European Journal of Sport Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tejs20

Measuring the effectiveness of offensive match-play in professional soccer

Albin Tenga [a], Lars T. Ronglan [a] & Roald Bahr [a]
[a] Norwegian School of Sport Sciences, Oslo, Norway
Published online: 09 Jun 2010.

To cite this article: Albin Tenga, Lars T. Ronglan & Roald Bahr (2010) Measuring the effectiveness of offensive match-play in professional soccer, European Journal of Sport Science, 10:4, 269-277, DOI: 10.1080/17461390903515170

To link to this article: http://dx.doi.org/10.1080/17461390903515170

European Journal of Sport Science, July 2010; 10(4): 269–277

Routledge
Taylor & Francis Group

ORIGINAL ARTICLE

Measuring the effectiveness of offensive match-play in professional soccer

ALBIN TENGA, LARS T. RONGLAN & ROALD BAHR

Norwegian School of Sport Sciences, Oslo, Norway

Abstract
The broader measures of offensive effectiveness, such as scoring opportunities and shots at goal, are commonly used as an alternative to goals scored due to the naturally low probability of scoring (about 1%) in soccer match-play. These measures may enable soccer practitioners to objectively see behind single match results, which are often influenced by chance. The purpose of this study was to examine the relationship between broader measures (scoring opportunities and score box possessions) and the ultimate measure (goals scored) of offensive effectiveness. We analysed data from videotapes of 163 of 182 (90%) matches played in the Norwegian men's professional league during the 2004 season. Multiple logistic regression analyses showed very similar results when comparing the effectiveness of different offensive tactics, regardless of which outcome was used. For example, counterattacks were more effective than elaborate attacks in producing goals (odds ratio OR = 2.07, 95% confidence interval: 1.40 to 3.05), scoring opportunities (OR = 2.30, 95% confidence interval: 1.28 to 4.15), and score box possessions (OR = 2.12, 95% confidence interval: 1.39 to 3.25). The receiver-operating characteristic function statistical procedure was used to examine the association between each of the three measures of offensive effectiveness: scoring opportunities, score box possessions, and goals scored. No significant difference was observed between the area under the curve (AUC) for the broader measures (scoring opportunities and score box possessions) and the ultimate measure (goals scored) of offensive effectiveness. The 95% confidence interval of the AUC for both scoring opportunities (0.74–0.84) and score box possessions (0.68–0.76) includes the AUC for goals scored (0.74). Thus, the results are very similar regardless of which outcome measure for offensive effectiveness is used. This indicates that scoring opportunities and score box possessions (shooting opportunities) can be used as a proxy for goals scored when comparing the effectiveness of different playing tactics in soccer. Compared with goals scored, using scoring opportunities or score box possessions requires smaller match samples for meaningful analyses, and may therefore be more feasible alternatives.

Keywords: *Match performance analysis in soccer, logistic regression, ROC curves, goals scored, scoring opportunities, score box possessions*

Introduction

Goal scoring is the ultimate objective measure of offensive effectiveness in soccer match-play and has been extensively used in match performance research (e.g. Bate, 1988; Grehaigne, 1991; Hughes, 1990; Hughes & Franks, 2005; Pollard & Reep, 1997; Reep & Benjamin, 1968). However, goals provide few data points for an entire match and, consequently, large samples of matches/team possessions are required for meaningful analyses. Indeed, there was a 1% scoring probability based on an average of three goals and 280 team possessions per match played in the Norwegian men's professional league during the 2004 season. Pollard and Reep (1997) reported a similar probability of scoring, as they observed just 47 goals for ~6000 team possessions recorded during international matches in the 1986 World Cup finals. Furthermore, Franks (1988) reported that each team performs approximately 200 team possessions on average in a single match and that successful teams (league champions) normally score an average of two to three goals per match, giving a scoring probability of about 1%.

Thus, such a low scoring probability precludes study designs that include an adequate sample size of randomly selected team possessions leading to goals.

