Match Analysis procedures in Team Sports: a study of relationships among Technical, Tactical, Physical Parameters and final outcomes in elite soccer matches as analyzed by a semiautomatic video tracking system.

Bruno RUSCELLO, PhD

Università di Roma “Tor Vergata”
Corsi di laurea in Scienze Motorie
Dipartimento di Ingegneria Industriale
Match Analysis

Match Analysis is a major subject among coaches, team managers and sport scientists and it is gaining an increasing relevance day by day. **Information** and the relevant **data processing** are the key factors while referring to this aspect of the sport training methodology.

Currently the concept of “**Match Analysis**” is used in several countries to define the process of observing and evaluating a “whole of behaviours” performed by the players during a match, applying different methodologies and using specific instruments and tools, in order to:

1. **collect and process** the relevant **data** concerning the different features of games or athletic disciplines, under different points of view;
2. **provide relevant presentations**, appropriately formatted, in order to show the collected and processed data in an **accessible way** to all the concerned people, at different levels (i.e. coaches, players, sport scientists, officials, managers, journalists, etc.);
3. **provide an interpretation** of the collected and processed data, in order to define better some specific feature of the investigated performance (i.e. the physiological side of the performance or the biomechanics or the tactical features of a match or a game) with the ultimate aim of improving these aspects through the appropriate administration of the relevant training processes.
Match Analysis

- Match Analysis refers to the objective recording and examination of behavioural events occurring during competition. It may be focused on the activity of one player, or may include the integration of actions and movements of players around the ball.

- Match Analysis may range in sophistication from discrete data about the activity of an individual player, or of each member of the team as an individual profile, to a synthesis of the interplay between individuals in conformity to a team plan.

C. Carling, M. Williams, T. Reilly – Handbook of Soccer Match Analysis
Routledge – Taylor & Francis Group
London and New York
Match Analysis, in situation sports, is a branch of the Sport Pedagogy and the Sport Sciences. Several disciplines, at different levels and extensions, combine to bring descriptions, classifications, eventually explanations and also to provide possible predictions (probabilistic approach) about some of the most significant situations that could be marked during sport events or matches.

B. Ruscello, Doctoral Thesis - 2008
MATCH ANALYSIS DOMAINS

Physiological - Conditioning
(Match Analysis - 1° level)
(Physical Trainer – Conditioner)

Biomechanics - Techniques
(Match Analysis - 2° level)
(Trainer – Instructor)

Strategical – Tactical – Mental
(Match Analysis - 3° level)
(Coach)

A. Sacripanti (2007)
A Study of Relationships among Technical, Tactical, Physical Parameters and Final Outcomes in Elite Soccer Matches as Analyzed by a Semiautomatic Video Tracking System

Cristoforo Filetti\textsuperscript{1,2,3}, Bruno Ruscello\textsuperscript{1,4,5,6}, Stefano D'Ottavio\textsuperscript{1,4,7}, and Vico Fanelli\textsuperscript{8}

Keywords
match analysis, video tracking, passing skill, likelihood, efficiency
The performance of a soccer team depends on many factors such as decision-making, cognitive and physical skills, and dynamic ever-changing space-time interactions between teammate and opponents in relation to the ball.
Seventy (n = 70) matches of the Italian SERIE A season 2013-2014 were investigated to analyze the mean performance of 360 players in terms of physical (physical efficiency index; PEI) and technical-tactical (technical efficiency index; TEI) standpoints.
Using a semiautomatic video analysis system that has incorporated new parameters able to measure technical-tactical and physical efficiency (Patent IB20 I 0/002593, 20 I 1-ISA), the correlation between these new variables and how much it relates to the likelihood of winning were verified.
Correlations between TEI and PEI were significant (n = 140, r = .60, p < .001), and TEI showed a higher likelihood of winning than PEI factors (p < .0001 vs .0001, CI 95% [1.64, 3.00] vs. [1.28, 2.07]).
Higher TEI and TEI + PEI differences between the teams were associated with a greater likelihood of winning, but PEI differences were not.
Key performance indicators and this performance assessment method might be useful to better understand what determines winning and to assist the overall training process and match management.
Going deeper...

What was our research question?
Introduction

Soccer is one of the most investigated sports in the world, with its scientific analysis growing continuously. Contemporary match analysis procedures (Barros et al., 2007; Glazier, 2010; Gregson, Drust, Atkinson, & Di Salvo, 2010) have provided insiders a great quantity of information.

