PROFILING ELITE MALE SQUASH PERFORMANCE USING A SITUATION AWARENESS APPROACH ENABLED BY AUTOMATED TRACKING TECHNOLOGY

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

STAFFORD RON MURRAY

LONDON SPORT INSTITUTE
FACULTY OF SCIENCE & TECHNOLOGY
MIDDLESEX UNIVERSITY

JULY 2018

ABSTRACT

The pioneering research into squash performance by Sanderson and Way (1977) tested the hypothesis that "an individual exhibits a pattern of play which is relatively stable over time and independent of the opponent". This belief influenced research in this area for many years. Many studies attempted to discover these stable patterns of play e.g. Hughes, 1985; McGarry and Franks, 1996; Hughes, Evans and Wells, 2001; and presented, what were suggested as indicative, playing profiles e.g. Murray and Hughes, 2001; Hughes, Watts, White and Hughes, 2006.

This PhD aimed to analyse squash match play at a more detailed level by using movement data to supplement shot information related to shot types, court area etc. Elite male squash matches were filmed using a fixed overhead camera and images processed in software to semi-automatically calculate player movement information as well as allow an operator to manually input shot information. The two data streams were synchronised in Matlab before reliability and accuracy testing. Good levels of reliability were found for all court locations and shot information (agreement > 90%) although when an operator coded a long match without a break some percentage agreements had less than 90% agreement, presumably due to fatigue effects. Error testing, using a series of queries, specific to each data type, following data collection and prior to data analysis, discovered multiple errors in the data which were corrected.

The physical demands and rally characteristics of elite-standard men's squash had not been well documented since recent rule changes (scoring and tin height). Rallies were split into four ball-in-play duration categories using the 25th (short), 75th (medium), 95th percentiles (long) and maximum values. The frequencies of shots played from different areas of the court had not changed after the adoption of new rules but there was less time available to return shots.

Chapter 5 considered how expert squash players use Situation Awareness (SA) to decide on what shot to play. Shot type, ball location, players' positions on court and movement parameters were captured 25 times per second for shots that achieved their objective. Six SA clusters were

named to relate to the outcome of a shot ranging from a defensive shot played under pressure to create time to an attempted winner played under no pressure with the opponent out of position.

The pressure exerted on a squash player is a coupling of the two players' SA abilities. The same variables used for Chapter 5 were used except all shots (excluding serves and rally ending shots) were used, producing five main SA clusters, where a greater proportion of shots were categorised in the greater pressure clusters and less in the lower pressure ones. Individual matches were presented using cluster profile infographics which demonstrated how individual players played differently in different matches.

ACKNOWLEDGEMENTS

My fantastic (and patient!) wife Sarah Murray and awesome son Bowdon Murray for being loving, supportive and always helping me throughout this project (and in life!).

My mother Lynda Stokes and brother Professor Warwick Murray (with his wife Ximena and the girls Franky and Gabriella) for making me believe that I had the ability to complete a PhD, always being there for me and making life fun.

Nanny Jean and Grandad Ray for always believing in me whatever I chose to do. I still have Granddads voice in my head now when I make decisions in life, RIP.

All the England Squash Team Players and Coaches (especially David Pearson and Paul Carter for their friendship and mentoring) for making the 20 years of working with them a privilege, fascinating, motivating and most importantly fun.

The English Institute of Sport and the Professional Squash Association for helping fund this project and allowing access to all the elite players on the world circuit.

High Performance Sport New Zealand (especially Peter Pfitzinger and Kerry Spackman) for supporting me through the final stages of completion.

Professor Mike Hughes for all the tireless support, incredible knowledge and life changing opportunities he has provided me over the past 30 years.

Finally, and most importantly to me, Professor Nic James and Dr Goran Vučković. Their thankless and endless support, drive, knowledge, patience and friendship has made this journey possible, challenging, a complete honour to be part of and most importantly a great experience in life; learnings from this journey and these relationships will stay with me and help guide me for life. I cannot thank these two enough….

PUBLICATIONS FROM THIS THESIS

Full SCI Academic Journal Papers

1. Murray, S., James, N., Perš, J., Mandeljc, R. & Vučković, G. (2018). Using a situation awareness approach to determine decision-making behaviour in squash. *Journal of Sports Sciences*, 36, 12, 1415-1422.
2. Murray, S., James, N., Hughes, M.D., Perš, J., Mandeljc, R. & Vučković, G. (2016). Effects of rule changes on physical demands and shot characteristics of elite-standard men's squash and implications for training. *Journal of Sports Sciences*, 10, 2, 129-140.
3. Vučković, G., James, N., Hughes, M., Murray, S., Milanović, Z., Perš, J. & Sporiš, G. (2014). A New Method for Assessing Squash Tactics Using 15 Court Areas for Ball Locations. *Human Movement Science*, 34, 81-90.
4. Vučković, G., James, N., Hughes, M., Murray, S.R., Sporiš, G. & Perš, J. (2013). The effect of court location and available time on the tactical shot selection of elite squash players. *Journal of Sports Science and Medicine,* 12, 66-73.

Research book

1. Murray, S., Hughes, M., James, N. and Vučković, G. (2016). *The science of sport: Squash*. Wiltshire: The Crowood Press Ltd.

