1-1-2009

Physical performance and decision making in association football referees: A naturalistic study.

Duncan R. Mascarenhas
Glyndwr University, d.mascarenhas@glyndwr.ac.uk

C Button

D O'Hara

M Dicks

This is the author's post print version after refereeing. 'This article was published in The Open Sports Sciences Journal, in 2009, by Bentham Sciences. The published article is available at http://www.bentham-open.org

Recommended Citation

Mascarenhas, D. R. D., Button, C., O'Hare, D. & Dicks, M. (2009) 'Physical performance and decision making in association football referees: A naturalistic study'. Open Sports Sciences, 2(9), 1-9

Method

Participants

Match Officials. Five New Zealand Football Championship (NZFC) referees, officiating in seven NZFC games agreed to participate in the study. Two referees participated twice in games involving different teams. All held the NZ badge qualification and two were Fédération Internationale de Football Association (FIFA) qualified referees. This represents half of the referees who officiate in this league. They were all male, aged from 31-43 yrs old (mean = 38.2 yrs, s = 5.89) and all had refereed in the National League for at least 4 years. New Zealand Soccer provided the estimated maximal oxygen uptake ($\dot{V}O_{2\,max}$) scores achieved by each referee completing a 12-minute run[see 36] at the beginning of the season (mean = 55.4 $ml.kg^{-1}.min^{-1}$, s = 5.90). All the referees received a detailed explanation of the purpose of the research and were assured of the confidentiality of their datasets prior to the study commencing.

Referee Panel. A separate panel of five experienced referees also participated by providing independent judgments of selected incidents in the games from edited video clips. Participants were recruited from members of the local referees' society, all of whom officiate alongside the subject sample during New Zealand's winter competitions. All held the NZ badge qualification and one was a FIFA qualified referee. The age and experience of these referees (mean age = 38.2 yrs, sd = 12.6; mean experience = 9.5 yrs, sd = 7.1) also closely matched that of the subject sample. These individuals were sent a copy of the DVD and a set of rating sheets and asked to complete the ratings in their own time. The rating sheet included instructions asking them to watch each clip as many times as they felt necessary in order to be confident in their decision and to indicate the number of viewings required on the sheet. They also indicated whether the clip was of sufficient quality and held enough information to make an informed decision and any additional comments they felt necessary to include about the clip.

Procedure

Data were collected at seven home games of Otago United Football Club in the NZ Football Championship (played between November 2005 and February 2006). Each match was recorded by two JVC-2000 DV cameras from elevated positions at the top of the main grandstand and on the opposite side of the pitch. The pitch side camera was manually operated to provide a detailed close-up view of the action, while the grandstand camera, also manually operated, recorded a wide-angle view including both the active play and the position of referee.

Each referee was fitted with a HR monitor and a SPI-10 Global Positioning System (GPSports, Fyshwick, Australia) transmitter 45 minutes before the start of each game. The GPS transmitter was worn in a light harness placed over the shoulders and under the referee's shirt. The GPS equipment collects and stores positional data at a sampling frequency of 1 Hz by comparing signals from between 6 and 9 satellites. The equipment also records, by radio telemetry, the HR signals from a strap worn around the referee's chest. The initiation of video and GPS recording was manually synchronised 30 minutes before the kick-off of each game to ensure that all data had the same timeline. The data recording ended as the referee left the pitch at the end of each game.

GPS Analysis

The GPS transmitter recorded the referee's position, speed of movement, and HR at 1-second intervals throughout each game. A recent assessment of the validity of this GPS system has found a relatively small systematic overestimation of absolute distance (within 4.8% ± 7.2% [37]). From the raw data, a number of other variables related to physical demand were calculated by the GPS[1] equipment's software (such as frequency and distance of sprints/jogs/walks, and percentage of time spent engaging at different exercise intensities). The following locomotor categories were used; standing and walking 0-7 $km.h^{-1}$; jogging 7–12 $km.h^{-1}$; moderate running 12-18 $km.h^{-1}$; and sprinting above 18 $km.h^{-1}$ (adjusted from speeds used by Drust, Reilly, & Cable[38]). These four categories were subsequently grouped into two locomotor categories: (1) low intensity activity encompassing all activities below 12km/h; and (2) high intensity running that encompasses all running above 12 km/h. The percentage of maximum HR (%HRmax) was calculated by dividing each participant's raw HR scores with their maximal HR value achieved throughout the game, since maximum HR has been shown to be consistently higher during match observations than during lab based tests[19].