Correspondence: A. Tenga, Norwegian School of Sport Sciences, Ullevaal Stadion, Sognsveien 220, PO Box 4014, N-0806, Oslo, Norway. E-mail: albin.tenga@online.no

ISSN 1746-1391 print/ISSN 1536-7290 online
DOI: 10.1080/17461390903515170

The broader measures of offensive effectiveness such as scoring opportunities (e.g. Olsen & Larsen, 1997), shots at goal (e.g. Harris & Reilly, 1988; Hughes & Snook, 2006; Pollard, 1986; Reep & Benjamin, 1968), and entry into the final third of the pitch (e.g. Bate, 1988) have been proposed as alternatives. In addition, compared with the ultimate measure of "goals scored", these broader measures have the potential to provide additional process information linked to their outcome. Moreover, the use of such broader measures may enable soccer practitioners to see objectively behind single match results, which are subject to chance. Reep and Benjamin (1968) demonstrated the existence of random chance, meaning that despite an excess of shots by one team in any single match, the opposing team can still score more goals and thus win the match. However, they also showed that, in the long run, the team producing the most shots tends to score more goals with a goal-to-shot ratio of approximately 1:10. This implies that the actions and outcomes in soccer matches can be described on the basis of probability. However, whether broader measures are able to reflect goals scored over a series of matches has not been examined.

Thus, the main aim of this study was to examine the relationship between broader measures (scoring opportunities and score box possessions) and the ultimate measure (goals scored) of offensive effectiveness.

Methods

Materials

Videotapes of 163 of 182 (90%) matches played in the Norwegian men's professional league during the 2004 season were analysed. The league involved 14 teams and followed a double round robin competition format, which means that each team played 26 matches, 13 at home and 13 away.

First, a sample ("cases") of all goals scored by counterattacks and elaborate team possessions (see Table I for definitions) was identified, including 203 (43%) of the total 476 goals scored.

Second, to obtain an equal number of team possessions for each team, playing home and away, we selected the same number of team possessions from each available match (20 per match were needed according to our power calculations, see below). To account for potential differences in playing tactics and their effectiveness between different stages of the match (first and second half, early or late in the half), the start time of the match period being analysed was selected at random. This was done by assigning to each match a computer-generated random decimal number between 0 and 1, which was multiplied by 86 to indicate the beginning

Table I. Descriptions of variables and definitions of category used in the team match performance analysis

Variables and categories

1. Team possession type

Def. Degree of offensive directness by levels of utilization or creation of imbalance in the opponent's defence to achieve penetration (i.e. how quick penetration is attempted after winning the ball). Penetration is achieved when a pass goes towards the opponent's goal past opponent player(s) while maintaining a high degree of control over the ball. High degree of control over the ball means enough space and time that makes it easier to perform intended actions on the ball.

A. Counterattack ("direct play"): starts by winning the ball in open play and progresses by either (a) utilizing or attempting to utilize a degree of imbalance from start to the end, or (b) creating or attempting to create a degree of imbalance from start to the end by using early (i.e. first or second, evaluated qualitatively) penetrative pass or dribble. Utilizing degree of imbalance means seeking penetration in such a way that a defending team fails to regain a high degree of balance from start to the end of team possession. Counterattacks progress relatively quickly.

B. Elaborate attack ("possession play"): starts by winning the ball in open play and progresses either (a) without utilizing or attempting to utilize a degree of imbalance, or (b) by creating or attempting to create a degree of imbalance by using late (third or later, evaluated qualitatively) penetrative pass or dribble. Not utilizing a degree of imbalance means seeking penetration in such a way that a defending team manages to regain a high degree of balance before the end of team possession. Elaborate attacks often progress relatively slowly.

2. Starting zone

Def. Area across the playing field in which team possession starts (Figure 1).

A. First third: that third of the playing field estimated from own goal line to middle third.

B. Middle third: that third of the playing field estimated from end of the first third to final third.

C. Final third: that third of the playing field estimated from end of the middle third to opponent's goal line, excluding score box.

D. Score box: Prime scoring area in front of the opponent's goal defined as an imaginary prolongation of the penalty area from 16 m to 30 m estimated distance from opponent's goal line.

3. Pass number

Def. Series of passes between players of the attacking team.

A. Short possession: one or two passes per team possession.

B. Medium possession: three or four passes per team possession.

C. Long possession: five or more passes per team possession.

4. Pass penetration

Def. Penetrative passes, i.e. passes towards the opponent's goal past opponent player(s) while maintaining control over the ball and otherwise for non-penetrative passes.

A. Penetrative: only penetrative pass.

B. Mixed: combination of penetrative and non-penetrative passes.

C. Non-penetrative: only non-penetrative pass.

5. Team possession outcome

Def. Degree of offensive success by dichotomy and discrete levels of effectiveness.

