Using a Situation Awareness approach to determine decision-making behaviour in squash

Stafford Murray¹, Nic James², Janez Pers³, Rok Mandelje³ & Goran Vučković⁴

¹ High Performance Sport New Zealand, Millennium Institute of Sport & Health, 17 Antares Place, Mairangi Bay, Auckland 0632, NZ.

² London Sport Institute, Faculty of Science and Technology, Middlesex University, The Burroughs, Hendon, London, UK.

³ Faculty of Electrical Engineering, University of Ljubljana, Tržaška c. 25, 1000 Ljubljana, Slovenia.

⁴ Faculty of Sport, University of Ljubljana, Gortanova 22, 1000 Ljubljana, Slovenia.

Corresponding author: Nic James, +44 780 961 8153, n.james@mdx.ac.uk

Key words: Situation Awareness; decision-making; squash shots; movement parameters

Running title: Decision-making in squash

1. Introduction

Intelligent decision-making in complex environments requires domain expertise, acquired through many hours of practice and experience. This allows experts to represent problems at a deeper level (Chi, Feltovich, & Glaser, 1981), have faster and more accurate pattern recognition (Chabris & Hearst, 2003) and make more accurate predictions of what others are likely to do (Raab & Johnson, 2007) compared to novices. Most research that has considered how deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993) i.e. tasks are repeated to achieve specific goals in the presence of feedback to refine knowledge and skills, have been undertaken in the domains of music and sport (Ericsson, 2006). Repetitive practice has also
been found to be useful for perceptual diagnosis of abnormalities and surgery (Ericsson, 2004). However, to counteract automaticity and gain high-level control of performance, conscious effort is considered necessary to go beyond routine behaviour and achieve real expertise (Ericsson, 2006; Ericsson, 2004).

In complex sports, situation awareness (SA; Endsley, 1995) refers to the awareness of all relevant sources of information, an ability to synthesise this information using domain knowledge gained from past experiences (Abernethy, Gill, Parks, & Packer, 2001) and the ability to physically respond to the situation i.e. demonstrate expert behaviour (Williams, Davids, & Williams, 1999). Decision-making behaviour is therefore usually considered from the athlete’s perspective, which is closely coupled to the environmental conditions at that time. Relevant sources of information are likely to be related to events previously encountered (historical and within the game being played), opponent movements (visual cues) and probabilistic information such as a heuristic “in this situation it is likely that” (James & Patrick, 2004). Decision-making is therefore viewed as emerging from the interaction between two or more players, under environmental constraints, over time, towards specific goals (Araujo, Davids, & Hristovski, 2006).

In an overview of different SA models Neville & Salmon (2016) suggest Klein’s (1993) recognition-primed decision model (RPD) as the most appropriate for examining decision-making under time pressure in a naturalistic setting (for an overview see also Kermarrec & Bossard, 2014). This model is an alternative to the information-processing approach and utilises a naturalistic decision-making (NDM) framework which proposes that problem solvers under time pressure, assess real world situations by recognising familiar patterns that experience has shown to be useful. For example, Macquet (2009) showed seven professional volleyball players video clips of their performance (within 5 days of the match) and asked them to explain what thoughts and feelings they were experiencing at that time.
Results supported the recognition hypothesis with typical situations matched to a response and atypical situations diagnosed. The cognitive activities related to this process were (a) to identify relevant cues e.g. player positions, (b) form expectancies e.g. possibility of a player movement, (c) determine plausible goals e.g. decide between two likely outcomes, and (d) adopt typical actions e.g. use a familiar technique as the response. Kermarrec and Bossard (2014) used a similar methodology to assess the decision-making of four elite football defenders and found similar results to Macquet (2009) except for a much higher incidence of simulation i.e. cognitively estimating the likely turn of events if a course of action was followed. The explanation for the higher incidence may have been due to the availability of more time than in other studies.