Thanks to low-cost, high-technology standards, this information is available to an always increasing audience of stakeholders (performance analysts, coaches, team managers, and exercise physiologist) with the ultimate aim to continuously enhance soccer performance.
Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science

Robert Rein* and Daniel Memmert

DOI 10.1186/s40064-015-1462-0

Review of the tactical evaluation tools for youth players, assessing the tactics in team sports: football

Sixto González-Villora¹, Jaime Serra-Olivares², Juan Carlos Pastor-Vicedo³ and Israel Teoldo da Costa⁴
Aims

The aim of the present study was to investigate the relationship between some technical-tactical and physical parameters and final game outcomes in elite soccer matches.

In these contests, soccer was considered as a non linear dynamical system that required a novel mathematical analysis (such as that provided by Prozone®) to investigate the relevance of the main key performance indicators (KPIs) that were chosen for the game's final outcome.

Our hypothesis, according to previous studies (Carling & Dupont, 2011; Diaz del Campo, Gonzalez Villora, Garcia Lopez, & Mitchell, 2011; Hoppe, Slomka, Baumgart, Weber, & Freiwald, 2015), was that there is a different contribution of these KPIs on matches outcome, such that the technical-tactical domain is of primary importance.
The novelty in our approach moves the analysis from the classic scouting approach to observing technical-tactical abilities merged with physical intensity carried out descriptively (e.g., by counting the number of passes, crosses, shots, and interceptions or in a deeper analysis the direction of the passes or the efficiency of the shots [n of successful shots/n attempted shots]) to a mathematical modeling, based on probability theory, that might estimate the whole team's efficiency in terms of right decision-making and, subsequently, the right technical performance, in the context of dynamic interactions.
How did we study the problem?
Method

Experimental Approach to the Problem

In this observational study, we tried to answer self-posed research questions regarding the influence of different domains of performance - the *technical tactical domain and the physical domain* - on the final outcome of a soccer match.

To achieve this, we introduced a *new performance analysis procedure*, based on probability theory, unveiling an innovative technology that allowed us to investigate the inner structure of the game. We set as *independent variables* some KPIs provided by the manufacturer of the software (passes; ball speed; ball trajectory; distance between the holder players, the teammates, and the opponents) and studied their impact on the final outcome (*dependent variable*) of the match.
These parameters, suggested and patented by K-Sport®, are based on mathematic intelligence applied to the technical-tactical performance in soccer.
One of the first approaches in this field was produced by Grehaigne, Bouthier, and David (1997), still far from the actual mathematic intelligence but able to merge three main variables:

- **space and time**;
- **information**, as the ability to manage all the changing parameters to enhance the certainty of the teammates and the uncertainty of the opponents;
- **organization**, intended as the continuous exchange of individual to collective tasks.
More simply, these authors suggested a **microstate** defined by a distribution of the players on the pitch's playing area according to their **positions**, **orientations**, and **speeds**.

Mathematically, considering the area of play in attack which is composed of **free space** + **interacting space** + **surfaces covered by the defenders**, they determined the available area for the attackers (**movement** + **available space**) to obtain the **likelihood of passes**.
These and others parameters have been processed deeply and the technical tactical indicators were calculated applying the following formulas:

**Formula 1-(K-analysis - C-Pass [Passing Difficulty Coefficient])**

\[
C_n = a' \cdot f'(A_n) + b' \cdot g'(T_n^{solo}) + c' \cdot h'(D_n^{pass}) + d' \cdot l(D_n^{T_{Tr-a}} , X_1) \cdot e' \cdot m'(D_n^{T_{Tr-C}} , X_2)
\]

This formula addresses the concept of the **different levels of difficulty in passing**, and it considers several interacting variables that should be taken into account in each changing decisional process performed by the players.
This formula addresses the concept of the optimal passing (KP) considering this skill as the product of manifolds interacting variables that should be taken into account in each changing decisional process performed by the players.
Figure 1. (a) K-regions software elaboration: Video animation made by the software. (b) The Voronoi theorem in an instant of a game, where each zone is assigned to each player according to the time and space and their derivate parameters.
The Voronoi Theorem applied to soccer
Figure 2. C-pass software elaboration: (a) video animation shows the distance (D) of the pass. The midfielder has to pass the ball to the player on the left side. If we consider only this parameter, the C-pass I is equal to the %: (D/Dmax, represented by the magnitude of the field dimension) x 100. In the maximized figure (b), we add the distance between the opponents and the ball trajectory. The software calculates automatically all the distances between the opponents and the ball trajectory, and it selects the two players closer to the passing lane from the right and left side. So considering the right distance (rd) and the left distance (ld) in respect to the ball trajectory, the second part of the C-pass is calculated as: 100 – {[(rd + ld)/D max] x 100}. Merging the formulas, C - pass = (C - pass I + C - pass2)/ 2.
Summarizing, the K-analysis takes into account **three key aspects**:

