*Journal of Sports Sciences*, 2002, **20**, 845–852

# Temporal pattern analysis and its applicability in sport: an explanation and exemplar data

ANDREW BORRIE,[1]* GUDBERG K. JONSSON[2,3] and MAGNUS S. MAGNUSSON[3]

[1]*Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK*
[2]*University of Aberdeen, Aberdeen, UK and* [3]*University of Iceland, Reykjavik, Iceland*

Accepted 17 April 2002

Quantitative analysis of sports performance has been shown to produce information that coaches can use within the coaching process to enhance performance. Traditional methods for quantifying sport performances are limited in their capacity to describe the complex interactions of events that occur within a performance over time. In this paper, we outline a new approach to the analysis of time-based event records and real-time behaviour records on sport performance known as T-pattern detection. The relevant elements of the T-pattern detection process are explained and exemplar data from the analysis of 13 soccer matches are presented to highlight the potential of this form of analysis. The results from soccer suggest that it is possible to identify new profiles for both individuals and teams based on the analysis of temporal behavioural patterns detected within the performances.

*Keywords*: match analysis, pattern, performance, temporal.

## Introduction

Within the coaching process, great emphasis is placed on the coach's ability to observe and recall all the critical discrete incidents from a sport performance. However, it has been shown that coaches cannot accurately observe and recall all of the detailed information that is required for a complete understanding or interpretation of performance (Franks and Miller, 1986). The coaching process is, therefore, enhanced by the provision of additional information that describes sport performance in detail beyond that which coaches can provide through recall of personal observations. Detailed quantitative analyses can enhance performance through the improvement of performer feedback if the feedback is provided in an appropriate form (Franks, 1997). Additional quantitative information about a performance can also impact on the coaching process by enhancing coaches' interpretation of performance.

Traditional analysis methods have used frequency of event occurrence as their index of performance. For example, analysts have recorded the number of passes made from particular playing zones or how many times a team or individual makes an unforced error. However, if one accepts the argument that sport performance consists of a complex series of interrelationships between a wide variety of performance variables, then simple frequency data cannot necessarily capture the full complexity of a performance. The challenge for the performance analyst is to find additional analysis methods that can generate alternative representations of performance.

One aspect of performance that has not been adequately studied is the temporal structure and interrelationships between events within sport performances. Our aims here are to introduce, and explain, a data analysis method that examines temporal structure and to show how it can be applied to sport performance. This analysis technique can identify consistent temporal patterns that exist within a sport performance, thus providing a different picture of the complex interrelationships between discrete events within that performance.

Traditional frequency analyses of performance have provided, and continue to provide, valuable information that coaches and performers use to enhance the coaching process. It is not our assertion that an analysis of temporal structure is better than other approaches,

* Address all correspondence to Andrew Borrie, Research Institute for Sport and Exercise Science, Liverpool John Moores University, 15–21 Webster Street, Liverpool L3 2EF, UK. e-mail: a.g.borrie@livjm.ac.uk

merely that analysis of temporal structure provides an additional fresh perspective for performance analysts to consider and use.

## T-pattern detection and analysis

The new analysis approach presented here is based on a very general type of time pattern called a 'T-pattern'. The corresponding detection algorithm for a T-pattern allows detection of repeated temporal and sequential structures in real-time behaviour records. The pattern type and algorithm were developed and tested extensively, outside of sport, based on the assumption that complex streams of human behaviour, such as sport performance, have a sequential structure that cannot be fully detected through unaided observation or with the help of standard statistical and behaviour analysis methods. Given that observational records of human behaviour, including sport performance, have a hidden temporal-sequential structure, an analysis tool that can discover and describe this structure will enhance understanding of the behaviour being studied. A generic observational software package called Theme has been developed specifically to operationalize T-pattern detection as an analysis process (Magnusson, 1996, 2000).

A schematic representation of a T-pattern is shown in Fig. 1. If one assumes that the letters in line 1 correspond to specific performance events (e.g. pass, tackle and shot in soccer) that appear on the line in proportion to the time of their occurrence, then line 1 is a visual representation of the temporal structure of a sport performance. Within the upper line, a sequence of four event-types – A, B, C and D – recurs, but masked by the occurrence of two other event-types, W and K. An event-type is noted using upper-case and its instances using lower-case – that is, an instance of A is noted a. If a performance analyst or coach had been visually inspecting the data string, it is unlikely that the pattern would have been detected. In many cases, frequency counts, lag-sequential analysis or time-series analysis would not identify complex patterns of this nature in real behaviour records. However, a T-pattern analysis would identify the pattern because of its consistent temporal structure. The T-pattern detection algorithms allow an analyst to discover repeated temporal patterns even when various other event-types occur in between the elements of the pattern.