Chapters in conference books

1. Murray, S., James, N., Dineen, P., Hughes, M. and Vučković, G. (2013). The effect of changing the scoring system on game related activity in squash. In Peters, D.M. and O'Donoghue, P. (eds) *Performance Analysis of Sport IX*. Oxon: Routledge, pp. 151-155.

Conference presentations

1. Murray, S., James, N. & Vučković, G. (2017). Shot selection in squash: A decision-making approach using cluster analysis. Research paper presented at the 8th International Scientific Conference on Kinesiology, Opatija, Croatia, May
2. Murray, S., James, N. and Vučković, G. (2016). Shot types played by the World No. 1 squash player progressing through 4 matches to the final of the British Grand Prix 2011. Research paper presented at the 11th World Congress of Performance Analysis of Sport, Alicante, Spain, November.
3. Vučković, G., James, N., Murray, S. and Hughes, M. (2013). Feet placement for back corner shots in elite squash. Research paper presented at the 18th Annual Congress of the European College of Sport Science, Barcelona, Spain, June. In N. Balagué, C. Torrents, A. Vilanova, J. Cadefau, R. Tarragó and E. Tsolakidis, (Eds.), *Book of Abstracts of the 18th Annual Congress of the European College of Sport Science*, Barcelona, Spain, pp. 829.

Keynote presentation

1. James, N., Murray, S. & Vučković, G. (2017). The evolution of Performance Analysis in Squash: From simple notation to the discrimination of complex decision-making. Keynote presentation at the 2nd Asia Pacific Conference on Performance Analysis of Sport, Langkawi, Malaysia, October.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

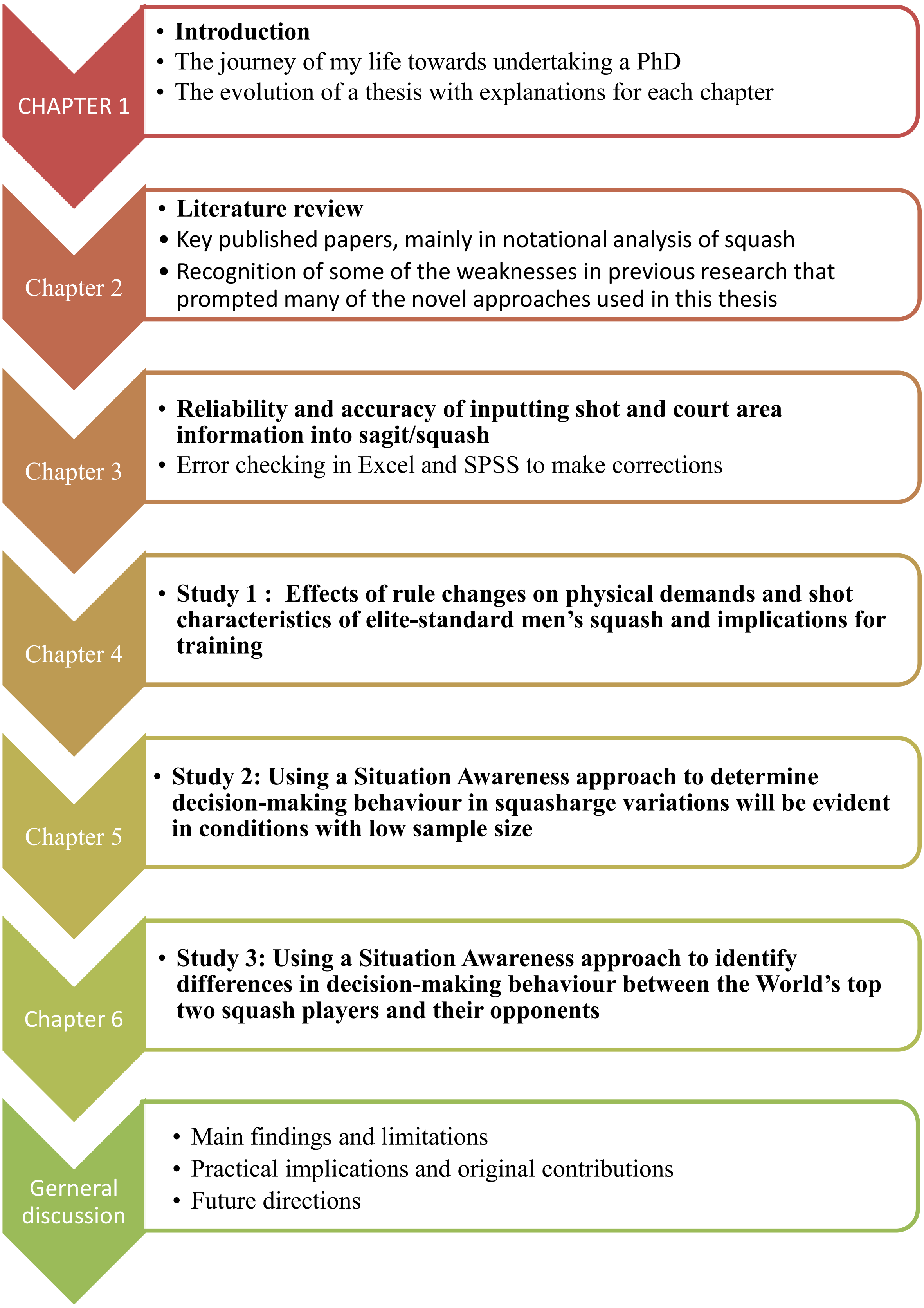
OVERVIEW OF THESIS

This thesis starts with an introduction (Chapter 1) to my journey towards starting a PhD as a consequence of playing and coaching squash coupled with the academic pursuit of degrees in Sports Science. With professional involvement in elite squash I was putting my academic work into practice and was determined to both assess my current practice and develop better and more valid notational analysis practices. This chapter briefly summarises how the studies evolved finishing with the aims and objectives of the study.