Video Analysis

The videotapes of each match were subsequently analysed by an experimenter (a qualified football association referee) who edited together all the clips where a foul, or potential foul had occurred. These tapes were then reviewed by two other experimenters, one of whom was a qualified referee, to identify referee decisions when contact occurred between opposing players or a potential handball offense occurred. These included fouls and misconduct incidents, with a range of challenges that required the referee to decide if players' had been tripped, kicked, pushed, charged, jumped, or held, or committed a handball offense (as stipulated by the laws of the game, FIFA, 2007[39]), and they included incidents where the referee apparently missed a decision, as adjudged by the experimenter. We also controlled for the influence of the assistant referees in helping the referee during certain incidents. For example we chose not to analyse offside situations and instead focused on incidents in which the referee alone made the decision. In all except one of the cases (1 of 144) the pitch-side, close-up camera perspective was used to identify tackle and handball incidents.

Incidents from the seven games were professionally edited using Final Cut Pro (version 3.0.4 for Mac OSX). Each clip was preceded by a title explaining the clip number and included approximately 5-seconds of 'lead-in' of preceding action to orientate the viewer to the context of the game, and each clip finished approximately 1-second after the incident[cf. 15]. Thus the clips ranged from about 6 to 10 seconds in length. Immediately after each incident the volume was dubbed out to remove crowd and player reactions. In cases where the match referee was in the frame he was digitally occluded with a black rectangle, again after the incident had occurred, to ensure that the viewers were not able to see or be influenced by his decision, as the panel's role was to adjudicate on the decision, not the match referee's performance. Subsequent analysis of the referee panel's responses on the quality of each clip and their additional comments revealed that this did not affect the panel's decisions or the number of viewings required. At the start of each set of clips for each game a head title came up on screen depicting the start of the game, with a "half-time" title after the first half clips to allow the viewers to orient themselves to the direction of play.

From this editing process 144 foul incidents from the 7 games (approximately 21 per game) were then transferred onto DVD, with each clip indexed so that viewers could easily review each clip by the push of a button. Copies of the DVD were then sent to an independent group of 5 experienced referees. Using a pre-prepared questionnaire this 'expert panel' gave independent decisions for each of the video clips, indicating their decision, the number of viewings required to arrive at this decision and a space to comment on the quality of the clip.

We used the number of times that panel members had to view each clip to provide an indirect indication of decision difficulty. It should be acknowledged that it is possible that this difficulty might reflect inadequacies in the video clip rather than the inherent difficulty of assessing the situation. However, none of the clips were reported to be of insufficient quality to be able to make an informed decision. There were also very few critical comments from the panel members about the clip quality and the panel subsequently confirmed that repeated viewings were necessary for more difficult incidents. Therefore it is reasonable to assume that those incidents that were more difficult for the panel members to judge (due to the speed of events, nature of the player contact etc) would also have been more difficult for the match referees.

Statistical Analysis

Various aspects of the GPS and HR data were summarised for each referee and compared between the first and second half with paired sample t-tests. The expert panel questionnaires were collated to identify incidents where a consensus decision had been made. The panel was deemed to have reached agreement when at least 3 of the 5 judges awarded possession to the same team. The uniformity between the judges for each clip was further quantified using correlation. Only the clips in which panel agreement was achieved (n = 127) were submitted to further analysis. This sub-set of incidents was contrasted with the match-day referees' actual decisions. 'Accurate' and 'inaccurate' decisions made by the match-day referee were grouped and the GPS and HR data associated with these incidents were compared with independent t-tests. For each dependent variable, the assumption of homogeneity of variance was confirmed prior to any further statistical analyses being conducted. Non-parametric Friedman's analysis of variance was used to compare the differences in decision accuracy in each period of the game. For categorical data (i.e., difficulty of decisions) chi-square tests were employed. The level of statistical significance was set at $P = 0.05$.

Results

Movement and Heart-Rate Analysis

In the present study referees covered 10,323 m on average (s = 486 m) during a game. Whilst the referees appeared to cover more ground in the first half than the second half (see Table 2) this difference was not statistically significant ($P > .05$). Despite the trend for distance covered being greater in the first half, the second half of games typically lasted longer than the first half (on average by 1 min 29 sec). The referees' average HR during playing time was 163 bpm (s = 8.6 bpm, 84% HRmax), with a higher mean HR in the first half in comparison to the second half ($P < .05$). In one case, the referee's HR was 15 bpm less in the second half, perhaps as a consequence of this referee sustaining a particularly high HR (175 bpm) in the first half.