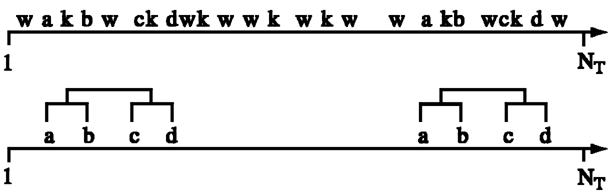


**Fig. 1.** The most regular T-patterns even in extremely simple data may be hard to spot. The T-pattern on the lower axis is present in the upper axis where a few occurrences of w and k make it hard to see. The defining characteristics of T-patterns are apparent: fixed order of components occurring with similar distances between them at each pattern occurrence. The binary tree structure indicates the bottom-up detection strategy, which may reflect inherent hierarchical structure.

### *T-pattern definition and detection*

A T-pattern is essentially a combination of events in which the events occur in the same order with the real-time differences between consecutive pattern components remaining relatively invariant (i.e. the time difference between A and B will be $x \pm y$) with respect to an expectation assuming, as a null hypothesis, that each component is independently and randomly distributed over time. As stated by Magnusson (2000), 'that is, if A is an earlier and B a later component of the same recurring T-pattern then after an occurrence of A at $t$, there is an interval $[t + d1,\ t + d2]$ $(d2 \geq d1 \geq 0)$ that tends to contain at least one occurrence of B more often than would be expected by chance' (p. 94). The temporal relationship between A and B is defined as a *critical interval* and this concept lies at the centre of the pattern detection algorithms (Magnusson, 1996, 2000).

The pattern detection algorithms can be used to analyse both ordinal and temporal data. However, for the algorithms to generate the most meaningful analyses, the raw data must be time-coded – that is, an event must be coded according to time of occurrence as well as event-type. The pattern in Fig. 1 illustrates how a larger pattern ((AB)(CD)) is detected as a combination of the two simpler patterns (AB) and (CD). Even in moderate data sets, the number of potential T-patterns is very high. When, for example, the potential number of event-types is 100, the number of potential event patterns involving up to 10 event-types is many orders of magnitude greater than $100^{10}$ if all possible time windows are also considered. Even for supercomputers, it becomes an impossible task to search for each possible temporal pattern separately. To deal with this problem, simple patterns are detected first – that is, identifying relationships between two event-types such as the (AB) relationship in Fig. 1 – while more complex patterns are detected as patterns of patterns (a so-called 'bottom-up' search strategy). The simpler patterns (AB) and (CD) are detected first and then the larger pattern ((AB)(CD)) as a combination of these. The new larger pattern may then become a part of even more complex patterns as it combines with other simple or complex patterns.

A further stage of the detection process deals with completeness competition between all the detected patterns. In this stage, those patterns that are less complete versions of one or more alternate patterns are deleted. That is, during the detection process, a pattern Q = (ABCDE) may be partially detected as, for example, (ACDE) or (BDE) or (ABCE); since elements of Q are missing, these three patterns constitute less complete descriptions of the underlying patterning. A newly detected pattern $Qx$ is thus considered equally or less complete than an already detected pattern $Qy$ if $Qx$ and $Qy$ occur equally often and all events in $Qx$ also occur in $Qy$. In this case, $Qx$ is eliminated. This completeness competition ensures that only the most complete patterns survive and constitute the result of the detection process.

The only 'meaning' of any higher (binary non-terminal) node in a pattern is a critical interval relationship between the occurrences series of its left (preceding) and right (following or concurrent) branch. The detection algorithm continues until no more critically related pairs of event-types or patterns can be found to form larger patterns. The number and composition of the patterns is thus decided by the data. This kind of pattern is more a matter of repeated relative timing and may implicate high transition probabilities (i.e. from just above zero to 100%). If one takes an easy linguistic example to demonstrate the point, the high frequency of the standard phrase (pattern) 'thank you' would be, *a priori*, highly unlikely and would form a very significant T-pattern, assuming, as a null hypothesis, that each word is distributed independently with its observed frequency. However, the forward transition probability of 'you' given the occurrence of 'thank', as well as the backward probability of a preceding 'thank' given the occurrence of 'you', would still both be very low indeed.