Chapter 2 summarises the key published papers, mainly in notational analysis of squash, related to the studies undertaken in this PhD. The chapter is not a list of every research project detailing aspects of squash performance because most of them are outdated due to rules changes (lower tin and different scoring systems). Many have also followed similar paths in regards the methods and presentation of results, both of which this chapter critically appraises. Indeed, it was the recognition of some of the weaknesses in previous research that prompted many of the novel approaches undertaken in the subsequent studies.

The following four chapters (3-6) are the research studies written in a format for publication with the exception of the methods sections which referred to previous studies where repetition occurred. The first study was not submitted for publication, as presented here (Chapter 3), as two parts of this study were included in two publications (Human Movement Science and Journal of Sports Science and Medicine) where data from a previous PhD study (one of the supervisors) were also used. The focus on this study was the reliability of the data collection systems, a fundamental prerequisite enabling the following studies to be carried out. Chapters 4 and 5 have been published in the Journal of Sports Sciences whereas the final study will be submitted after this thesis has been submitted.

The final chapter is a general discussion where I reflect on the main findings of the thesis. The original contributions from both applied and academic perspectives are indicated along with the key limitations. Finally, ideas for future research are suggested.



CHAPTER 1

- **Introduction**
- The journey of my life towards undertaking a PhD
- The evolution of a thesis with explanations for each chapter

Chapter 2

- **Literature review**
- Key published papers, mainly in notational analysis of squash
- Recognition of some of the weaknesses in previous research that prompted many of the novel approaches used in this thesis

Chapter 3

- **Reliability and accuracy of inputting shot and court area information into sagit/squash**
- Error checking in Excel and SPSS to make corrections

Chapter 4

- **Study 1 : Effects of rule changes on physical demands and shot characteristics of elite-standard men's squash and implications for training**

Chapter 5

- **Study 2: Using a Situation Awareness approach to determine decision-making behaviour in squasharge variations will be evident in conditions with low sample size**

Chapter 6

- **Study 3: Using a Situation Awareness approach to identify differences in decision-making behaviour between the World's top two squash players and their opponents**

Gerneral discussion

- Main findings and limitations
- Practical implications and original contributions
- Future directions

Figure 0.1. Infographic showing outline of thesis and brief overview of chapters

Chapter 1: Introduction

It is often uncertain when the journey of a PhD starts, and this one certainly didn't when I first started playing squash aged 8. However, a very successful junior career, England Junior Captain and World junior number 1, culminated in becoming a professional player on the International circuit, certainly developed an intense interest in the game, particularly, the nature of expertise.

Having played professionally for 2 years, I decided to change focus and increase my scientific knowledge. At Cardiff Metropolitan University (1993-2009) I gained graduate and post graduate qualifications (BA (Hons), PG Cert, MSc) in the areas of Human Movement, Coaching Science and Analytics in Sport whilst continuing playing squash (World student team champion). Simultaneously, developing my coaching qualifications and experience in squash (Level 5 (International) Coach and Wales and England National Senior Squads) whilst undertaking and presenting my research at International conferences.

Opportunities arose outside of squash due to my growing expertise in performance analysis of sport (South African Cricket, Welsh Rugby, Welsh Football, England Badminton) culminating in the role of Head of Sports Analysis, Skill Acquisition and Biomechanics at the English Institute of Sport (2009-2015). This job enabled me the time to continue working in elite squash as England Squash Team Leader (2009-2014) and enrol for a PhD (2009).

Studying for a part-time PhD whilst working full-time provided many challenges, mainly devoting specific time to this enormous task, whilst also undertaking consultancy roles with some of the biggest global names in sport and technology (McLaren FI, England Cricket, Qatar Aspire, GB Tennis, Cisco Communications and Esso Petroleum). Consequently, whole weeks were devoted to research activity either in Swansea, Wales (Prof James' home), Kranj, Slovenia (Dr Vučković's home) or Manchester, England (my home).

At the outset of the PhD, data collection and reliability dominated proceedings. The first task was to decide what data collection procedure to use. The obvious strategy was to use SAGIT/Squash software, which had initially been developed by one of my supervisors (Dr Goran Vučković) and continually refined by (Dr Goran Vučković and Professor James). This was the only software that enabled the tracking of player movements in an indoor court at the time. It also enabled the recording of events e.g. shot types which could be synchronised with the movment data. Alternative softwares such as Dartfish had the capability of recording these events and could have been synchronised to the movement data although this type of software did not allow the adition of the postiion of the ball at the time of each shot. This was possible in SAGIT/Squash and a clear advantage for subsequent analyses, hence the decision to use SAGIT/Squash.