I. Score box: Levels of offensive effectiveness within the score box (Figure 1).

A. Goal scoring: scoring attempt ending with a goal approved by a referee.

B. Scoring opportunity: scoring attempt with relatively high scoring probability (e.g. from shorter distances, from wider angles, with poor goalkeeper positioning) as well as with near-scoring situations such as corner kick direct on crossbar.

Table I (*Continued*)

Variables and categories
C. Score box possession: entry into score box with high degree of control over the ball or set play given to the attacking team as a result of entry into score box. High degree of control over the ball means enough space and time that makes it easier to perform intended action on the ball.
II. Not score box: Levels of offensive effectiveness outside the score box (Figure 1).
D. No score box possession: entry into score box with low degree of control over the ball. Low degree of control over the ball means not enough space and time that makes it more difficult to perform intended action on the ball.
E. Final third: ending up in the final third of the playing field.
F. Middle third: ending up in the middle third of the playing field.
G. First third: ending up in the first third of the playing field.

(in minutes) of a match period from which a total of 20 team possessions would be extracted. This was based on the assumption that 20 team possessions last 6.5 min on average, and that there is 2–3 min of extra time added at the end of each match.

Of 3260 team possessions, only 1688 possessions that began by winning the ball in open play (excluding 1572 team possessions initiated by set-plays) were selected. These 1688 team possessions were then analysed for team possession outcome as follows: scoring opportunity, score box possession, not score box possession, final third, middle third, and first third (Figure 1). Finally, team possessions leading to scoring opportunities ($n = 80$, 4.7%), to score box possessions ($n = 167$, 9.9%), and to other team possession outcomes ($n = 1441$, 85.4%) were collected. Therefore, the final sample could be analysed in three ways: team possessions leading to goals (203 cases vs. 1688 random controls), team possessions leading to scoring opportunities (80 cases vs. the remaining 1608 random controls), and team possessions leading to score box possessions (167 cases vs. the remaining 1521 random controls).

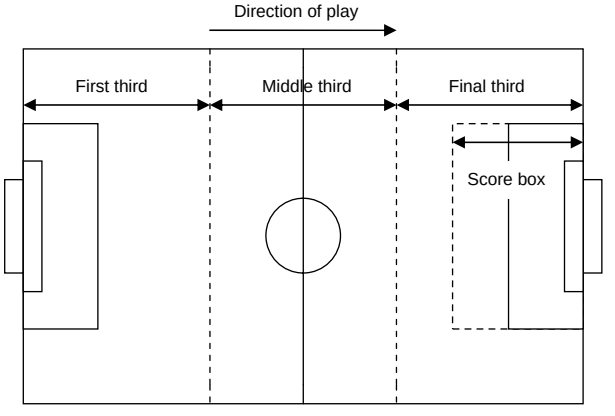

Figure 1. Zones of the playing field include the first third, middle third, final third, and score box.

Team match performance analysis

A team possession was used as the basic unit of analysis and was defined according to Pollard and Reep (1997, p. 542):

> A team possession starts when a player gains possession of the ball by any means other than from a player of the same team. The player must have enough control over the ball to be able to have a deliberate influence on its subsequent direction. The team possession may continue with a series of passes between players of the same team but ends immediately when one of the following events occurs: a) the ball goes out of play; b) the ball touches a player of the opposing team (e.g. by means of a tackle, an intercepted pass or a shot being saved). A momentary touch that does not significantly change the direction of the ball is excluded; c) an infringement of the rules takes place (e.g. a player is offside or a foul is committed).

The following five multidimensional ordered categorical variables were used in this study: team possession outcome (one dependent variable); team possession type, starting zone, pass number, and pass penetration (four independent variables) (Table I). The dependent variable, team possession outcome, had six primary values: scoring or not scoring, scoring opportunity or not scoring opportunity, and score box possession or not score box possession.

Video analysis

The video footage, in DigiBeta video format, was reviewed using a computer-controlled Sony DigiBeta video machine. All goals and 20 consecutive team possessions were extracted from each of the 163 matches using a G4 Mac machine and the program FinalCut Pro version 9.0. A total of 476 goals and 3260 random team possessions in Mac format were stored, and then converted from Mac format to WMV PC format to enable further analysis using Windows Media Player. A soccer coach/researcher (A.T.) experienced in match performance analysis and a soccer coach/master student (D.K.) each analysed about half of the goals and control team possessions based on five variables used in the team match performance analysis. Previously, these two analysts recorded reliability within acceptable limits for all variables used, with the kappa values from an inter-observer test for three variables showing "very good" reliability (0.81–1.0) and for two variables "good" reliability (0.61–0.80) according to Altman (1991). The video analysis data were registered

directly in SPSS (version 15.0, SPSS Inc., Chicago, IL). The study received ethics approval from the Norwegian Social Science Data Services (NSD).