### *Issues, advantages and characteristics of T-pattern analysis*

*Intervening behaviours: accounting for their influence.* The number and type of behaviours that may occur between the components of a T-pattern can vary greatly between occurrences of the same pattern. For this reason, methods that depend only on the order of events – disregarding the temporal distances between them – have great difficulty detecting such patterns. For example, in a soccer match, an attack down the right side of the pitch by team A may be temporally related to the same team conceding a corner a short time later. The right wing back is pulled out of position in the first attack and the defence then becomes vulnerable to a counter-attack, resulting in the concession of a corner. In between making the initial attack and conceding a corner, team A may have lost possession in many ways. However, the attack and the concession of the corner were causally related. The variability in loss of possession in the middle of the pattern would mean that the pattern is only detectable by its temporal characteristics.

Current data collection approaches, combined with analysis methods in major statistical packages, were not developed for the detection of T-patterns and generally cannot identify them. The same is true for specialized behaviour research software, such as The Observer (Noldus, 1991; Noldus *et al.*, 1999) or GSEQ (Bakeman and Quera, 1995), which are based on such methods. Analysis of the characteristics of standard statistical methods, which make them inadequate for T-pattern detection, has been described elsewhere (Magnusson, 2000).

*Acyclical patterns.* Although T-patterns can occur cyclically, this is not one of their defining characteristics. The temporal distances between occurrences of patterns, not events, may just as well be irregular – for example, a pattern of play may occur three times within the first 20 min of a soccer match and then not re-occur until the last 5 min. The within-pattern intervals between events will have remained relatively invariant but overall pattern occurrence will have been irregular – that is, detection does not rely upon cyclical occurrence.

*Causality.* Causality is also an issue that may relate to any identified pattern. The analysis process as it stands does not infer that events within a pattern are causally related. The detection algorithm is neutral to this issue, which concerns the interpretation of detected patterns. If one considers a simple tennis example, this issue is easily highlighted. If a tennis player bounces the ball before each serve and, when the ball is served wide to the opponent's backhand, the return is played down the line, then this is likely to constitute a temporal pattern. However, within this pattern are both non-causal and potentially causal relationships. The act of bouncing of the ball will not have been causally related to the direction of the serve, but the direction of the serve may have been causally related to the direction of the return – that is, the opponent can only play shots down the line when the ball is served wide. One cannot assume that just because a pattern exists the elements within it are causally related.

## Exemplar data

### *Methods*

A team sport, soccer, has been analysed with the intention of identifying whether T-pattern detection has relevance as an analytical method within performance analysis. The research used multiple game analysis with each game being treated as a single case. Thirteen

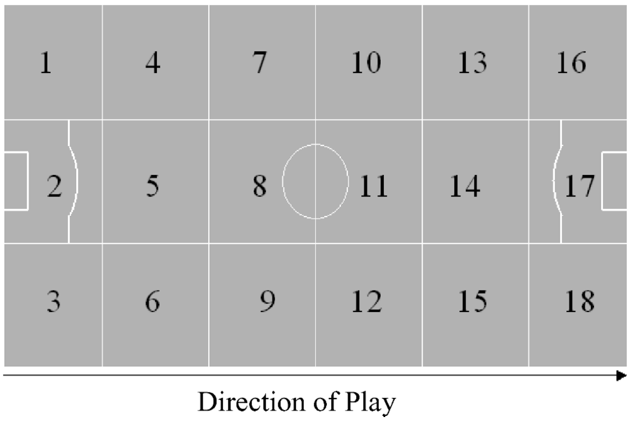


**Fig. 2.** A schematic representation of the zones identified for analysis of ball movement in soccer.

soccer matches, four club and nine international matches, were coded using a combination of the soccer match analysis system developed at Liverpool John Moores University and ThemeCoder, enabling detailed coding of digitized video files (25 Hz). Coding included information on pitch position, player and match events. Pitch position was classified according to the pitch divisions shown in Fig. 2. The primary event categories for data collection were pass, tackle, header, run, dribble, clearance, shot, cross, set-play, lost control and foul. Additional qualifying statements could be tagged to each event category. All data were analysed using the Theme software package.