Table 2. Summary of mean locomotion data comparing first and second half performance. * denotes statistically significant difference between first and second half ($P < .05$)

	First half		Second half		Statistical comparison
	M	SD	M	SD	
Distance Covered (m)	5283	449	5040	525	t(6)=1.97, P = .10, d = 0.48
Duration (mins: sec)	47:34	1:03	49:06	0:57	t(6)=2.34, P = .06, d = 1.53
*Heart Rate (b/min)	166	7.4	160	9.4	t(6)=2.48, P < .05, d = 0.71
Heart rate max (%)	85	3.0	82	5.3	t(6)=2.23, P = .07, d = 0.70
Time spent sprinting (%)	3	1.4	2	1.3	t(6)=.86, P = .43, d = 0.74
Time spent running (%)	12	3.3	11	3.7	t(6)=1.74, P = .13, d = 0.29
*Time spent jogging (%)	22	2.8	20	2.2	t(6)=4.81, P < .01, d = 0.79
*Time spent walking (%)	63	4.8	67	6.6	t(6)=3.01, P < .03, d = 0.69

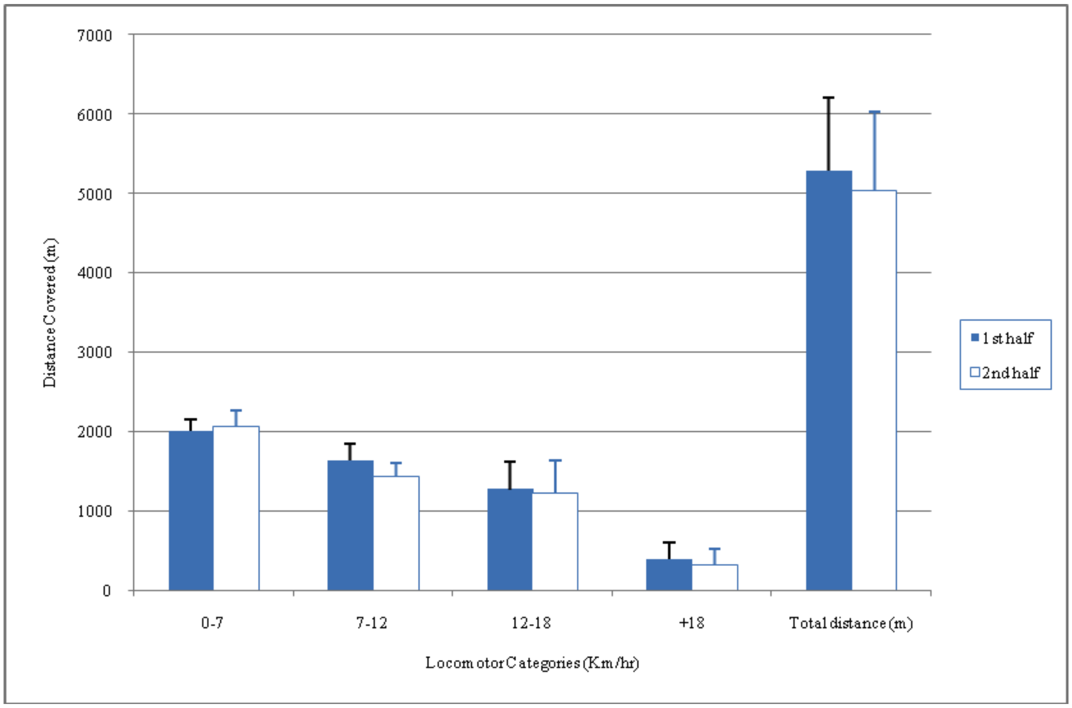

Figure 1: Distances covered in different locomotor categories between the 1st and 2nd halves (error bars indicate standard deviation amongst referees).

The referees performed similar levels of high-intensity running in the first and second half (mean time = 30% vs. 31% respectively). Throughout the game referees spent 65% (s = 5.9) of the time standing or walking, 21% (s = 2.8) jogging, 12% (s = 3.4) moderate running and 2% (s = 1.3) sprinting. In the first half, the referees spent proportionally less time standing and walking and more time jogging than in the second half (see Table 2). In terms of distance covered within each speed zone there were similar distributions in each half (see Figure 1).