Statistical analysis

We estimated that a difference ($\triangle$) of 27% between the proportions of counterattacks and elaborate attacks culminating in a goal might be detected for a sample size of 609 (203 goals and twice as many control team possessions, i.e. 406) under various initial conditions. This was calculated based on the sample size formula $n = 2(\sigma/\triangle)^2 f_{(\alpha,\ \beta)}$, assuming an alpha of 0.05, a beta of 0.10, and a standard deviation (σ) of 2.22. The standard deviation of 2.22 was estimated using the equation $\sigma^2 = p(1-p)(Q_e^{-1} + Q_c^{-1}) + p_e(1 - p_e)Q_e^{-1} + p_c(1 - p_c)Q_c^{-1}$, where $p = Q_e p_e + Q_c p_c$ (Lachin, 1981), for two independent samples with different sizes. p_e and p_c are the proportions of counterattacks and elaborate attacks in the two groups and Q_e and Q_c are sample fractions in the two groups. Twice as many control team possessions (406) was assumed to be adequate because adding four controls, for example, instead of two controls per case will only marginally increase the efficacy. This means that a change of split from 50% versus 50% to 36.5% versus 63.5% (i.e. $\triangle$ = 27%) is then possible to detect with the same power. Thus, a total of 203 goals and 1688 controls from either counterattacks or elaborate attacks begun by winning the ball in open play were estimated to be sufficient. To obtain 1688 controls, a total of 3260 (20×163) team possessions were collected and further analysed into counterattacks and elaborate attacks. Since the probability of scoring from team possessions initiated by winning the ball in play (0.9%) was much lower than that of producing scoring opportunities (4.7%) and score box possessions (9.8%), the addition of extra cases from either scoring opportunities or score box possessions will only marginally increase power and therefore a difference of 27% ($\triangle$) will still be possible to detect with this power. The alternative hypothesis was tested against the null hypothesis, that there would be no difference in association between scoring opportunities, score box possessions, and goal scored, using a chi-square analysis to determine whether there was a difference in associations between playing tactics and the probability of producing each of the three team possession outcomes.

This hypothesis was tested further by multiple logistic regression analysis in which the dependent variable was whether a goal was scored or not, whether a scoring opportunity was created or not, and whether a score box possession was generated or not. As independent variables, we used four offensive playing tactics, each of which had two categories. These included counterattack versus elaborate attack, final third versus first third, long possession versus short possession, and penetrative pass versus non-penetrative pass. We used a multivariate regression model in which all four independent variables were entered and tested in a single step. In this way, we were able to investigate the relationship between each independent variable on the probability of producing a goal, scoring opportunity or a score box possession, adjusted for the other independent variables. From these models, odds ratio and their respective 95% confidence limits were calculated.

To examine the association between scoring opportunities, score box possessions, and goals scored, a receiver-operating characteristic (ROC) curve statistical procedure was used. The area under the curve (AUC) in ROC analysis reflects the agreement between the offensive tactics and each of the three types of team possession outcomes. First, the predicted probabilities for scoring a goal, making a scoring opportunity, and creating a score box possession were obtained from a multiple logistic regression model using the four categorical variables as independent variables (see null hypothesis test explained above). These three scores (i.e. predicted probabilities) were saved in separate columns as new variables. Each of these new variables was then selected as the test variable and a corresponding team possession outcome as the state variable, with 1 as its positive value. The area under the ROC curve (AUC) for each of the three team possession outcomes was then computed. It is important to note that the ROC area is made by graphing sensitivity versus 1 minus specificity when varying the split point in the distribution of estimated probabilities. However, when endpoints have different prevalence as is the case here, comparisons of ROC areas across endpoints may be more difficult to perform. Statistical significance was set at $P < 0.05$ in all tests.

Results

Descriptive analysis

Each of the 14 teams involved contributed an average of 15 goals (range 8–21), 6 scoring opportunities (range 3–10), and 12 score box possessions (range 9–14) to the final sample. The same differences were observed between proportions of control team possessions produced by offensive playing tactics in all variables.