## Results and discussion

The coding of soccer play here only concerns player behaviour and position most directly related to the ball. Information about the positions and behaviours of all players during the match would be a welcome addition to allow deeper insights, for example, through the detection of possibly inter-individual patterns of positioning, which might, moreover, connect to the patterns already detected. The results show that repeated play patterns can be detected even without such additional information.

The results show that many temporal patterns exist in soccer. The number, frequency and complexity of the detected patterns indicates that sport behaviour is more synchronized than the human eye can detect. This synchrony was found to exist on different levels, with highly complex time structures that extended over considerable time spans within performances, with patterns occurring both cyclically and acyclically.

The results taken from soccer show three discrete examples that identify within-team patterns (e.g. Fig. 3, ball movement) and interactive patterns involving both teams (Fig. 4, goal scoring; Fig. 5, play preceding critical free kicks).

A typical within-team pattern of events from the soccer analysis is shown in Fig. 3. This figure displays a detected T-pattern that occurred three times during the first half of a European Championship qualifying match in 1998. The three boxes in Fig. 3a show the same observation period. The upper-left box shows the hierarchical construction of the pattern. The tree structure identifies the simple patterns on its right-hand edge and, as the tree builds towards the left edge of the box, shows how the simple patterns are linked together to form the more complex pattern. The upper-right box displays the time of each event-type in the pattern and their pattern connection based on the critical interval relationship between their occurrence series. The bottom box shows the pattern, as a hierarchical structure, expressed in relation to the observation period – that is, when it occurred during the match; only complete patterns are shown in this box.

The pattern describes how player A moves the ball towards the opponent's goal by receiving the ball in, and then passing it out of, pitch zones 8, 11 and 14 consecutively. Player A then completes the sequence by passing it to player B who receives it in zone 15. The pattern describes an attacking movement through the middle of the pitch, which opponents would clearly wish to prevent. Traditional frequency analyses of passing would have identified the ball reception and subsequent pass from each zone as discrete events, but are far less likely to have linked the consecutive actions in the four zones. The movement from zone 11 to 14 also occurred on another five occasions during the first half (Fig. 3a, upper right box), further suggesting that player A was working effectively through the central channel of the pitch. This integrated analysis would potentially enhance the information given to the coach.

Figure 4 also shows a T-pattern from a club match (English Premiership) in which the pattern involves players from both teams and relates to the critical incident of shots on goal. The pattern that is shown occurred on three occasions during the second half of the match and includes two shots on goal within each pattern. The total time covering the three patterns includes six shots by team A, which represent 75% of their total shots on target during the second half. Even more significantly, two of the three pattern occurrences resulted in goals. From the defending team's perspective, identification of the antecedents to pattern occurrence will allow appropriate defensive strategies to be created.

The results in Fig. 5 show the capacity of the analysis process to identify longer and more complex patterns that involve extensive interactions between opposing teams. In Fig. 5, a consistent temporal pattern precedes