Decision Making Analysis

Expert Panel. The coefficient of correlation between the judges' ratings indicated a high degree of uniformity (range = 0.3 to 0.63; all statistically significant at $P < .018$). Agreement was achieved on 88% of the clips (127/144), with an average of 4 out of the 5 experts agreeing on each decision. The difficulty of each decision, which was determined from the number of viewings reported by each judge, was evenly distributed across game time (0-15 mins = 1.7 viewings, 15-30 mins = 1.6, 30-45 mins = 1.5, 45-60 mins = 1.7, 60-75 mins = 1.6, 75-90 mins = 1.7).

Accuracy of Decisions. The match-day referees made the same decision as the panel on 64% of occasions (awarding 81 out of 127 clips correctly). From the occasions where the referee and panel's decision did not concur (n = 46, 36%), 54% (n = 25) arose because the panel saw no infringement and thus the referee erred in penalising the challenge, and 41% (n = 19) were for missed decisions. Therefore, there did not seem to be any strong bias towards over-penalising or under-penalising amongst incorrect decisions.

There was also a reasonable balance between the match-day referees and the expert panel in awarding decisions to either the home or away teams. The match-day referees awarded 45 decisions (35%) in favour of the home team and 39 decisions (31%) to the away team and decided there was no-infringement in 43 cases (34%). This finding could possibly suggest a small refereeing bias towards the home side. However, the expert panel who were arguably less susceptible to influential factors such as the crowd, awarded a similar distribution of decisions, i.e., 44 (35%) should have resulted in home free kicks and 37 (29%) should have been awarded to the away team, and 46 (36%) should have been play-on situations. When referees were inaccurate in their decisions, once more this did not seem to bias one team over another as exactly half of the mistakes were given in favour of the home team (n = 23) and half in favour of the away team (n = 23).

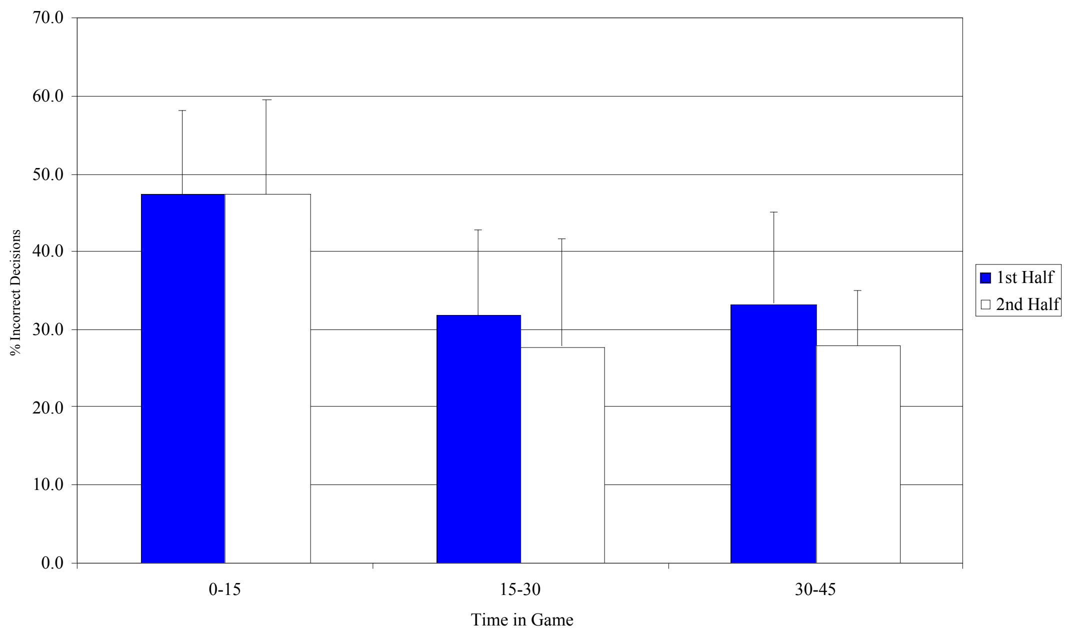

Figure 2. Proportion of incorrect decisions made by match-day referee at different phases of the game (error bars indicate standard deviation amongst referees).

The referees were less accurate in the opening 15 minutes of each half (1st 15 minutes, mean = 51% correct; 2nd 15 minutes, mean = 69%; 3rd 15 minutes, mean = 70%) than they were at any other period (see Figure 2). The small sample size (n = 7) meant that these differences were not statistically significant. However, the addition of only 4 participants (first half) or 5 (second half) following the same trend would have yielded highly significant ($P < .03$) differences.