- **Playing choices analysis**: KS indicates optimal solution (mathematically, the system assesses how the variables change according to the player performed choices; i.e., the choice is positively assessed if it allows to create events where the teammates are efficient), OffAgg (offensive aggressiveness) indicates if the player tends to attack, DefAgg indicates if the player tends to attack opponents in possession (mathematically, the system assesses whether the distance between the player and the opponent increases or decreases in the sample time considered as unit).
Passing analysis and ball possession: KP (this KPI represents a subset of the above KS; mathematically, the system assesses how the variables change according to the player performed pass, i.e., the passes are positively assessed if they allow to create events where the teammates are efficient), KP prec indicates passing accuracy, CP indicates difficult passing efficiency, Control indicates ball control, off 1vs1 indicates offensive one against one, and def 1vs1 indicates defensive one against one (Figure 3).
Figure 3. **1 vs 1 evaluation**: definition - the software recognizes 1 vs 1 when A (attacker) is in possession of the ball against one opponent defender (D); Condition 1 - inside the circle with radius $r$, there are only the players A and D; Condition 2 - D is located within the angle which has vertex in the ball and bisector passing through the centre of the goal. A is effective (last field in the figure) if keeps the ball possession and overcomes D, keeping the zone field that gives advantage for the team. D is effective if recovers the ball or channels the opponent outside the *dangerous zone*. 

Without ball (running) movements analysis: KM is the optimal movement without ball, KM prec is the movement accuracy, and CM is the difficult movements efficiency (these three KPIs represent a subset of the above KS; mathematically, the system assesses how the variables change according to the player performed movements; i.e., the movement are assessed positively if they allow to create high likely success KS or reduce the CP); Pressing indicates pressing efficiency (it is evaluated the ability to stop or slow down the opponents actions; mathematically, the system assesses the events where the player performs a high Def Agg index and it assesses them positively if the player in possession loses the ball control or he is forced to make a backside or sideways pass).
The aforementioned three points establish the **technical efficiency index (TEI)**.

These concepts have been related to the **physical efficiency index (PEI)**, defined as follows: considering that the number of duels in the elite football showed a decreasing between the first and the last 5' (Carling & Dupont, 2011), the above-mentioned patent tried to analyse deeper each parameter involved.
Often, during the matches, **two players have to challenge each other** to achieve the ball before the direct opponent: both perform **acceleration**, achieving high speed and keeping them for some time (sprint). When they achieve the ball, both have performed a high acceleration, a sprint, and then distance covered for the previous physical KPI. **Thus, the performance is similar, but the duel has always a winner: one of the players is more effective than the other.**

This concept can be detected by the tracking and the K-filter elaboration (above mentioned patent, 2011). The mathematic intelligence applied on the system is even able to save the best time performance in duel (recorded as 100%) and the worst one (recorded as 0%): all the values in this interval are normalized on 0-100% scale.
We divided the study into two types of analysis: one considered the absolute value of the data to investigate how the measured parameters might affect the game result, and the second considered the differences between the absolute values of the two teams for each game. This latter allowed us to study how the difference between the parameters of the two teams might influence either's winning likelihood.
Participants

We analysed 70 games of the Italian SERIE A during the season 2013-2014, considering the final outcome of the matches as the difference in goals scored at the end of each match (1.07 ± 0.98); Gòmez, DelaSerna, Lupo, & Sampaio, 2014, 2016; Gomez, Gasperi, & Lupo, 2016). All the computations were made taking into account the performances of all players (n=442, age 27.7±6.2, height 1.81 ±0.16m, weight 78.7±5.19kg) involved in those games.
**Procedures**

All the video data were collected by K-Sport Tracking Semi-Automatic System & AMISCO Pro Tracking System (Prozone Sport, Leeds, UK) and processed by K-SportOnline (K-Sport, Montelabbate, PU, Italy). K-analysis system (based on patent IB20 I 0/002593) was used to analyze the X, Y raw data in order to perform the report. The system has been based on three different cameras (Toshiba Camileo H30, full HD) positioned and fixed at least at 12 meters high at the central tribune of the stadiums for all 70 games analyzed. Each camera provided a complete footage of one third of the field, and the total size of the field was rebuilt via Prozone software managed by a laptop (required Windows 10, ram 8 mega, hard disk 500 GB, intel core i7), also managing all the tracking procedures. As Prozone is a semiautomatic system, the final tracking is always supervised by a trained human operator.
Analysis

Pearson's coefficient with range (-1, 1) was used to investigate the correlation between TEI and PEI. To study the effect of TEI/PEI on the match results, we used a logistic regression model, belonging to the general linear model class, with a binomial random variable (win-no win). To evaluate the TEI and PEI's significance, we considered the p value of likelihood ratio test at 5% significance level, and set confidence intervals at 95% (95% CI).