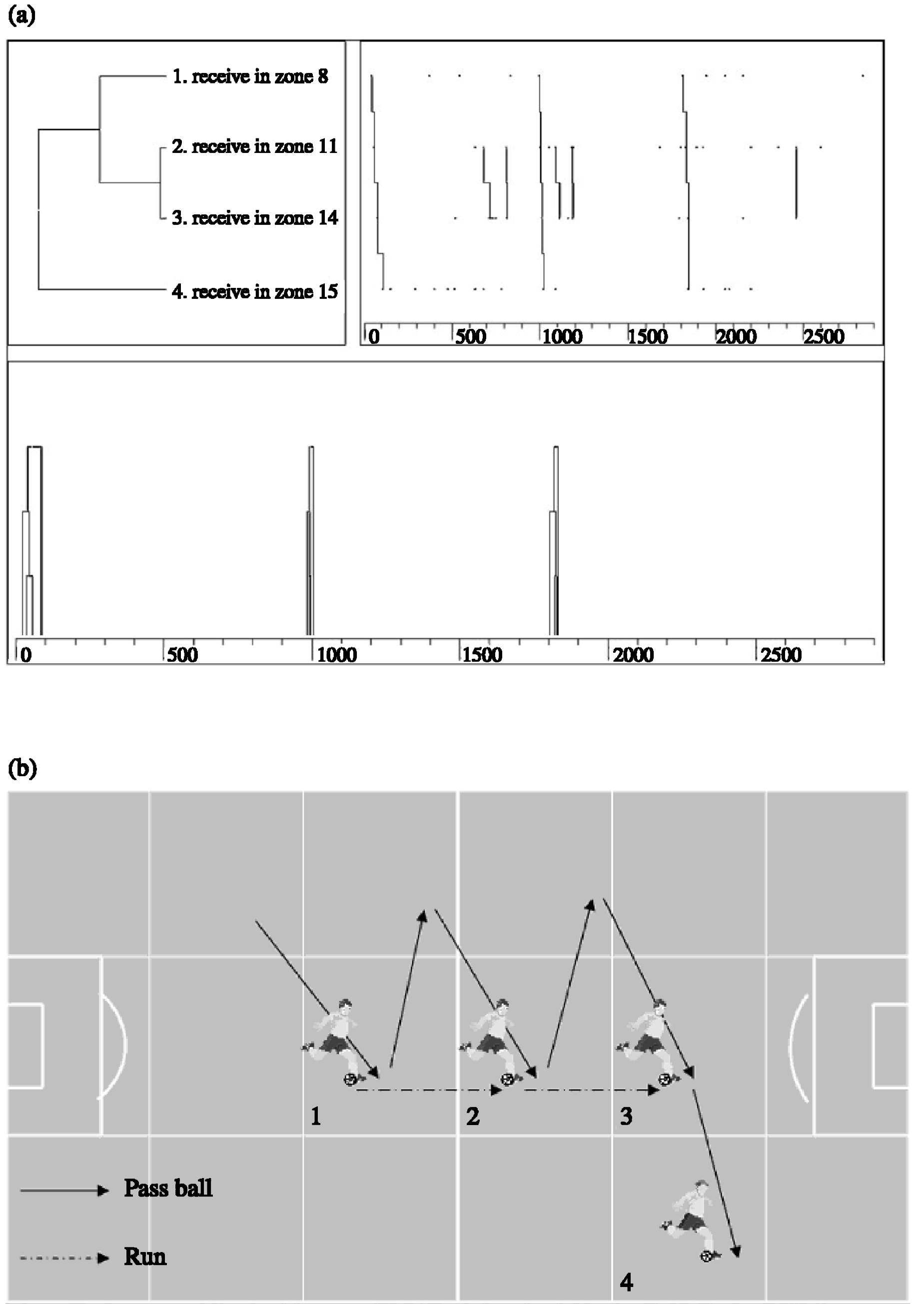


**Fig. 3.** A temporal pattern of attacking movement of the ball through the centre of the pitch. (a) Data output from Theme analysis software showing temporal and hierarchical representation of a T-pattern. The three boxes in this figure involve the same observation period and the T-pattern relates to ball movement in the centre of the pitch within a soccer match. (b) Schematic representation of the same data. (1) Player A receives the ball in zone 8, passes the ball to a team-mate and runs forward. (2) Player A receives the ball in zone 11, passes the ball to a team-mate and runs forward. (3) Player A receives the ball in zone 14, passes the ball to a team-mate in zone 15.

both goals scored in a European club match. The pattern covers an extended period within which the two teams exchange possession on two occasions. The duration of the periods between the events forming the pattern was such that other match events will have occurred between pattern events. The duration and nature of the pattern are such that one could question whether, given the extended time period of the pattern and the changes of possession, the pattern events were causally related. However, the consistency in the temporal pattern preceding each goal is quite clear. In this case, the information provided by the T-pattern analysis raises questions about the relationship between events that are spread over an extended period. This does not