Two International competitions taking place in Manchester, where I worked with the England squash team (2010/11), were selected and permission gained from the venue and the Professional Squash Association (PSA, Appendix 1.1) to place a camera above the all-Perspex show court and record the matches. During this period, I had to learn how to use the SAGIT/Squash software (see Appendix 1.2 for visualisations and screenshots of its use), which enabled semi-automated (operator supervision to correct tracking errors) player tracking using computer vision as well as allowing shot information to be added later. For this PhD this software was modified to allow some extra information to be recorded e.g. foot placement when playing a shot, although this variable was subsequently not used in this thesis but is planned for a future publication! It was also 5 years after starting the PhD before all the data had been processed and was ready for analysis. Whilst this was mostly due to the limited availability of time, other factors played a part. During this time my two supervisors were still analysing data collected for Dr Vučković's PhD (data collected in 2003) and they suggested that I reanalyse some of this data to both learn how to use the system but also to undertake reliability studies for inclusion in papers that were subsequently published in 2013 and 2014. This work also allowed the data collected for this PhD to be tested for both reliability and accuracy and this work has been presented in Chapter 3.

Many conversations over the 4 years of collecting, testing and checking the data led to some fundamental questions, inspired from a huge wealth of practical squash experience (both supervisors are National level players and International coaches) and academic knowledge. We all shared a passion for undertaking research that was academically rigorous but practically significant.

The first study (Chapter 4) investigated the effect that rule changes had made on how squash was played at the elite level. We had been developing training regimes to mimic match conditions with the England squash teams but realised that if the PSA's desire to make squash more entertaining by reducing game time through rule changes had had the desired effect then our training regimes needed changing.

The second study (Chapter 5) was inspired by a desire to challenge and change the usual approach to studying squash play. Utilising research in Psychology related to the nature of expertise in terms of decision-making we considered a situation awareness (SA) approach for determining the objective of a player's shot in terms of the pressure exerted on an opponent. This novel approach elicited 6 different SA clusters derived using a data mining approach.

The final study (Chapter 6) tried to address the "theory-practice" gap identified by McKenzie and Cushion (2013) and utilised the approach developed in Study 2 to discern individual differences in playing style. The top 2 players in the World, at the time of data collection, were contrasted both between and within their performances, to identify differences of practical relevance to players and coaches.

The work undertaken in this thesis has contributed to a number of peer reviewed International journal publications but has also inspired a book edited by this research team, conference presentations and keynote talks. On a personal front I have managed my professional life around that of a PhD student to include roles as Head of Analytics and Performance Insights for Team GB at two Olympic games (London 2012 and Rio 2016) as well as contributing to numerous teams competing at World championships, European Championships and Commonwealth Games for British and English teams. Recently I have taken on a new role as Innovation manager for High Performance Sport New Zealand (January 2017) meaning that weeks spent discussing squash research, analysing data and writing research papers with my supervisors has become more technological and less personal, something that is definitely not what a PhD should be all about.

1.1 Aims and objectives

The aims of the thesis were to:

1. Determine the accuracy of event coding using the SAGIT/Squash (updated version named Tracker) software

   Related objectives:

   - Undertake training on using software (O1a)
   - Determine correct methods for assessing reliability of notating: court areas where ball was hit, associated shot information e.g. shot type and the effects of prolonged notating matches (O1b)
   - Determine error checking methods to ensure data accuracy (O1c)

2. Explore elite male squash match play to discover any impacts that rule changes had made

   Related objectives:

   - Evaluate sufficient matches played under point on serve to 9 and point per rally to 11 to understand the variability of elite squash match play e.g. rally length (O2a)
   - Identify the best statistics to present match play variables to indicate the breadth of the game as opposed to the average (O2b)
   - Utilise effect size statistics to meaningfully compare differences (O2c)
   - Describe the effect of the difference in playing standard (based on World ranking) on match play characteristics (O2d)
   - Develop ghosting routines (training to mimic match play demands) for aspiring elite players (O2e)

3. Describe shots in squash according to their objective of placing the opponent under pressure using a novel situation awareness (SA) approach

   Related objectives:

    - Understand the factors (three tasks of SA: identification of relevant sources of information, synthesis using domain knowledge, physically respond) which facilitate decision-making (shot selection) in squash (O3a)
    - Determine variables (time, distance, speed, shot type, court area) that best discriminate the amount of pressure an opponent is put under following a shot that achieved its goal using data mining (cluster analysis) (O3b)
    - Present the amount of pressure placed on an opponent by different shot types played from different areas of the court (O3c)

4. Determine the decision-making behaviours (shot selections) of the top two squash players in the world playing against each other and against different ranked opponents

   Related objectives:

    - Use the SA approach determined in Study 4 to compare clusters formed using only shots that achieved their goal (as in Study 4) with all shots i.e. include shots that didn't achieve their goal (O4a)
    - Determine a methodology for comparing SA clusters derived from different data sets (O4b)
    - Develop an infographic for displaying the four variables used to determine SA clusters and use it to present the decision-making behaviours evident in different matches (O4c)

1.2 Operational definitions

The game of squash involves players (usually two competing against each other although doubles dies take place) hitting a ball against a front wall and returning an opponent's shot before the ball bounces twice. Shots are name according to the direction and speed they are played and whether the ball bounced on the floor (ground shot) or not (volleyed). The shot is also labelled according to whether the ball was hit from a forehand or backhand perspective. In this thesis shots that reached the back wall were discriminated from those that did not for some shots (labelled wall, Table 1.1).