Squash is an intermittent activity characterized by frequent bursts of near maximal activity in a range of directions with regular short recovery periods (Kingsley, James, Kilduff, Dietzig, & Dietzig, 2006) requiring specific fitness (Locke, Colquhoun, Briner, Ellis, O’Brien, Wollstein, & Allen, 1997). Players try to force an opponent to play shots under spatial and temporal pressure by accurate shot placement. Squash shots are named according to their direction, speed and trajectory e.g. straight drive. However, this simple categorisation scheme doesn’t discriminate the variety of shots that can be played under the same category name. For example, a straight drive can be hit at different speeds and trajectories from different parts of the court. Each change in any of these parameters has the possibility of changing the outcome of the shot in terms of its effect on the opponent.

Many factors influence match characteristics e.g. number of games played, player standard and game duration, which has been shown to be related to the difference in World ranking between the two players (Murray, James, Hughes, Perš, Mandeljc, & Vučković, 2016). Indeed, at the elite level, 25% of rallies last less that 7s but 25% last more than 22.7s (Murray, et al., 2016). This may account for the fact that the tactical use of different shot
types to move the opponent around the court has only been shown when court location and
time between shots (Vučković, James, Hughes, Murray, Sporiš, & Perš, 2013) as well as
preceeding shot type and court location (Vučković, James, Hughes, Murray, Milanović, Perš,
& Sporiš, 2014) have been considered. When these variables have been controlled elite
players tend to utilise similar strategies e.g. elite players were highly likely to play one of two
shots when in the back corners returning straight and crosscourt drives (Vučković, et al.,
2014). Consequently, significant differences in performance variables between similarly
ranked elite players are unlikely to be obvious, particularly gross measures such as frequency
of shot types played in different areas of the court.

Data mining is a collection of statistical techniques for converting data into
information in the computer science and artificial intelligence domains (Ofoghi, Zeleznikow,
MacMahon, & Raab, 2013). These techniques have also been used to analyse sports data
(Schumaker, Solieman, & Chen, 2010) leading to the increased use of the terms “sports
analytics” (e.g. www.analyticsinsport.com), “sabermetrics” (e.g. www.sabr.org/sabermetrics)
and “big data management” (www.sas.com/en_us/software/analytics.html;
www.ibmbigdatahub.com/tag/1647). Ofoghi et al. (2013) suggest that data mining techniques
can help avoid the pitfalls associated with reductionist approaches favoured in some
performance analysis research (Mackenzie & Cushion, 2013) enabling the discovery of
hidden or underlying relationships between many factors that either directly or indirectly
influence sports performance.

In this paper, squash is used as the exemplar sport to examine decision-making
behaviour in terms of the selection of a shot type i.e. a goal-driven action to move an
opponent into an area of the court under varying degrees of difficulty. This is determined by
the initial conditions (both players’ movements and locations during the time preceding a
shot; e.g. Macquet, 2009) and the intended outcome (between defence and attempting a
winning shot) which is based on the weighting of the importance of the interaction between both players’ situations (judged against previously encountered similar situation; Klein, 1993). Hence, shot types were quantitatively clustered using multiple parameters related to players’ positions (court areas) and movements (velocity, distance and time) between an opponent playing the shot prior to the player’s shot and the opponent’s following shot. This approach facilitates a better understanding of the decision-making process and could lead to determining the small differences in behaviour between elite players. This novel approach can be adopted in other disciplines to aid the understanding of the nature of expertise.

2. Methods

2.1 Participants
Matches at the 2010 (n = 14) and 2011 (n = 27) Rowe British Grand Prix, held in Manchester, UK were recorded and processed using Tracker software (Vučković, et al., 2014) that is a newer version of the SAGIT/Squash software (Perš, Kristan, Perše, & Kovačič, 2008). Thirty-four full-time professional players of mean age 27.7 years (SD = 3.85) who were ranked in the world’s top 75 participated. Ethics approval for the study was provided by the sports science sub-committee of Middlesex University’s ethics committee in accordance with the 1964 Helsinki declaration.