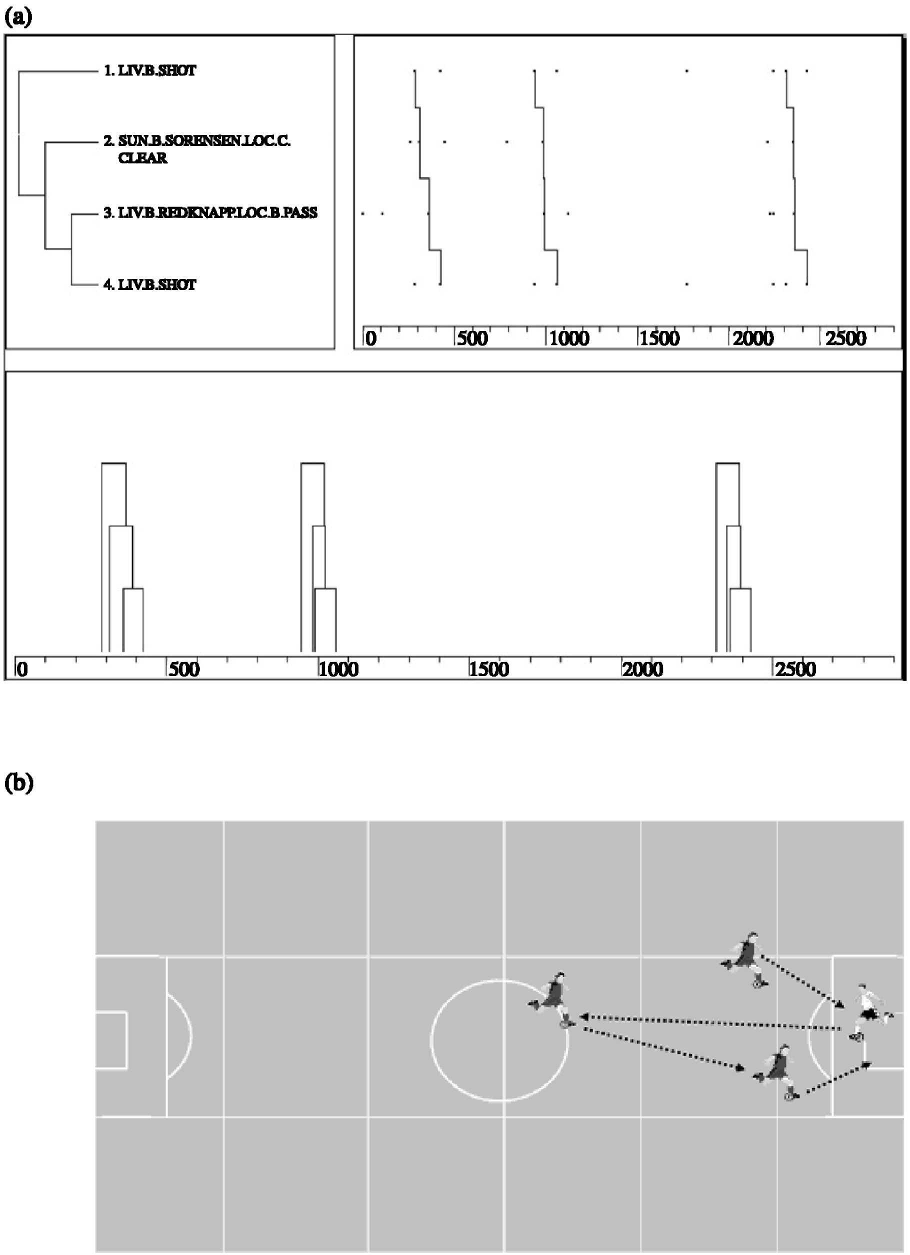


**Fig. 4.** A T-pattern incorporating regularity in shots on goal in a FA Premiership match. The pattern includes six of team A's eight shots. Pattern occurrences 2 and 3 resulted in a goal.

provide a coach with clear and simple answers about the nature of the goals scored, but without the analysis the potential significance of the relationships between events would not have been revealed. This analysis may, therefore, prompt a coach to revisit the video footage of passages of play to identify causally linked elements of performance that might have been missed had the temporal pattern not been identified. Although in the case of shots on goal the coach might normally revisit, using video, the preceding 10–20 s of play if the temporal pattern covers a larger time frame, as in this example, the coach might not be viewing all relevant action. The analysis may, therefore, help to define the time for a visual review of specific aspects of performance.

It is difficult to conceive of coaching scenarios in which the information identified in Figs 3, 4 and 5, and the further analyses that they may stimulate, would not be of value in enhancing coach knowledge. At the very least, the analyses presented provide a perspective on team performance that is rarely, if ever, attained using traditional frequency counts of discrete events within a match.