Of the increased errors in the first 15 minutes, 45% (n = 9) were due to penalising when the referee should not have, 20% (n = 4) were for not awarding a home free kick and 35% (n = 7) were for not awarding a free kick to the away team. Thus, there was no clear pattern of over penalising (45%) or under penalising (55%) in the first 15-minute period of each half.

Speed of Movement. The referees' accuracy did not vary with the speed of their movement (t (125) = 0.08, P = 0.9, Cohen's d =0.02), as the average speed for correct (mean = 6.9 km/hr, s = 4.3) and incorrect decisions (mean = 7.0 km/hr, s = 5.3) was not significantly different.

Heart Rate and Distance Covered. There was no significant difference in the average HR between correct (mean = 165.5 bpm, s = 12.5 bpm) and incorrect decisions (mean = 165.6 bpm, s = 13.3 bpm; t(125) = .058, P = 0.9, Cohen's d = 0.01). Similarly, there were no significant differences between the cumulative distance covered (correct mean = 5493 m, s = 2912, incorrect mean = 4709 m, s = 2915 m) and the quality of their decisions (t(125) = 1.46, p = 0.1, Cohen's d = 0.27).

Difficulty of Decisions. There was a strong relationship between incident difficulty and the correctness of the match referees' decisions (χ^2 (1, 127) = 15.5, P < .0001). For the 33 most difficult incidents (i.e., viewed most often by the panel) match referees' decisions were only 36% correct compared with 75% correct for the remaining 94 clips. There was no statistically significant difference between the most difficult decisions and the others in terms of referees' HR, or locomotion.

Discussion

The aim of this study was to examine the DM and locomotor performance of top football referees. In accordance with other investigations[e.g. 26, 31], football referees covered on average nearly 10.5km during a game, with the majority of the distance covered in the first half. There was no difference in the proportion of high intensity running performed although average heart rates dropped from the first half (mean = 166 bpm, s = 7.4) to the second half (mean = 160 bpm, s = 9.4). Previous research has questioned whether the reduction of physical activity in the second half may be due to referee fatigue or possibly because the tempo of the game decreases as players get tired[e.g. 25, 27, 29, 40]. Taken together the movement and HR scores presented here lend partial support to both interpretations. For example if locomotor activity was limited by referee fatigue in the second half, one might expect higher HR values. However, the referees had higher mean HRs in the first half than the second half. As player work rates were not monitored in the present study, we are unable to confirm that the tempo of the game decreased in the second half. However, the fact that referees heart rates were lower in the second half than in the first half indicates that the referees were conserving energy as they became more fatigued[see also 26]. Further research in which the work-rates of players and officials are measured simultaneously would be necessary to confirm our interpretation of the data.

Castagna and D'Ottavio[29] found that elite Italian referees sprint for 13% of match time, run for 25%, jog for 44%, walk for 9%, and move backwards for 9% of the time. However, the present study found a relatively smaller distribution of sprinting and running activities (2% and 12% respectively) with more jogging and walking/standing (21% and 65% respectively). This discrepancy may be explained partly by the slightly different locomotor categories preferred by Castagna and D'Ottavio (e.g. low-intensity running categorised as < 13km.hr^{-1}, whereas sprinting was > 24km.hr^{-1}) but more likely by the increased total distance covered by the Italian referees (approximately 11.5 km per game). Indeed, the intensity of play in the premier league matches played in Europe and New Zealand may be different which has been shown to influence referee work rate[27].

Examination of referees' DM performance revealed figures in line with previous research, with match-day referees achieving on average 64% accuracy from the incidents selected (Fuller *et al.*[17]: 70% accuracy; Gilis *et al.*[18]: 60% accuracy). Interestingly, further inspection of the data revealed that referees were on average only 51% accurate in the opening 15 minutes of each half, and 70% accurate at all other times. Our analysis suggests that with only a moderately larger sample (e.g., N = 12 referees) these differences would have been statistically significant. This is an intriguing possibility with important practical applications.