2.2 Procedure
Matches took place on a court set up with a PAL video camera (Sony HDV handy camera HVR-S270, Japan) with a specially adapted 16 mm wide angled lens (Sony NEX SEL16F28) attached to the ceiling above the central part of the court such that the entire floor and part of the walls were within the field of view. A similar camera (used by the Professional Squash Association to record matches) was located on a tripod 15 m behind the court and 5 m above ground level. The camera placement and techniques for transferring video images into
Tracker were identical to SAGIT/Squash i.e. automatic processing with operator supervision, and have been well documented (Vučković, Perš, James, & Hughes, 2009). Similarly, the reliability for resultant calculations of distance and speed for each player (Vučković, Perš, James, & Hughes, 2010) and positions on court (Vučković, et al., 2009) have been published. The exact camera location for the overhead camera (both vertically and horizontally) was not critically important, as subsequent calibration for image capture accounted for its position.

Data were collected 25 times per second.

Fig. 1 Squash court floor divided into 15 cells

2.3 Data used

The shot type (n=30) and ball location (cell, Figure 1) for each shot (denoted player A), excluding serve, return of serve and rally ending shots (winners, errors, lets and strokes), were recorded along with the same information for both the preceeding shot (B-1) and following shot (B+1). Additional information regarding time, speed and distance were
recorded both between shots and at the time player A hit the ball (Figure 2). Information collected following player A’s shot (variables related to player B’s movements) were used as measures of the action (shot type) selected as the response due to the player’s SA (Macquet, 2009). On this basis serve and return of serve were excluded due to the lack (for the serve player B is stationary) or abundance (the return of serve can be any shot type) of variability in player B variables. Similarly, rally ending shots (winners, errors, lets and strokes) could be any shot type and hence response variables would be extremely varied and hence the analysis complex. For this reason these were excluded from this analysis with the recommendation that a separate study examines these variables.

Fig. 2 Variables collected and those found to discriminate shot types in cluster analysis

Key: *Variables used in cluster analysis (bold and italic)*
- Variables not used
- Ball next to player hitting shot

**Fig. 2** Variables collected and those found to discriminate shot types in cluster analysis
Every shot in squash has a primary function to move the opponent away from the centre of the court. This is not always achieved as inaccurate shots or an opponent’s anticipation can result in shots being played in the centre of the court or non-intended locations. For example, a drive aimed to a back corner may be volleyed near the centre of the court because the opponent anticipated the shot trajectory. As this shot did not achieve its intention the performance parameters following the shot would be very different from a shot that did achieve its aim. Consequently only shots that did achieve their intention i.e. the opponent returned the shot from the cell where the shot was aimed, were selected for analysis (Table 1). Similarly, time and distance parameters associated with shots played from and to the front, middle or back of the court would be different and so these shot types were separated according to the cell played from. Finally, if the return shot of a lob was a volley or a ground stroke the corresponding parameters would be quite different and hence only shots which were not volleyed were included due to insufficient data for volley returns.

2.4 Statistical analyses
Cluster analysis is a data mining technique that enables the formation of groups within a data set based on maximising the homogeneity of cases within a group and the heterogeneity between clusters (Hair, Anderson, Tatham, & Black, 1995). For example, Vaz, Rooyen, & Sampaio (2010) used cluster analysis to determine close, balanced and unbalanced game final score differences in rugby union. Cluster analysis begins with all cases as separate groups and the two “most alike” cases are combined in the first step using the most appropriate distance measure. The two cases with the smallest distance measure will then cluster together and a group mean (cluster centroid) can be calculated and used in the next step. The next two most alike cases (or groups once cases have been clustered) are then combined. This process continues until an optimal cluster solution is obtained, although this may be determined from
a practical standpoint as there are no objective methods for determining the optimal number of clusters (Hair, et al., 1995).