In addition to a case-by-case consideration of each analysed performance, the data were also considered in

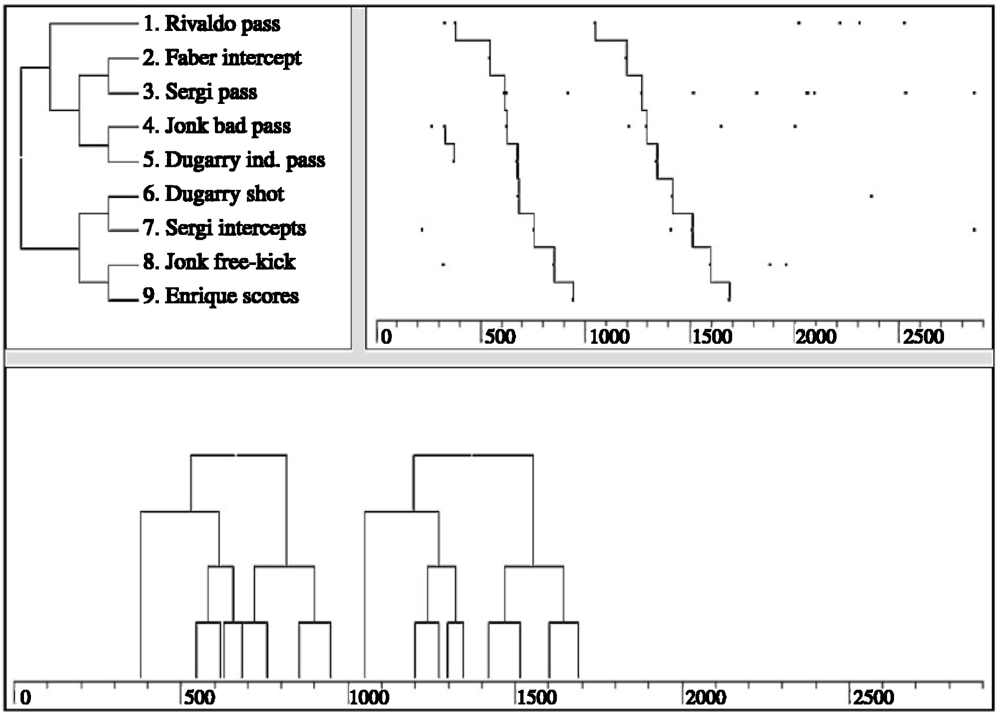


**Fig. 5.** A T-pattern incorporating nine event-types that occurred twice during the second half of a Champions League match in 1997. The pattern occurs over an extended period and concludes with the scoring of a goal on each occasion. (1) Rivaldo passes the ball. (2) Faber intercepts pass. (3) Sergi passes the ball. (4) Jonk makes a bad pass. (5) Dugarry makes an individual successful act. (6) Dugarry takes a shot. (7) Sergi intercepts a pass. (8) Jonk takes a free kick. (9) Enrique scores a goal.

relation to differences in performance standard (club *vs* international competition) and the effect of event frequency on pattern frequency. Although these analyses were relatively simple in nature, they emphasized the potential significance of temporal structure within sport performance.

*Event frequency and pattern frequency*

In all matches analysed, given the coding system used, the total number of coded events was in excess of 1000. The first issue considered, therefore, was whether the number and frequency of detected patterns was simply due to the many data points in each game. To answer this question, results from three international matches were manipulated to construct randomized data files, which were created such that the occurrence times of each event-type in the original data were replaced by the same number of random occurrence times. The randomized data contain the same number of series, each with the same number of occurrences as the original data, and the size of the data set remains the same. A T-pattern analysis was then conducted on the randomized data files and the original coded match data. We found that the randomly distributed files produced fewer patterns than were detected in the original files, as shown in Fig. 6. This suggests that the temporal configuration of play events is due to synchronization and cooperation between players (including interaction with opponents), their actions and movements, rather than being a simple consequence of the number of data points in a complex performance.

*Performance standard*

The effect of performance standard on pattern occurrence was also considered by comparing the mean number of events coded per game, the number of individual patterns identified and the total number of pattern occurrences in three club and three international matches. The matches were randomly selected from the pool of club and international matches that had been analysed. The results in Fig. 7 show that the difference in the number of patterns found per game – and in pattern occurrence per game – between club and international matches was far greater than the difference in the number of discrete events coded per game. This suggests that, within the international matches, individual events were more often part of larger temporal patterns and that more consistency in temporal structure exists within international matches.