Finally, independent t test was used to find differences between means of TEI/PEI in winners and not winners. Significance was set as p < .05. Effect size as Cohen's d was also calculated, considering a large range (d > .8). Before applying the t test, we verified assumptions of normality of samples with the test of Shapiro and homoscedasticity, at 5% significant level. All statistical analyses were performed using R 3.0.0.
What did we find?
Results
As a first step of the study, we performed the correlation between TEI and PEI (Figure 4).

Figure 4. Correlation between TEI and PEI within the analyzed sample (n = 70 games) with Pearson's coefficient $R$ on the left corner.
We performed individually the variable TEI, in relation to the likelihood of winning: with each increased unit of TEI, the relationship between the likelihood of winning and likelihood of not winning increased by about 115%. Consequently, we considered only the variable PEI, and it too was statistically significant (about 60% success with increasing units). When TEI and PEI were considered together, because of the correlation between the two parameters, the PEI lost relevance, and the TEI was the only significant variable. In particular, when combining TEI and PEI, the relationship between the likelihood of winning and the likelihood of losing increased by about 117%, that is, the likelihood of winning increased. Of this increase, however, a 111% increase was attributed only (per unit) to TEI, while only 3% was attributed to increasing units of PEI.
In the second part of the study, we considered the difference between the parameters of the two teams (i.e., $\text{TEI} = \text{TEI Team A} - \text{TEI Team B}$), proceeding with the same variables analysis adopted above.

Table 1 shows the maximum likelihood estimate and 95% CI that better clarifies the importance of our specific analyses compared to the universal standard.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>OR</th>
<th>p value*</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{TEI}^a$</td>
<td>2.15</td>
<td>.001</td>
<td>[1.64, 3.00]</td>
</tr>
<tr>
<td>$\text{PEI}^b$</td>
<td>1.60</td>
<td>.001</td>
<td>[1.28, 2.07]</td>
</tr>
<tr>
<td>$\text{TEI}^c$</td>
<td>2.11</td>
<td>.001</td>
<td>[1.52, 3.12]</td>
</tr>
<tr>
<td>$\text{PEI}^d$</td>
<td>1.03</td>
<td>.848</td>
<td>[0.75, 1.41]</td>
</tr>
<tr>
<td>$\text{TEI (diff)}^e$</td>
<td>2.25</td>
<td>.001</td>
<td>[1.75, 3.13]</td>
</tr>
<tr>
<td>$\text{PEI (diff)}^f$</td>
<td>1.92</td>
<td>.001</td>
<td>[1.50, 2.60]</td>
</tr>
<tr>
<td>$\text{TEI (diff)}^g$</td>
<td>2.06</td>
<td>.001</td>
<td>[1.55, 2.93]</td>
</tr>
<tr>
<td>$\text{PEI (diff)}^h$</td>
<td>1.21</td>
<td>.249</td>
<td>[0.88, 1.72]</td>
</tr>
</tbody>
</table>

Note. $\text{OR} =$ odds ratio; $\text{CI} =$ confidence interval; $\text{TEI} =$ technical efficiency index; $\text{PEI} =$ physical efficiency index.

$^a$Explanatory variable of one-term logistic regression (TEI team’s value).

$^b$Explanatory variable of one-term logistic regression (PEI team’s value).

$^c$Explanatory variable of two-term logistic regression (TEI team’s value).

$^d$Explanatory variable of two-term logistic regression (PEI team’s value).

$^e$Explanatory variable of one-term logistic regression (TEI team’s difference).

$^f$Explanatory variable of one-term logistic regression (PEI team’s difference).

$^g$Explanatory variable of two-term logistic regression (TEI team’s difference).

$^h$Explanatory variable of two-term logistic regression (PEI team’s difference).

$p < .05$. 
Considering the **TEI individually**, the likelihood of winning increased **125%**. If we considered only **PEI**, it increased around **92%**. But if we considered both TEI and PEI together, the increase in the ratio between the likelihood of winning and the not winning likelihood they increased almost **150%**. Deeply, the increase due to **TEI considered individually is about 106% against PEI of 21%**. Because of the correlation between the two parameters, **PEI lost practical relevance, even though its influence on the likelihood of winning was not negligible.**
Next, we analyzed TEI and PEI differences in relation to the score between the teams within each game. Table 2 shows these correlations.