Intuitively one might assume that referees begin the game by "laying out their stall," and refereeing strictly to the letter of the law, an approach that is often propounded by referee associations. However, there was no trend towards either over-penalising or under-penalising during the early period of each half. Alternatively, Adams[41] might attribute such a dip in performance, after a period of rest, to some form of warm-up decrement. Anecdotally, the referees' performed no obvious mental warm-up techniques in association with their physical warm-up immediately prior to each half, thus potentially reducing their initial DM performance levels[41]. More likely, these phases of the game present periods of relative instability where teams have yet to settle into established patterns of play, and equally the referee attempts to find appropriate solutions to the game s/he is presented with, by setting boundaries that are synchronous with the game[42]. Regardless of the reasons for poorer performance during the opening period of each half, referees need to ensure that they conduct warm ups that combine physical and mental demands to ensure that they are primed for the challenge that the game is likely to present. It is also worth pointing out that the accuracy data presented are probably underestimates for all the decisions a referee makes in a game as only a selection of incidents were analysed in the present study.

Investigation of the balance of both correct and incorrect decisions (i.e., whether the decisions favoured the home or away team) revealed no bias, despite contrary evidence from previous research[e.g. 20, 21]. The match day referee and the experienced panel (who were not susceptible to player or crowd coercion) gave a similar distribution of decisions in favour of the home and away teams.

Although research into referees and ARs has shown that speed of movement can affect decision accuracy[22, 35] the current study did not replicate these findings. In the case of referees, the discrepancy might arise due to the different levels of football considered (youth vs. senior) and the fact that our referees participated in the whole match versus 20 minute segments[35]. In relation to ARs perhaps the speed-accuracy decrement arose in Oudejans *et al's* study[11] because the ARs were trying to adjudicate on the relative positioning of players (essentially a perceptual judgment), which became difficult when their own speed increased. Referees are responsible for trying to adjudicate if a foul has occurred or not (i.e., a more complex type of decision drawing upon cognitive judgment) so speed of movement appears to affect these sorts of decisions less. The lack of any simple relationships between DM with speed of movement, HR, and cumulative distance found in the current study indicates that none of these variables in isolation can be used to predict whether a correct or incorrect decision is more likely. Instead, a more complex, multivariate relationship between DM and physical performance is likely to underpin performance in naturalistic environments.

As we predicted the match referees were less accurate as the decisions became more difficult, as objectively indicated by the number of repeated viewings required by the expert panel. Given the increasingly common use of television match officials (TMOs) in sports such as cricket, rugby union and rugby league the present study raises some interesting questions applicable in a number of sports. From a DM perspective, it would be of interest to analyse the accuracy of the video official's decision having watched an incident several times relative to the frequency of accurate decisions made from the first viewing (i.e., are gut instincts most often correct?). Also, video officials are relatively impartial to the nuances of the game such as player and crowd reactions. They are required to make relatively passive judgments from a number of different perspectives compared to the active DM processing of the on-pitch referees who may see and hear information that the video official cannot, and can also account for the context of the game [12]. To what extent does this passive presentation of information provide helpful (or conflicting) information from which to inform decisions? Furthermore, does the increasing presence of video officials take some of the control (and players' respect) away from the match-officials? Further research will be necessary to address such issues but for the time-being, it is likely that video officials will remain solely a useful aide to the match official/s for certain types of decisions and sports.

Future research should investigate innovative methods to train referee DM whilst maintaining the naturalistic elements of the task. It is possible that officials could enhance their training by viewing multiple video clips of difficult incidents in the way that the TMO does in certain sports. A similar strategy has been shown to be effective amongst rugby football referees, leading to improved judgment accuracy[see 2]. Unfortunately, within the present study we did not collect any data on the subsequent DM performance of those referees who acted as expert judges to ascertain if they benefited from the experience. It would also be worthwhile investigating other training methods that could improve DM. One factor that has been investigated in other performance domains is the time allowed to make each judgment. There is some evidence that time compression in training certain skills (i.e., speeding up the rate at which events occurs) can enhance subsequent performance. This is known as 'above-real-time-training' and has shown some utility in training pilots and air traffic controllers[43]. Similar strategies are used regularly in police and army training, where recruits are progressively placed in stressful real-life situations to 'inoculate' them to extreme DM demands[44].

Conclusion

This is the first study to investigate both the DM accuracy and physical performance of football referees officiating in competitive matches. A number of findings reported (e.g. distance covered, HR and DM accuracy) support research conducted with top-level football referees in other countries. From the small sample of referees who participated, there appear to be no clear simple relationships between activity levels and DM performance. For example, it was anticipated that as referees fatigue their DM performance might deteriorate. However, there were no significant differences between variables such as the cumulative distance covered or HR, and the quality of decisions. These findings should not be construed as implying that the DM and physical performance of the referee are independent. Instead we argue that these processes are intricately connected but their complex relationship is heavily influenced by a number of other factors (such as situational context), which remain to be studied together. Whilst the referees levels of accuracy may seem low at 64% it should be pointed out that only a selection of incidents from the game were analyzed. In fact, the level of decision accuracy found in the present study was exactly in line with those found in previous studies[17, 18].

