career (Buekers, Borry, & Rowe, 2015). Further work is clearly needed to clarify the connections between perceptual-cognitive skills, training and the level of attained expertise in refereeing and other domains of expertise.
In this paper, we reported a novel attempt to evaluate the relative contribution of domain-generic and domain-specific perceptual-cognitive skills to expert performance in the domain of association football referees. A unique combination of 6 domain-generic and 4 domain-specific measures of perceptual-cognitive skill was employed. While there remains debate in the literature in regards to the relative importance of domain-generic and domain-specific skills, our findings highlight the much greater contribution of domain-specific perceptual-cognitive skills to expert performance in association football referees. These findings have significant implications for the development of systematic testing and training programs to enhance perceptual-cognitive skill in sports officials, and potentially in many other domains where anticipation and decision making are paramount to superior performance.
References


Figure captions

Figure 1: Example of a video clip of an open play situation showing the moment of occlusion in the anticipation task.

Figure 2: A. Example of a video clip of a corner kick situation showing the exact moment and place of the foul (hand ball). B. Response sheet for the recall task with a blank football field showing the same viewing perspective as the original video clip.

Figure 3: A. Navon Level–Switching Task: Examples of a global circle trial and a local square trial. B. Example of the Coherent Motion Test: The black arrows are used to illustrate the global direction of motion and are not present in the actual test. C. Example of the Embedded Figures Test: The local target figure is only present in one of the complex patterns.
### Table 1

**An overview of the domain-generic perceptual-cognitive skill measurements**

<table>
<thead>
<tr>
<th>Test</th>
<th>Measured skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parametric Go/No-Go Test:</td>
<td>Sustained attention, inhibitory control, set-shifting, processing speed,</td>
</tr>
<tr>
<td>Corsi Block-Tapping Task:</td>
<td>Working memory</td>
</tr>
<tr>
<td>Reaction Time Test:</td>
<td>Processing speed: simple reaction time</td>
</tr>
<tr>
<td>Navon Level-Switching Task:</td>
<td>Ability to switch between global and local levels of hierarchical visual stimuli</td>
</tr>
<tr>
<td>Coherent Motion Test:</td>
<td>Ability to integrate multiple local motions in the perception of global motion</td>
</tr>
<tr>
<td>Embedded Figures Test:</td>
<td>Ability to find a local target in global patterns</td>
</tr>
</tbody>
</table>
Table 2

Mean scores for the domain-specific and domain-generic perceptual-cognitive skill measures (standard errors) for the elite and sub-elite group of referees with corresponding F-values, p-values, and effect sizes

<table>
<thead>
<tr>
<th></th>
<th>Elite</th>
<th>Sub-elite</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_{p}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain-specific perceptual-cognitive skill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision-making accuracy (%)</td>
<td>63.1</td>
<td>55.4</td>
<td>6.993</td>
<td>.012</td>
<td>.146</td>
</tr>
<tr>
<td>Anticipation accuracy (%)</td>
<td>63.9</td>
<td>57.4</td>
<td>4.910</td>
<td>.032</td>
<td>.107</td>
</tr>
<tr>
<td>Recognition accuracy (%)</td>
<td>63.6</td>
<td>68.6</td>
<td>2.849</td>
<td>.099</td>
<td>.065</td>
</tr>
<tr>
<td>Recall deficit X-axis (pixels)</td>
<td>110.3</td>
<td>94.8</td>
<td>2.788</td>
<td>.103</td>
<td>.064</td>
</tr>
<tr>
<td>Recall Deficit Y-axis (pixels)</td>
<td>68.3</td>
<td>92.7</td>
<td>8.970</td>
<td>.005</td>
<td>.180</td>
</tr>
<tr>
<td><strong>Domain-generic perceptual-cognitive skill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple reaction time (s)</td>
<td>.393</td>
<td>.399</td>
<td></td>
<td>.463</td>
<td>.163</td>
</tr>
<tr>
<td>Corsi Block working memory score</td>
<td>6.5</td>
<td>6.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency score Go/No-Go Test</td>
<td>16.4</td>
<td>17.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coherent motion accuracy (%)</td>
<td>65.8</td>
<td>66.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navon LG reaction time cost (s)</td>
<td>0.079</td>
<td>0.073</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navon GL reaction time cost (s)</td>
<td>-0.206</td>
<td>-0.133</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embedded figures decision time (s)</td>
<td>5.46</td>
<td>5.50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3

Standardized canonical discriminant function coefficients and pooled within-group correlations for the five predictor variables and cumulative proportion of explained variance

<table>
<thead>
<tr>
<th></th>
<th>Standardized coefficient</th>
<th>Correlation</th>
<th>Cumulative proportion of explained variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall Deficit (Y-axis)</td>
<td>-.580</td>
<td>-.371</td>
<td>18%</td>
</tr>
<tr>
<td>Decision-making accuracy</td>
<td>.910</td>
<td>.327</td>
<td>30%</td>
</tr>
<tr>
<td>Efficiency score Go/No-Go Test</td>
<td>-.832</td>
<td>-.279</td>
<td>47%</td>
</tr>
<tr>
<td>Recognition accuracy</td>
<td>-.546</td>
<td>-.209</td>
<td>55%</td>
</tr>
<tr>
<td>Anticipation accuracy</td>
<td>.513</td>
<td>.513</td>
<td>61%</td>
</tr>
</tbody>
</table>
A. Global circle trial

Local square trial

B.

C.

Target figure

Correct

Alternative 1

Alternative 2
• The importance of perceptual-cognitive skills for football referees is examined
• Elite referees do not excel in domain-generic perceptual-cognitive skill measures
• Similar to players, elite football referees excel in domain-specific skill measures
• A combination of five variables were significant predictors of group membership