Table 1.1: Operational definitions for squash shot types

Shot type	Definition
Serve	1st shot of the rally (not analysed in this thesis)
Straight drive (wall added if ball touched back wall, volley added if ball hit before ball bounced on floor)	Shot aimed to the back corner on the same side of the court as it was hit from
Crosscourt drive (wall added if ball touched back wall, volley added if ball hit before ball bounced on floor)	Shot aimed to the back corner on the opposite side of the court as it was hit from
Straight drop (volley added if ball hit before ball bounced on floor)	Low soft shot aimed at the front of the court on the same side of the court as it was hit from
Crosscourt drop (volley added if ball hit before ball bounced on floor)	Low soft shot aimed at the front of the court on the opposite side of the court as it was hit from
Straight kill (volley added if ball hit before ball bounced on floor)	Low hard shot aimed at the front of the court on the same side of the court as it was hit from
Crosscourt kill (volley added if ball hit before ball bounced on floor)	Low hard shot aimed at the front of the court on the opposite side of the court as it was hit from
Two wall boast (volley added if ball hit before ball bounced on floor)	Shot hit onto the nearest side wall such that it then hits the front wall and bounces on the floor before touching the opposite side wall
Three wall boast (volley added if ball hit before ball bounced on floor)	Shot hit onto the nearest side wall such that it then hits the front wall and then hits the opposite side wall before bouncing on the floor
Straight lob	High soft shot aimed to the back corner on the same side of the court as it was hit from
Crosscourt lob	High soft shot aimed to the back corner on the opposite side of the court as it was hit from

Chapter 2: Review of literature

2.1 Introduction

The notation of sports events to create an objective record of behaviour can be traced back to the notation of dance (Laban, 1948), baseball (Fullerton, 1912) and basketball (Messersmith and Corey, 1931). However, in the UK and Europe, probably the most influential and certainly the most controversial notational analyst was Charles Reep, who died in 2002 (Pollard, 2002) having devoted over 50 years to analysing football (and other sports) in great detail. The first hand notation system for racket sports in the UK was developed for tennis (Downey, 1973) but due to its complexity was never used to gather data (Hughes, Hughes and Behan, 2007). The works of Reep and Downey inspired researchers at Liverpool Polytechnic who started the first sports degree, independent of Physical Education, in the UK in the mid-1970s. From the staff, some influential researchers emerged, including Reilly and Thomas (1976) who coded football players' movements into standing, walking, trotting, running and sprinting categories. This relatively simple analysis had profound consequences as football coaches were able to match training schedules to actual match demands for the first time. Similarly, Sanderson and Way (1977) adapted Jake Downey's (1973) notation system for tennis to pioneer the analysis of squash. Mike Hughes, a new colleague at Liverpool Polytechnic (1981) and squash coach, became fascinated at the possibilities of this approach. He consequently developed undergraduate academic courses and, along with a multitude of students, notation systems for a wide range of sports. Whilst much of this work was unpublished, or published within proceedings of conferences, the impact of Mike's work was profound. He promoted Performance Analysis, most notably notational analysis, by instigating the International Society of Performance Analysis of Sport in 1992 (formerly known as the International Society of Notational Analysis) and later the International Journal of Performance Analysis in Sport in 2001. By founding these two important outlets for academic work in Performance Analysis Mike led the rapid growth in this area. His influence today reaches across the world and the large expansion of academics in this area has seen a similar rise in publications, both textbooks and research papers, published in a wide range of high impact International journals.

2.2 Notational analysis of squash

Sanderson and Way (1977) hand notated squash matches live and hence recognised the need for speed in their recording methods. Their solution was to pre-print schematics of the court floor and to record just one stroke per schematic using illustrative symbols. Seventeen stroke symbols plus a system for recording winners, forced and unforced errors enabled descriptive statistics to be produced. They found that the backhand drive was by far the most favoured stroke, which is still true today as shots played from the backhand side (for right handers) are still the most prevalent in elite squash (Murray, James, Hughes, Perš, Mandeljc and Vučković; 2016) and the drop shot the most erratic with a high incidence of both winners and errors.

Sanderson and Way (1977) also tested the hypothesis that "an individual exhibits a pattern of play which is relatively stable over time and independent of the opponent". Using the product-moment correlation coefficient their results indicated that players showed a higher degree of similarity when winning compared to losing. This concept of a pattern of play has largely persisted in the literature where its meaning relates to a description of the relative frequency of each stroke a player made in the matches analysed. Fundamental to this is the assumption that if players demonstrate a relatively stable playing pattern then opponents can make use of this information to their advantage. Hughes (1985) was the first to produce a computerised notation system, similar to Sanderson and Way's (1977), with adaptions made in regard to recording player's positions (16 rectangular cells, Figure 2.1) with a view to identifying patterns of play in squash.

Figure 2.1: Squash court floor split into 16 cells as used by Hughes and co-workers.