Similar significant differences between percentages of attempts for each of the three possession outcomes were observed in all variables except team possession type. For the variable team possession

Table II. Number of goals ($n=203$) and controls ($n=1688$) and percentage of goals scored by playing tactics according to offensive variables

Variable	*N* (%)	Goals	Controls	Goal%	*P**
Team possession type					
Counterattack	792 (41.9)	106	686	13.4	0.002
Elaborate attack	1099 (58.1)	97	1002	8.8	
Starting zone					
Final third	55 (2.9)	18	37	32.7	<0.001
Middle third	860 (45.5)	101	759	11.7	
First third	976 (51.6)	84	892	8.6	
Pass number					
Short possession	884 (47.9)	67	817	7.6	<0.001
Medium possession	572 (31.0)	68	504	11.9	
Long possession	388 (21.0)	56	332	14.4	
Pass penetration					
Penetrative pass	173 (9.4)	44	129	25.4	<0.001
Mixed	1043 (56.6)	138	905	13.2	
Non-penetrative pass	626 (34.0)	9	617	1.4	

*Pearson chi-square.

type, only the percentage of attempts on goal (13.4%) produced by counterattacks was significantly higher than that produced by elaborate attacks (8.8%). The proportion of control team possessions produced by elaborate attacks for goals (1002/1688 = 59.4%), scoring opportunities (961/1608 = 59.8%), and score box possessions (904/1521 = 59.4%) was higher than that produced by counterattacks for all three team possession outcomes (686/1688 = 40.6%, 647/1608 = 40.2%, and 617/1521 = 40.6%, respectively) (Tables II, III and IV).

Logistic regression analyses

There were significant differences between the offensive tactics in the odds ratio for producing each of the three measures of offensive effectiveness for all variables except the variable starting zone for score box possession. For example, compared with elaborate attacks, counterattacks registered a higher odds ratio for scoring a goal (OR = 2.07, 95% confidence interval: 1.40 to 3.05; $P<0.001$), creating a scoring opportunity (OR = 2.30, 95% confidence interval: 1.28 to 4.15; $P=0.005$), and making a score box possession (OR = 2.12, 95% confidence interval: 1.39 to 3.25; $P=0.001$) (Table V).

ROC curves procedure

No significant difference was observed between AUC for the broader measures (scoring opportunities and

Table III. Number of scoring opportunities ($n=80$) and controls ($n=1608$) and percentage of scoring opportunities produced by playing tactics according to offensive variables

Variable	*N* (%)	Scoring opportunities	Controls	Scoring opportunity%	*P**
Team possession type					
Counterattack	686 (40.6)	39	647	5.7	0.13
Elaborate attack	1002 (59.4)	41	961	4.1	
Starting zone					
Final third	37 (2.2)	8	29	21.6	<0.001
Middle third	759 (45.0)	47	712	6.2	
First third	892 (52.8)	25	867	2.8	
Pass number					
Short possession	817 (49.4)	25	792	3.1	0.003
Medium possession	504 (30.5)	28	476	5.6	
Long possession	332 (20.1)	25	307	7.5	
Pass penetration					
Penetrative pass	129 (7.8)	20	109	15.5	<0.001
Mixed	905 (54.8)	52	853	5.7	
Non-penetrative pass	617 (37.4)	6	611	1.0	

*Pearson chi-square.

Table IV. Number of score box possessions ($n = 167$) and controls ($n = 1521$) and percentage of scoring box possessions produced by playing tactics according to offensive variables

Variable	*N* (%)	Score box possessions	Controls	Score box possession%	*P*⋆
Team possession type					
Counterattack	686 (40.6)	69	617	10.1	0.85
Elaborate attack	1002 (59.4)	98	904	9.8	
Starting zone					
Final third	37 (2.2)	5	32	13.5	0.007
Middle third	759 (45.0)	93	666	12.3	
First third	892 (52.8)	69	823	7.7	
Pass number					
Short possession	817 (49.4)	47	770	5.8	<0.001
Medium possession	504 (30.5)	53	451	10.5	
Long possession	332 (20.1)	62	270	18.7	
Pass penetration					
Penetrative pass	129 (7.8)	18	111	14.0	<0.001
Mixed	905 (54.8)	123	782	13.6	
Non-penetrative pass	617 (37.4)	20	597	3.2	

⋆Pearson chi-square.

score box possessions) and the ultimate measure (goals scored) of offensive effectiveness. The 95% confidence interval of the AUC for both scoring opportunities (0.74–0.84) and score box possessions (0.68–0.76) includes the AUC for goals scored (0.74) (Table VI). The plot of the multiple curves shows goals scored overlapped both scoring opportunities and score box possessions (Figure 2).