Given such findings and the relative ease of GPS data collection, this investigation highlights the value of such technology as a valuable tool for sports scientists to investigate athletic performance. With further developments in GPS miniaturization, this technology may also present another practical solution to measuring referee and player locomotion in their naturalistic environment without compromising their safety or performance.

References

[1] Craven BJ. A psychophysical study of leg-before-wicket judgements in cricket. British Journal of Psychology. 1998;89:555-78.

[2] Mascarenhas DRD, Collins D, Mortimer P, Morris RL. A naturalistic approach to training accurate and coherent decision-making in rugby union officials. The Sport Psychologist. 2005;19(131-147).

[3] Jones MV, Paull GC, Erskine J. The impact of a team's aggressive reputation on the decisions of association football referees. Journal of Sports Sciences. 2002;20:991-1000.

[4] Ericsson KA, Lehmann AC. Expert and exceptional performance: Evidence of maximal adaptation to task constraints. Annual Review of Psychology. 1996;47:273-305.

[5] Williams AM, Davids K, Williams JG. Visual Perception and Action in Sport. London: E & F. N Spon 1999.

[6] Klein GA. An overview of naturalistic decision making applications. In: Zsambok CE, Klein GA, eds. *Naturalistic decision making*. Mahwah, NJ: Laurence Erlbaum Associates (1997):49-60.

[7] Royal KA, Farrow D, Mujika I, Halson SL, Pyne D, Abernethy B. The effect of fatigue on decision making and shooting skill performance in water polo players. Journal of Sports Sciences. 2006;24(8):807-15.

[8] Vickers J, Reeves M-A, Chambers KL, Martell S. Decision training: Cognitive strategies for enhancing motor performance. In: Williams AM, Hodges NJ, eds. *Skill acquisition in sport: Research, theory and practice*. London: Routledge, Taylor and Francis 2004:103-20.

[9] Dicks, M, Button, C Davids, K. Representative task designs for the study of perception and action in sport. International Journal of Sport Psychology, 2008. Special issue on Skill Acquisition and Sport Performance. In press.

[10] Davids K, Button C, Renshaw I, Araújo D, Hristovski R. Movement models from sport provide representative task constraints for studying adaptive behavior in human motor systems. Adaptive Behavior. 2006;14(1):73-95.

[11] Oudejans RRD, Verheijen R, Bakker FC, Gerrits JC, Steinbrückner M, Beek PJ. How position and motion of expert assistant referees in soccer relate to the quality of their offside judgments during actual match play. International Journal of Sport Psychology. 2005;36:3-21.

[12] Mascarenhas DRD, Collins D, Mortimer P. The art of reason versus the exactness of science in elite refereeing: Comments on Plessner and Betsch (2001). Journal of Sport and Exercise Psychology. 2002;24:328-33.

[13] Reilly T, Gregson W. Special populations: The referee and assistant referee. Journal of Sports Sciences. 2006;24:795-801.

[14] Anderson JR. Acquisition of cognitive skill. Psychological Review. 1982;89:396-406.

[15] Mascarenhas DRD, Collins D, Mortimer P. The accuracy, agreement and coherence of decision making in rugby union officials. Journal of Sport Behavior. 2005;28:253-71.

[16] Mascarenhas DRD. The psychology of refereeing: Rugby football league coherence training results. Edinburgh, Scotland: University of Edinburgh, PESL Dept; 2004.

[17] Fuller CW, Junge A, Dvorak J. An assessment of football referees’ decisions in incidents leading to player injuries. American Journal of Sports Medicine. 2004;32(Suppl.1):17S-22S.

[18] Gilis B, Weston M, Helsen WF, Junge A, Dvorak J. Interpretation and application of the laws of the game in football incidents leading to player injuries. International Journal of Sport Psychology. 2006;37(2-3):121-8.

[19] Helsen W, Bultynck J-B. Physical and perceptual-cognitive demands of top-class refereeing in association football. Journal of Sports Sciences. 2004;22:179-89.

[20] Nevill AM, Balmer NJ, Williams M. The influence of crowd noise and experience upon refereeing decisions in football. Psychology of Sport and Exercise. 2002;3:261-72.