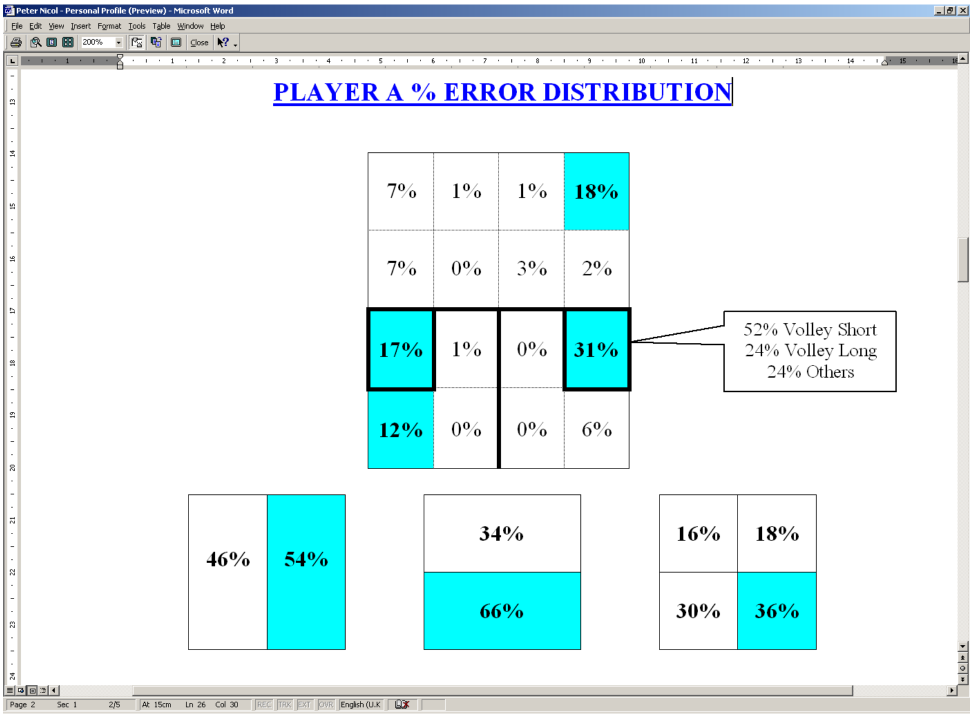
This approach dominated this type of research for the next 15 years or so (e.g. Hong, Robinson, Chan, Clark and Choi, 1996; Hughes, Wells and Mathews, 2000). Murray and Hughes (2001) presented the most evolved version of these methods, which they used working with England Squash, at the 5th World Congress of Performance Analysis of Sport. Two computerised systems were used, a real-time winner and error analysis that included the number of shots in a rally and a post-event full analysis. A Microsoft Access database allowed filtering of the data such that individual players, shot type and cell locations could be presented. For example, winner error ratios and error locations for a player in rallies lasting 15 shots or more were presented. Similarly, error rates by cell and shot directions from the four corner cells were combined over a number of matches. The meaningfulness of this type of information is primarily based on the assumption that there is some likelihood that these patterns are representative of the player or groups of players being presented. However, this was never verified in a peer reviewed publication and hence the validity of this type of study has yet to be scientifically ascertained.

2.3 The characteristics of elite squash

At the outset of the emergence of notational analysis of squash, fundamental questions relating to how squash matches manifested themselves were asked. Pertinent questions relating to rally length, game and match duration as well as ball-in-play time, winner and error rates and shot selection were recorded with a view to aiding training regimes for match preparedness, determining playing strategy changes due to rule changes and identifying individual differences. For example, Hughes, Watts, White and Hughes (2006) found that the average rally length for elite male matches played under point-per-rally rules to 15 was 25.4s, point-per-rally rules to 11 was 26.0s compared to point-on-serve to 9 which was 18 seconds (Hughes, 1985) or 21 seconds (Hughes and Robertson, 1998). This type of finding was indicative of the time, in that average values were presented, measure of dispersion usually omitted, and sometimes the data subjected to inferential statistics enabling a suggestion that change had or had not taken place. Whilst these studies undoubtedly advanced the layman's knowledge of the structure of the game of squash the (usually) large divergence within a playing standard was not presented. Furthermore, little consideration was made regarding the validity of the sample to generalise to a population. Indeed, in many instances sample sizes were quite small, suggesting that to some extent the findings may have only be applicable to that particular sample.

In relation to determining adequate sample sizes, debate centred on the number of matches required to create a stable (the authors also used the word normative) profile (Hughes, Evans and Wells, 2001). This paper advocated methods that purportedly determined how many matches were required to derive an accurate average i.e. supposedly a value deemed to be representative of typical performance (presumably population statistics) and referred to as a performance profile for any single variable e.g. rally duration. The chosen method was based on whether the moving average changed by less than 5% when an additional match was added to the data set. This was clearly an attempt at answering a fundamental question, namely, how much data is needed to provide a reasonably accurate assessment of any particular variable? The moving average methodology average has been commented upon by O'Donoghue (2005) who suggested that using the average as the weighting variable may be inappropriate for some variables where small differences could be significant but deemed tolerable. He also suggested that weighting more recent values higher than older ones could be sensible. O'Donoghue (2005) therefore presented "profiles" for tennis players where performances for many players (e.g. 143 elite female players) were presented as percentiles. This approach is more akin to the view here, that a simple average is not particularly informative. This concurs with the old adage used by statisticians which suggests that a mean (arithmetic average) is meaningless without a standard deviation!

A typical example of an attempt at presenting a player's profile was given by Murray and Hughes (2001), here the error distribution of a player was "normalised" over 5 matches (Figure 2.2). This profile was suggested to be representative of where on court a player "typically" made errors.



PLAYER A % ERROR DISTRIBUTION

7%	1%	1%	**18%**
7%	0%	3%	2%
17%	1%	0%	**31%**
12%	0%	0%	6%

52% Volley Short
24% Volley Long
24% Others

46%	**54%**

34%
66%

16%	**18%**
30%	**36%**

Figure 2.2: Example of a "normalised" (average over 5 matches) error distribution for a player (taken from Murray and Hughes, 2001).