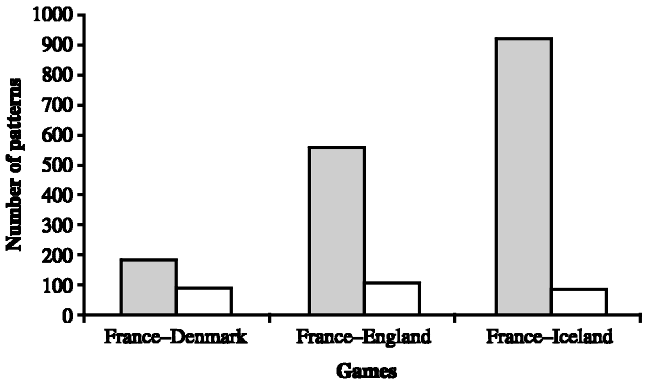


**Fig. 6.** A comparison of the number of T-patterns detected in data files containing randomized coded events and real coded events from three international soccer matches. ■, real data; □, random.

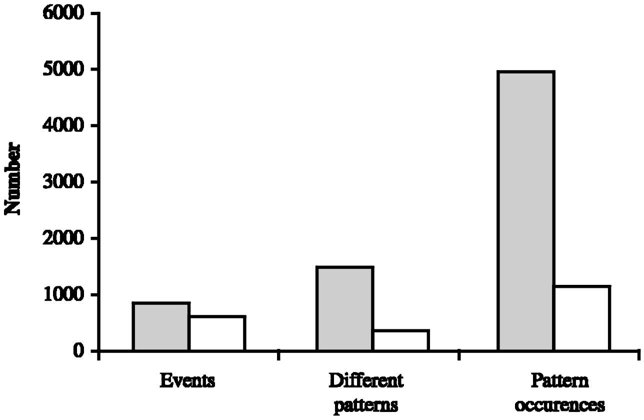


**Fig. 7.** A comparison of club ($n = 3$) and international ($n = 3$) soccer matches in terms of the mean number of separate events coded, the number of different T-patterns detected within the events coded and the total number of occurrences of all T-patterns. ■, international matches; □, club matches.

## Conclusions

Studies that focus on simple frequency analysis cannot identify temporal patterns within a sport performance. This investigation of soccer has shown that temporal patterns, defined here as T-patterns, exist in team sports. The results reported here highlight the potential for T-pattern analysis to move beyond the constraints of traditional frequency-based analyses of performance and make a significant contribution to a deeper understanding of sport performance. The results suggest a very real possibility of discovering new kinds of profiles (complex intra- and inter-individual patterns) for both individuals and teams using the detected behavioural patterns in combination with elementary statistics. Without such an analysis, meaningful information will not be made available to the coach and it is possible that performance will not be optimized. Consequently, we believe that T-pattern analysis has potential as an effective research and support tool in performance analysis.

## References

Bakeman, R. and Quera, V. (1995). *Analyzing Interaction: Sequential Analysis with SIDS and GSEQ*. New York: Cambridge University Press.

Franks, I.M. (1997). Use of feedback by coaches and players. In *Science and Football III* (edited by T. Reilly, J. Bangsbo and M. Hughes), pp. 267–268. London: E & FN Spon.

Franks, I.M. and Miller, G. (1986). Eyewitness testimony in sport. *Journal of Sport Behaviour*, **9**, 39–45.

Magnusson, M.S. (1996). Hidden real-time patterns in intra- and inter-individual behavior: description and detection. *European Journal of Psychological Assessment*, **12**, 112–123.

Magnusson, M.S. (2000). Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments and Computers*, **32**, 93–110.

Noldus, L.P.J.J. (1991). The Observer: a software system for the collection and analysis of observational data. *Behavior Research Methods, Instruments and Computers*, **23**, 415–429.

Noldus, L.P.J.J., Trienes, R.J.H., Hendriksen, A.H.M., Jansen, H. and Jansen, R.G. (1999). The Observer Video-Pro: new software for the collection, management and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments and Computers*, **32**, 197–206.