Discussion

The main outcome of this study was that the results are very similar regardless of which outcome measure for offensive effectiveness is used. However, the failure to identify a difference does not necessarily prove equivalence. Nevertheless, we would argue that our results (see Table V) show such a high degree of similarity that shooting opportunities and score box possessions can be used as a proxy for goals scored when comparing the effectiveness of different playing tactics in soccer. Thus, compared with goals scored, using these broader measures requires smaller match samples for meaningful analyses and may therefore be more feasible. In addition, these intermediate outcome measures are capable of generating process variables that may provide useful additional information.

The strength of this study was that we used a case-control design with cases selected both randomly

Table V. Odds ratio (OR) and 95% confidence interval (95% CI) from multivariate analysis for goals scored ($n = 203$), scoring opportunities ($n = 80$), and score box possessions ($n = 167$) by the two opposite playing tactics according to offensive variables

	Goals scored		Scoring opportunities		Score box possessions	
Variable	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Team possession type						
Counterattack vs. elaborate attack[a]	2.07 (1.40–3.05) 1	<0.001⋆	2.30 (1.28–4.15) 1	0.005⋆	2.12 (1.39–3.25) 1	0.001⋆
Starting zone						
Final third vs. first third[a]	5.76 (2.56–12.93) 1	<0.001⋆	11.57 (3.82–35.03) 1	<0.001⋆	2.34 (0.75–7.25) 1	0.14
Pass number						
Long possession vs. short possession[a]	2.63 (1.52–4.56) 1	0.001⋆	6.56 (2.58–16.67) 1	<0.001⋆	4.60 (2.52–8.39) 1	<0.001⋆
Pass penetration						
Penetrative pass vs. non-penetrative pass[a]	20.70 (9.72–44.09) 1	<0.001⋆	22.18 (8.22–59.87) 1	<0.001⋆	4.72 (2.39–9.34) 1	<0.001⋆

Note: The odds ratio (OR) reflects the chance of producing a goal, scoring opportunity or score box possession compared with the reference category[a]. In contrast with goals, scoring opportunities and score box possessions were selected randomly.
⋆Offensive tactics included in the model.

Table VI. Area under the ROC curve with 95% confidence interval (95% CI) showing the association between goals scored, scoring opportunities, and score box possessions

Team possession outcome	Area	95% CI	*P**
Goals scored	0.74	0.71–0.78	<0.001
Scoring opportunities	0.79	0.74–0.84	<0.001
Score box possessions	0.72	0.68–0.76	<0.001

Note: Scoring opportunities and score box possessions were selected randomly.
*Pearson chi-square.

(scoring opportunities and score box possessions) and not randomly (goals) and a large sample of randomly selected control team possessions. This increases the validity and ability to generalize the findings. Moreover, in contrast to most previous studies, we used logistic regression analysis, the appropriate statistical method for comparisons of categorical differences associated with binary response variables (Nevill, Atkinson, Hughes, & Cooper, 2002). The use of multidimensional categorical data enabled us to analyse factors related to match performance that usually are difficult to measure directly, as well as their interdependency.

That there were very similar associations between each of these three measures of offensive effectiveness was not surprising. This is because scoring opportunities and score box possessions are both characterized by factors closely linked with goals scored. *Scoring opportunities* are used to achieve a more refined measure of scoring attempts than simply counting the number of shots. A scoring opportunity is based on the multidimensional qualitative evaluation of the environment around a scoring attempt. This involves analyses of other essential features, well recognized in practice, associated with improving chances for a shot to produce a goal. These include factors such as shooting distance, shooting angle, goalkeeper's position, opponent's defensive pressure, and shooting technique (Franks, 1988; Hughes, 1990; Olsen, Larsen, & Semb, 1994). Olsen and Larsen (1997) reported an average of 3.4 scoring opportunities per goal for the Norwegian national team in the 1993 season. *Score box possessions* are thought to be strongly correlated with shooting opportunities. Most goals are reported to be scored from the area that extends beyond the penalty box (Hughes, 1990; Worthington, 1980). Shots originating from within this prime scoring area (score box) are known to have higher quality in terms of accuracy on target and scoring potential than those from outside this area (Bate, 1988; Dufour, 1993). For example, according to Dufour (1993), shots from a distance of 30 m or more (i.e. outside score box) have a scoring efficiency of nearly 0%, while shots from within 16.5 m and 5.5 m (i.e. inside score box) have a scoring efficiency of 10% and 15%, respectively. Thus, possessions with a high degree of ball control inside the score box may lead to a goal being scored in up to 15% of cases.