As last analysis, we used a t test to compare the averages of TEI and PEI's differences between winning and non winning teams (Table 3).
What do this findings mean?
DISCUSSION

The main aim of the present study was to investigate the relationship of a TEI and a PEI on soccer game results through a K-Sport-Prozone device.

The significant correlation (r = .60, p < .001, n = 140) between TEI and PEI (Figure 4) confirms the existing relationship between the physical and technical-tactical aspects in the game.

In the literature of the last decade, many studies (Dellal et al., 2010; Diaz del Campo et al., 2011; Hoppe et al., 2015) tried to identify the real link between these two categories, but the findings were often divergent because of the several variables involved in the game performance.
To better investigate the influence of the TEI and the PEI (set as independent variables) on the game result (dependent variable), we performed a binary logistic regression (Figure 5), considering TEI and PEI individually and both together. The black solid curve (TEI + PEI) has a similar trend to the dotted black one (TEI) with the likelihood ratio of winning and non-winning ranging from 115% to 117%.
Figure 5. Win likelihood curves considering the two parameters individually and together within analyzed sample (n = 70 games).

The hatch black curve (PEI) grows slower (92%), showing slower influence of the PEI on the likelihood of winning.
Going deeper in our analysis, we performed the second regression analysis trying to understand if the TEI and PEI score difference (i.e., TEI Team A - TEI Team B) between the two teams in the single game might relate to the final result. The solid black curve (TEI Team A - TEI Team B + PEI Team A - PEI Team B) confirms the result highlighted in Figure 5.

Considering both the parameters together, the likelihood of winning increases 25% more with respect to considering only the TEI and 60% more with respect to the PEI.
Furthermore, it is very interesting to note whether the TEI and PEI difference is $\sim 10$, when the likelihood of winning carne to almost 100%.

Finally, in the case of two similar level teams, the likelihood of winning came to around its theoretical value of 33% (Figure 6).
Conclusion

The overall aim of this study was to determine whether the new parameters (TEI and PEI) that are part of the patent recorded by Marcolini and K-Sport (2011) might relate to the score of soccer games when considered both singularly or in combination.

All the technical parameters involved in TEI seem to have a greater relevance on the final result than the PEI.

Our findings also suggest that, considering the TEI difference score within the single game, a team that achieves a higher TEI compared to the opponent's has a higher likelihood of winning the game.
Although the PEI index seems to play a similar role as TEI, its predictive power on the final result seems to be of a lesser relevance.

Indeed, we also found teams able to win the game although they achieved a low PEI score.

Although the PEI index seems to play a similar role as TEI, its predictive power on the final result seems to be of a lesser relevance. Indeed, we also found teams able to win the game although they achieved a low PEI score. All these procedures are aimed at evaluating the individual performance of each player, crucial to improve the team efficiency as a whole: these data may help coaches plan possible interventions with the ultimate aim of improving the performance of players, both singularly and as a team.
As a practical application, the subparameters involved in TEI (Formulas 1 and 2) provide deep information about the performance of each player according to each specific position in a soccer play system.

As shown in Figure 7, within the macroparameter TEI, the technical staff might monitor within a reasonably long period of time, each single subparameter that could be improved in the individual player so as to have a marginal gain on the team too.

In this example, in which the priority of the overall training process is the technical-tactical aspect of the game as planned by the coach, deeper data they suggested that one of the two midfielders were lacking in OffAgg, although the team showed a high PEI %.
Figure 7. Observed improvements over 2 years in different parameters of a single player's performance (note as the OffAgg improved as a consequence of a specific training process).

*Note.* KS = optimal solution; KP = optimal pass; KP prec = passing precision; KM = without-ball optimal movement; Km prec = movement precision; Control = efficiency in the ball control; Pressing = pressing efficiency; OffAgg = offensive aggressiveness; DefAgg = defensive aggressiveness; Speed = speed events efficiency; Ace = acceleration events efficiency; Dee = deceleration events efficiency; Power = power events efficiency.
The specific training process was aimed at improving the ability of attacking the space in both players. Thus, this assessment process was exploited in a dual mode: to improve individual ability and for whole team play, during ball possession.

Our study suggests that time spent stimulating the players in their decision-making processes during training should be greater than that usually spent toward physical gains. It would seem that the decision-making process and technical abilities are the main goals to pursue in elite soccer.
Thank you for your attention

bruno.ruscello@uniroma2.it