[21] Boyko RH, Boyko AR, Boyko MG. Referee bias contributes to home advantage in English Premiership football. Journal of Sports Sciences. 2007;25:1185-94.

[22] Borotikar BS, Newcomer R, Koppes R, McLean SG. Combined effects of fatigue and decision making on female lower limb landing postures: Central and peripheral contributions to ACL injury risk. Clinical Biomechanics. In press.

[23] Mascarenhas DRD, O'Hare D, Plessner H. The psychological and performance demands of association football refereeing. International Journal of Sport Psychology. 2006;37:99-120.

[24] Catterall C, Reilly T, Atkinson G, Coldwells A. Analysis of the work rates and heart rates of association football referees. British Journal of Sports Medicine. 1993;27:193-6.

[25] Mallo J, Navarro E, Garcia-Aranda J-M, Gilis B, Helsen W. Activity profile of top-class association football referees in relation to performance in selected physical tests. Journal of Sports Sciences. 2007;7:805-13.

[26] Krustrup P, Bangsbo J. Physiological demands of top-class soccer refereeing in relation to physical capacity: Effect of intense intermittent exercise. Journal of Sports Sciences. 2001;19:881-91.

[27] Weston M, Castagna C, Impellizzeri FM, Rampinini E, Abt G. Analysis of between-half work rates in English Premier League soccer referees. Journal of Science and Medicine in Sport. 2006;9:256-62.

[28] Weston M, Brewer J. A study of the physiological demands of soccer refereeing. Journal of Sports Sciences. 2002;20:59-60.

[29] Castagna C, D'Ottavio S. Effect of maximal aerobic power on match performance in elite soccer referees. The Journal of Strength and Conditioning Research. 2001;15:420-5.

[30] Harley R, Tozer K, Doust J. An analysis of movement patterns and physiological strain in relation to optimal positioning of association football referees. Journal of Sports Sciences. 1999;17:813.

[31] Asami T, Togari H, Ohashi J. Analysis of movement patterns of referees during soccer matches. In: Reilly T, Less A, Davids K, Murphy WJ, eds. *Science and Football* London: E & F. N. Spon 1988:341-5.

[32] Johnston L, McNaughton L. The physiological requirements of soccer refereeing. Australian Journal of Science and Medicine in Sport. 1994;26:67.

[33] D'Ottavio S, Castagna C. Physiological load imposed on elite soccer referees during actual match play. Journal of Sports Medicine and Physical Fitness. 2001;41:27-32.

[34] Button C, Petersen C. Quantifying the physiological demands of football refereeing with GPS tracking technology. Dunedin, New Zealand: University of Otago; 2005.

[35] Verheijen R, Oudejans RRD, Beek PJ, Bakker FC. Factors affecting decision-making of soccer referees. In: Spinks W, Reilly T, Murphy A, eds. *Fourth World Congress of Science and Football*. Sydney: Routledge: Taylor and Francis 2002:28.

[36] Eissmann HJ, D'Hooghe M. Sports medical examinations. *The 23rd man: Sports medical advice for football referees*. Leipzig: Gersone-Druck 1996:7-19.

[37] Edgecomb SJ, Norton KI. Comparison of global positioning and computer-based tracking systems for measuring player movement distance during Australian Football. Journal of Science and Medicine in Sport. 2006;9:25-32.

[38] Drust B, Reilly T, Cable NT. Physiological responses to laboratory-based soccer-specific intermittent and continuous exercise. Journal of Sports Sciences. 2000;18:885-92.

[39] F.I.F.A. Laws of the Game 2007/2008. 2007 [cited 2008 31 March]; Available from: http://www.fifa.com/mm/document/affederation/federation/laws_of_the_game_0708_10565.pdf

[40] Bangsbo J. Energy demands of competitive soccer. Journal of Sports Sciences. 1994;12:S5 - S12.

[41] Adams JA. The second facet of forgetting: A review of warm-up decrement. Psychological Bulletin. 1961;58:257-73.

[42] Mascarenhas DRD, Collins D, Mortimer P. Elite refereeing performance: Developing a model for sport science support. The Sport Psychologist. 2005;19:364-79.

[43] Miller L, Stanney K, Guckenberger D, Guckenberger E. Above real-time training. Ergonomics in Design. 1997;5(3):21-4.

[44] Gladwell M. Blink: The power of thinking without thinking. London: Little Brown 2005.