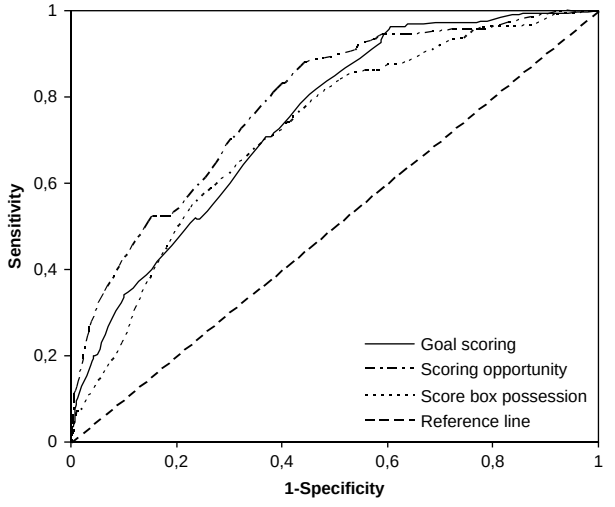

Figure 2. Multiple ROC curves for visual comparison of association between scoring opportunities, score box possessions, and goals scored.

Furthermore, it is natural that the probability of scoring a goal by team possessions begun by winning the ball in play (0.9%) is much smaller than that of creating both scoring opportunities (4.7%) and score box possessions (9.8%). Reep and Benjamin (1968) noted that as a team possession advances towards a goal, the defending opponents progressively converge into a more compact formation and hence become closer to the ball and therefore improve their chances of preventing penetration. Therefore, it is logical to consider goals scored, scoring opportunities, and score box possessions (shooting opportunities) as ordered categories of the variable team possession outcome. Based on the same understanding, Franks (1988) identified four similar sub-objectives for team possessions (or four priorities for loss of possession): goals, shots, shooting opportunities, and reaching the attacking third of the field.

Our finding that the offensive tactics of counterattack, final third starting zone, long possessions (five passes or more), and penetrative passes are more effective in producing goals, scoring opportunities and score box possessions (shooting opportunities) than their respective opposite tactics of elaborate attacks, first third starting zone, short possessions (two passes or less), and non-penetrative passes corresponds well with the findings of previous studies. Olsen and Larsen (1997) showed that breakdowns (counterattacks) produce more goals and scoring opportunities than longer attacks (elaborate attacks). That team possessions originating from the final third rather than first third are considerably more effective for goal scoring was also reported by

Bate (1988), Hughes (1990), and Hughes and Snook (2006). That there was no difference in effectiveness in producing score box possessions between team possessions started in the final third and first third in the present study was not expected. That penetrative passes were found to be more effective than non-penetrative passes lends support to the superior principle of play governing the game of soccer, namely penetration versus preventing penetration (Franks, 1988; Olsen et al., 1994).

Previous studies have either supported (e.g. Hughes & Churchill, 2004; Hughes & Franks, 2005; Hughes, Robertson, & Nicholson, 1988; Hughes & Snook, 2006) or opposed (e.g. Bate, 1988; Hughes, 1990; Reep & Benjamin, 1968) our finding that long possessions (five passes or more) are more effective than short possessions (two passes or less). However, Reep and Benjamin (1968) demonstrated that more goals were scored from shorter passing sequences (three passes or less) than longer ones, and Hughes and Franks (2005) reported that there were more instances of shorter passing sequences than longer ones. Thus, consistent with our results, longer passing sequences were considered to be more effective than shorter ones.

It is also apparent that treating *short possessions* and *long possessions* interchangeably with *direct play* and *possession play* as done by Hughes and Franks (2005) may be inappropriate. This is because simply counting the number of passes excludes other essential features in the analysis of these styles of attack (Franks, 1988; Olsen et al., 1994). In the current study, the use of multidimensional qualitative evaluation enabled us to define counterattack ("direct play") and elaborate attack ("possession play") as degrees of offensive directness judged by how quick the attempt(s) is to utilize or create imbalance in the opponent's defence. For the purposes of this paper, the terms "elaborate attacks" and "counterattacks" are considered to be synonymous with "possession play" and "direct play", respectively. Hence, the current results may suggest that a combination of quick utilization of imbalances in the opponent's defence (i.e. counterattack) and the use of five passes or more (i.e. long possession) is effective in producing goals, scoring opportunities, and score box possessions. Indeed, Hughes and Franks (2005) reported that longer passing sequences produced more goals per possession than shorter passing sequences for successful teams, probably because successful teams possess the skill to sustain long passing sequences when exploiting imbalances in the opponent's defence.