Some observations regarding this data presentation are called for. First, any small sample has the potential of not accurately reflecting the variability inherent in all, or any, of the performances of a player, which in themselves may not display stable properties at all (O'Donoghue, 2004). This could easily be tested using multiple small samples and comparing them for similarity, something not evident in the literature. Other factors of relevance to the likelihood that a sample of 5 matches was representative of a player's error distribution would be the opponent standard sampled i.e.

matches won and lost would likely display different patterns, time within a match i.e. errors may occur at the beginning and end of games more than the middle or form, injury and court conditions.

The second comment relates to the use of the average, in this example of 5 matches, which is provided without reference to the variability between matches. In Figure 1, 36% of errors were in the back right corner. However, it is unknown whether in individual matches the proportion of errors in this area of the court varied very little or a lot. It might be the case for example, that this player in one match made a large number of errors but in the other 4 matches relatively few. The underlying message would be markedly different in this scenario compared to if the error rate was similar for each match. This criticism can be answered by the addition of some measure of dispersion such as the standard deviation. However, given that the standard deviation is not well understood by many people, including some researchers, this thesis has used percentiles (not the same methodology as O'Donoghue, (2005) but with the same desire to display the variability of the data) to present the variability inherent in squash matches (Chapter 4).

The use of equally sized cells was presumably based on the need to make comparisons between cells fair. However, this methodology puts shots played in quite difficult situations (close to the walls) in the same cell as fairly easy shots (closer to the centre of the court). Some rationale for splitting the court floor into cells was presented by court Vučković et al. (2013) and adopted for this thesis.

2.4 Tactics: The search for invariant behaviours (shot types)

McGarry and Franks (1996) analysed the sequence of shots in squash during quarter-finals and beyond of ranking tournaments played in 1987 and 1988. They calculated shot responses e.g. boast, drive etc. according to different complexities of preceding conditions. The simplest, and therefore unlikely to be very informative, was to account only for the opponent's preceding shot e.g. a player's shot response to a boast was tallied, irrespective of where on the court the shot was played or what situation preceded this shot. Using this criterion, shots were typically variant i.e. multiple shot responses were likely, although some degree of invariance was shown for responses to the lob i.e.

shot responses to a lob was fairly predictable. This is logical since a well-placed lob is usually either volleyed straight or is played when the ball is at the back of the court and close to the back corner, hence limiting shot responses. The main finding however, was that at this basic level of analysis, invariance or predictability was generally not found. Consequently, McGarry and Franks attempted more complex analyses by considering not only the opponent's preceding shot but also the subsequent court location for the shot. Many of these situations did not suggest invariant shots responses although some situations for certain players did. This suggests that individual differences did occur for shot selection (providing some justification for the methods employed by Murray and Hughes, 2001). Finally, the most complex analysis of this study involved calculating shot responses for both the opponent's preceding shot and the shot prior to it i.e. the player's preceding shot. The results suggested that invariance was more evident than before although again to a limited degree. This important study was the first to challenge the basic concept of a pattern of play, that is the idea that players have recognisable and repeated aspects to the shots they play. This idea is actually contrary to the nature of expertise since an expert squash player would not wish to make it obvious to an opponent what shot they are going to play simply on the basis of the current situation. However, since experts have knowledge of squash tactics, a high degree of situation awareness (Endsley, 1995) and the ability to play shots accurately, it is not surprising that shot selection is based on seeking to gain advantage by putting the opponent under pressure. It is likely therefore, that a limited number of shot types are optimal for any given situation and that when a player is under some pressure the shot response options would be more limited than when under no pressure. In conclusion, a pattern of play can be hypothesised as likely under certain situations where potential shot responses are limited, and some shot types have a tactical advantage over others.

The analysis of shot responses by McGarry and Franks (1996) yielded good evidence of the complexity of elite sport. It is not surprising that elite squash players do not necessarily respond in the same way to similar shots as if they did they probably wouldn't be elite players. It would be interesting to repeat this study with sub-elite players to calculate the degree of invariance displayed at lower level squash as their limited ability would suggest less shot options and hence greater invariance. One of the players in the McGarry and Franks study was known by his fellow players to be relatively predictable, particularly in the opening game. His tactic was to keep the opponent on the court as long as possible in the first game and did not seem too worried if this cost him the game. This was because he played a ‘power’ game consisting predominately of drives to the back of the court volleying wherever possible to pressurise his opponent. This tactic was very evident in the first game of his matches which often took over 30 minutes to complete (compared to an average of about 10 minutes) and in one instance the first rally lasted some 5 1/2 minutes! These details are well known given his record of not losing a single match in over 5 years of competitions at the highest level. In contrast two other players in the study were well known for being aggressive shot makers playing a variety of shots in an attempt to surprise their opponents. Details such as these only serve to make finding invariant behaviours more difficult especially if the analyses group players together.

The most recent studies to assess squash tactics in elite players presented the different shots played in different areas of the court and for different time constraints (Vučković et al., 2013) with a follow up paper presenting shots played in different areas according to the preceding shot type; these were controlled for the availability of time and involved only matches played by right handed players (Vučković et al., 2014).

Vučković et al. (2013) presented the basic occurrence of shots in 15 different cells of the court (Figure 2.3) which highlighted the tendency to play shots to the back-left corner, supporting the large proportion of backhand drives found by Sanderson and Way (1977) seemingly to apply pressure on the backhand side, concurring with the findings of Hong et al. (1996) and Murray and Hughes (2001) whilst also supporting the idea that players play with a tactical plan (Gréhaigne and Godbout, 1995).