**Table 1** Shot types selected based on court cell played from and cell returned from

<table>
<thead>
<tr>
<th>Shot type</th>
<th>Cell played from</th>
<th>Cell ball returned from</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive(^a) and Drive wall(^b)</td>
<td>11 &amp; 12 or 14 &amp; 15 (front)</td>
<td>1 &amp; 2 or 4 &amp; 5 (back)</td>
</tr>
<tr>
<td>Drive crosscourt(^a) and Drive crosscourt wall(^b)</td>
<td>11 &amp; 12 or 14 &amp; 15 (front)</td>
<td>4 &amp; 5 or 1 &amp; 2</td>
</tr>
<tr>
<td>Drop</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10 (middle) 11 &amp; 12 or 14 &amp; 15 (front)</td>
<td>6, 7, 11 &amp; 12 or 9, 10, 14 &amp; 15</td>
</tr>
<tr>
<td>Drop crosscourt</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10 (middle) 11 &amp; 12 or 14 &amp; 15 (front)</td>
<td>9, 10, 14 &amp; 15</td>
</tr>
<tr>
<td>Boast attack and Boast defence</td>
<td>1 &amp; 2 or 4 &amp; 5 (back) 6 &amp; 7 or 9 &amp; 10 (middle)</td>
<td>14 &amp; 15 or 11 &amp; 12</td>
</tr>
<tr>
<td>Lob</td>
<td>11 &amp; 12 or 14 &amp; 15</td>
<td>1 &amp; 2 or 4 &amp; 5</td>
</tr>
<tr>
<td>Lob crosscourt</td>
<td>11 &amp; 12 or 14 &amp; 15</td>
<td>4 &amp; 5 or 1 &amp; 2</td>
</tr>
<tr>
<td>Kill</td>
<td>1, 2, 6, 7, 11 &amp; 12 or 9, 10, 14 &amp; 15</td>
<td>6, 7, 11 &amp; 12 or 9, 10, 14 &amp; 15</td>
</tr>
<tr>
<td>Drive (v)(^a) and Drive wall (v)(^b)</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10</td>
<td>1 &amp; 2 or 4 &amp; 5</td>
</tr>
<tr>
<td>Drive crosscourt (v)(^a) and Drive crosscourt wall (v)(^b)</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10</td>
<td>4 &amp; 5 or 2</td>
</tr>
<tr>
<td>Drop (v)</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10</td>
<td>6, 7, 11 &amp; 12 or 9, 10, 14 &amp; 15</td>
</tr>
<tr>
<td>Drop crosscourt (v)</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10</td>
<td>9, 10, 14 &amp; 15 or 6, 7, 11 &amp; 12</td>
</tr>
<tr>
<td>Boast attack (v) and Boast defence (v)</td>
<td>6, 7 &amp; 8 or 8, 9 &amp; 10</td>
<td>14 &amp; 15 or 11 &amp; 12</td>
</tr>
<tr>
<td>Kill (v)</td>
<td>1, 2, 6, 7, 8 or 4, 5, 8, 9 &amp; 10</td>
<td>6, 7, 11 &amp; 12 or 9, 10, 14 &amp; 15</td>
</tr>
<tr>
<td>Kill crosscourt (v)</td>
<td>1, 2, 6, 7, 8 or 4, 5, 8, 9 &amp; 10</td>
<td>9, 10, 14 &amp; 15</td>
</tr>
</tbody>
</table>

Key: Drives categorised by whether the return was hit after the ball bounced off the back wall\(^b\) or not\(^a\) and if the shot was volleayed (v)
A two-step cluster analysis using a probability based log-likelihood distance measure (SPSS) enabled both continuous (e.g. distance player B was from the T at the moment player A hit the shot (B distance from T)) and categorical (shot type) variables to be used. The distance between two clusters was related to the decrease in log-likelihood as they were combined into one cluster using the Bayesian information criterion. In calculating log-likelihood, normal distributions for continuous variables and multinomial distributions for categorical variables were desirable but not necessary (Norušis, 2011). All variables associated with player B returning the shot in the correct cell were input into the cluster analysis which produced a poor cluster quality. The input variable that was the least powerful predictor was then removed from the analysis. This was repeated until 4 continuous variables (bold and italic in Figure 2) produced cluster qualities rated as fair (n = 18805, 55.1% of all shots):

- Time between player A’s shot and player B returning the shot \((\text{Time})\)
- Distance player B moved to return player A’s shot \((\text{Distance for B})\)
- Maximum velocity of player B from the moment player A hit the shot to player B returning the shot \((\text{max speed for B})\)
- Distance player B was from the T at the moment player A hit the shot \((B_{\text{distance from T}})\)

The continuous variables used were largely independent with the highest correlation \((r = 0.59)\) between distance and the maximum speed of player B during the time between player A hitting shot and player B hitting the return shot. The noise handling option was selected so that outliers that were present could be included in a separate cluster. Match data were entered into SPSS randomly to reduce order effects (Norušis, 2011).