It is important to realize that goal scoring is and will always be the ultimate objective measure of offensive effectiveness in soccer match-play. The use of broader measures is recommended only in cases of obvious disadvantages caused by naturally low scoring probability.

The current findings may have some practical implications for match performance researchers when designing studies on playing effectiveness. The use of scoring opportunities or score box possessions (shooting opportunities) as a measure of offensive effectiveness instead of goals scored should encourage researchers to employ more appropriate study designs that include an adequate sample size of randomly selected events. In addition, these broader measures may provide additional process information useful to soccer practitioners.

In conclusion, our results show that scoring opportunities and score box possessions (shooting opportunities) can be used as a proxy for goals scored when comparing the effectiveness of different playing tactics in soccer. Compared with goals scored, using scoring opportunities or score box possessions requires smaller match samples for meaningful analyses, and may therefore be more feasible.

References

Altman, D. G. (1991). Some common problems in medical research. In D. G. Altman (Ed.), *Practical statistics for medical research* (pp. 403–409). London: Chapman & Hall.

Bate, R. (1988). Football chance: Tactics and strategy. In T. Reilly, A. Lees, K. Davids, & W.J. Murphy (Eds.), *Science and football* (pp. 293–301). London: E & FN Spon.

Dufour, W. (1993). Computer-assisted scouting in soccer. In T. Reilly, J. Clarys, & A. Stibbe (Eds.), *Science and football II* (pp. 160–166). London: E & FN Spon.

Franks, I. (1988). Analysis of association football. *Soccer Journal*, *33*(5), 35–43.

Grehaigne, J. F. (1991). A new method of goal analysis. *Science and Football*, *5*, 10–16.

Harris, S., & Reilly, T. (1988). Space, team work and attacking success in soccer. In T. Reilly, A. Lees, K. Davids, & W. J. Murphy (Eds.), *Science and football* (pp. 322–328). London: E & FN Spon.

Hughes, C. (1990). *The winning formula*. London: William Collins.

Hughes, M., & Churchill, S. (2004). Attacking profiles of successful and unsuccessful teams in Copa America 2001. *Journal of Sports Sciences*, *22*, 505.

Hughes, M., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of Sports Sciences*, *23*, 509–514.

Hughes, M., Robertson, K., & Nicholson, A. (1988). Comparison of patterns of play of successful and unsuccessful teams in the 1986 World Cup for soccer. In T. Reilly, A. Lees, K. Davids, & W. J. Murphy (Eds.), *Science and football* (pp. 363–367). London: E & FN Spon.

Hughes, M., & Snook, N. (2006). Effectiveness of attacking play in the 2004 European Championships. In: H. Dancs, M. Hughes, & P. G. O'Donoghue (Eds.), *Proceedings of the World Congress of Performance Analysis of Sport 7* (pp. 46–62). Cardiff: CPA UWIC Press.

Lachin, J. M. (1981). Introduction to sample size determination and power analysis for clinical trials. *Controlled Clinical Trials*, *2*, 93–113.

Nevill, A. M., Atkinson, G., Hughes, M. D., & Cooper, S.-M. (2002). Statistical methods for analysing discrete and categorical data recorded in performance analysis. *Journal of Sports Sciences*, *20*, 829–844.

Olsen, E., & Larsen, O. (1997). Use of match analysis by coaches. In T. Reilly, J. Bangsbo, & M. Hughes (Eds.), *Science and football III* (pp. 209–220). London: E & FN Spon.

Olsen, E., Larsen, O., & Semb, N. J. (1994). *Effektiv fotball*. Norway: Gyldendal Norsk Forlag A/S.

Pollard, R. (1986). Soccer performance analysis and its application to shots at goal. *Research Bi-annual for Movement*, *4*, 19–27.

Pollard, R., & Reep, C. (1997). Measuring the effectiveness of playing strategies at soccer. *The Statistician*, *46*, 541–550.

Reep, C., & Benjamin, B. (1968). Skill and chance in association football. *Journal of the Royal Statistical Society A*, *131*, 581–585.

Worthington, E. (1980). *Teaching soccer skills*. London: Henry Kimpton (Publishers) Ltd.