2.4.1 Determining the number of clusters

The two-step cluster analysis initially formed a four cluster (one for outliers) solution for the data, identifying the cluster qualities as fair. Whilst there are no objective methods for determining the optimal number of clusters (Hair, et al., 1995) five and six cluster solutions
were also produced (both cluster qualities fair). Two squash coaches, each with over 20 years’ experience at National level, discussed the practical value of the three different cluster solutions favouring the six cluster solution for its usefulness in an applied context.

3. Results

Six SA clusters were named to relate to the outcome of a shot (Figure 3). A “defensive” shot occurred when the player was under a lot of pressure (high, 23.4% of shots) and attempted to create time by playing a slow shot (low opponent pressure). At the other end of the scale, an “attempted winner” was played when the player was not under pressure (very low, 2.0% of shots) and the opponent was out of position (very high opponent pressure).

As shot types were classified according to which court areas they were played from (Table 1) some shot types occurred in just one cluster e.g. 99.8% of drives from the back of the court were pressing shots, whereas others were classified in more than one cluster e.g. drive from the front of the court was clustered into pressure (30.8%), attack (55.6%) or attempted winners (10.5%, Figure 4).
Key: All values are means (standard deviation underneath in italics)

Distance from T: at time of player A playing shot

Time, distance and speed (maximum attained): from when player A played shot
to when player B hit return

Fig. 3 Time, distance and speed parameters for the six SA clusters
Fig. 4 Proportion of shot types classified in six SA clusters
4. **Discussion**

Previous research in squash has typically considered the pattern of shots played from different areas of the court as being indicative of tactics (Vučković et al., 2013; Vučković et al., 2014). Whilst this presents a general pattern of the shot selection for a given situation the generality of the approach hasn't discriminated differences between similarly ranked players. A new approach which described shot selection from a SA perspective was presented here. Shot types were quantitatively clustered using parameters related to players’ positions (court areas) and movements (velocity, distance and time) between an opponent playing the shot prior to the player’s shot and the opponent’s following shot. These variables were selected as they included factors influencing a player’s decision-making (variables up to the point at which a shot was played) and variables which were a consequence of the shot (measures of the shot’s effectiveness). Six SA clusters were formed based on the opponent’s position at the time of playing the shot and the subsequent movement parameters related to returning the shot.

The first SA cluster “defence” occurred when a player was under high pressure and selected a variety of shots, determined by the location on court, to try to increase the shot to shot time to reduce the pressure imbalance (23.4% of shots). The second SA cluster “maintain stability” was the straight drive from back corner to back corner which hit the back wall before the opponent played a return (14.3% of shots). The next cluster “pressing” differed from the previous cluster in that it involved the same shot, the straight drive, but the shot did not reach the back wall (24.6% of shots). This meant there was slightly less time between the shot and the return (mean = 1.3s) than for the previous cluster (mean = 1.9s) and hence more pressure. The straight drive is the most prevalent shot in squash (Vučković et al., 2013) although previous research has not distinguished whether the shot reached the back wall or not. The next cluster “pressure” involved a variety of shots (n=10) which increased
the distance the opponent moved and his maximum speed (between the moment player A hit
the shot to player B returning the shot). These shots (20.3%) tend to put a little bit more
pressure on the opponent by increasing the tempo of the rally but involve more risk. The next
cluster “attack” differed from “pressure” shots in that the opponent tended to be further from
the T (mean = 1.0m compared to 0.6m) meaning that the opponent needed to cover more
distance and at higher speeds. This cluster (15.3% of shots) was the only one to contain the
straight drop (from the front and middle), straight kill and straight volley drop shots. These
shots reduce the number of return options for the opponent due to the limited time and space
(Vučković et al., 2013). The final cluster “attempt winner” occurred when the opponent was
well out of position (mean = 2.6m from the T) and the player hit into an open court away
from the opponent’s position (2.0% of shots). This situation rarely occurs in elite level
squash.

Decision-making in complex sports, like all complex tasks, requires fast processing of
many information sources (Chabris & Hearst, 2003). In squash, to achieve SA, the player
must determine where the ball will travel, move to an appropriate position, decide which shot
to play and execute the shot with precision, all under considerable time pressure (Chabris &
Hearst, 2003; Williams, Davids, & Williams, 1999). Expertise, it is reasonable to suggest, is
correlated with faster and better performance on each of these SA tasks. The same shot when
played from different areas of the court produced different outcomes (SA clusters) due to the
player having different objectives (SA tasks). Similarly, the same shot type played from the
same court area produced different outcomes (SA clusters). This latter finding would have
been expected if all shots had been analysed as poorly hit shots or opponent anticipation
would have explained the difference found. However, these shot outcomes were excluded
from the analyses. Therefore, the different SA clusters for the same shot type from the same
court area is consistent with players changing the pace and trajectory of the shot because of
different objectives (SA tasks). The discrimination of these differences are unique to this study and offer a new insight into the very subtle differences between players of quite similar expertise, unlike previous studies that typically show expert novice differences (Chi, Feltovich, & Glaser, 1981). For example, the very best players may play a higher proportion of one shot in one area of the court more often in the higher-pressure SA clusters. This may be due to better shot accuracy, hitting shots earlier due to more efficient movement to the ball (Chabris & Hearst, 2003; Raab & Johnson, 2007), or a different shot selection (pace and trajectory) due to recognising the situation differently or having a different response option available. Thus simply categorising a shot type in squash, without reference to other variables associated with time and movement variables, has been shown to be insufficient to capture the complexity of the task.

Winners and errors were not included in this study as these shots are usually a manifestation of the consequences of the most recent shots in a rally. As such they require a detailed analysis connecting shot to shot information and the interaction between players. Similarly, serves and return of serves were not analysed here. These are very important shots at lower levels of squash where rallies tend to be shorter than for elite matches (Hughes, Wells, & Mathews, 2000) and the impact of a good or weak serve or return of serve can dramatically affect the rally outcome. However, at the elite level where rallies are much longer the impact of the serve and return of serve are much less. These were excluded from this study because they are determined by the rules of the game in terms of where they take place on court and are thus quite different from the other shots played. The removal of these shots from the analysis (winners, errors, serves and return of serves) is a limitation of the study but their inclusion would have complicated the analysis significantly. It is therefore suggested that a future study examines these very important shots using a similar methodology to the one here.
This research has provided a methodology which can lead to determining the small differences in behaviour between elite players both in this sport and others as well as other disciplines to aid the better understanding of nature of expertise in terms of SA. For example, team sports where pre-planning can be undertaken prior to decision-making, such as handball, basketball and football, require decisions to be made based on opponent (and teammate) positions, pitch area and movements. These parameters, along with measures related to the performed action, could, as in this study, be used to discriminate expert behaviours. To achieve this in squash, future research needs to compare players of different elite standards, as well as different standards generally, to determine how the pattern of their SA clusters differ. This information can be used by coaches to determine areas for development of their players to reach the next level of performance.

5. Conclusion

Cluster analysis classified squash shots into six categories determined by the amount of pressure the players was under whilst playing the shot and the resultant pressure the opponent was under due to the shot. This resulted in the most prevalent shot in squash, the straight drive from the back of the court, being classified as either hitting the back wall (maintain stability) or not (pressing). This distinction, not previously identified in the literature, has the potential for discriminating expertise difference in both decision-making and skill level (movement and shot quality).

6. References