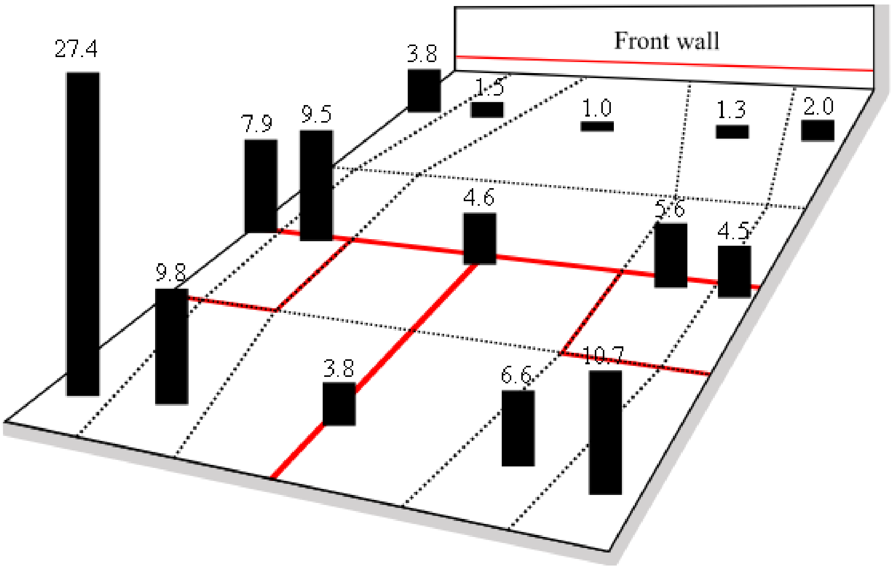


Figure 2.3: Proportionate occurrence of shots played in each area of the court (Vučković et al., 2013)

Vučković et al. (2013) also showed the percentage of different shots types played in each court area suggesting that shots tended to be hit low and hard at the back of the court to return the ball to the back areas whereas when shots were played further forward the drop shot (towards the front) and the lob (to the back) became more prevalent (Figure 2.4).

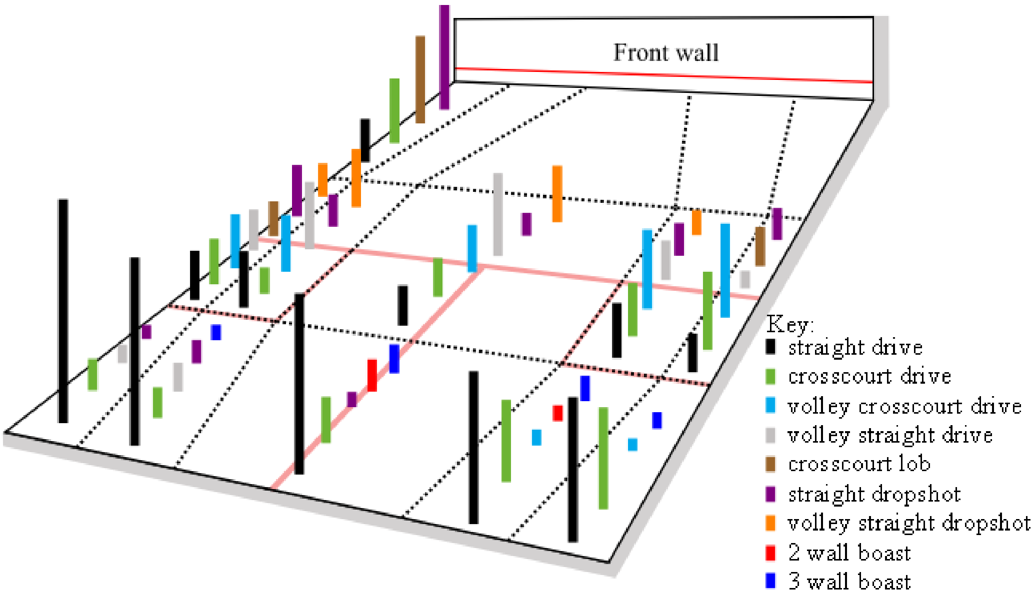


Figure 2.4: Percentage occurrence of shot types played in each area of the court (Vučković et al., 2013)

Vučković et al. (2013) also assessed shot types in response to different amounts of time available to play the shot. It was shown that as the time a player had to play a shot decreased the variation in shots played also decreased. The greatest variations in shot selection were found when shot times exceeded about 1.2s. In agreement with McGarry and Franks (1996) both variant and invariant behaviour were shown to exist in elite squash although for the first time it was suggested that the availability of time to play the shot contributed to the extent to which of these behaviours is evident. In terms of invariant behaviour i.e. a playing pattern, this was found despite the fact that the 10 matches analysed involved 15 different players, suggesting that typical responses could be identified (probably in situations when there was one shot type that was more tactically astute than any other) although this was most common when players had small response times (<1.2s).

The utility of the approach adopted by Vučković et al. (2013) can easily be identified using the example of shots played from the front corner of the court (Figure 2.5). In this situation when time was most limited (< 1.06s) most shots were crosscourt drives or crosscourt lobs (to allow the player time to return to the T). When more time was available the proportion of straight drives increased as did straight drop shots.

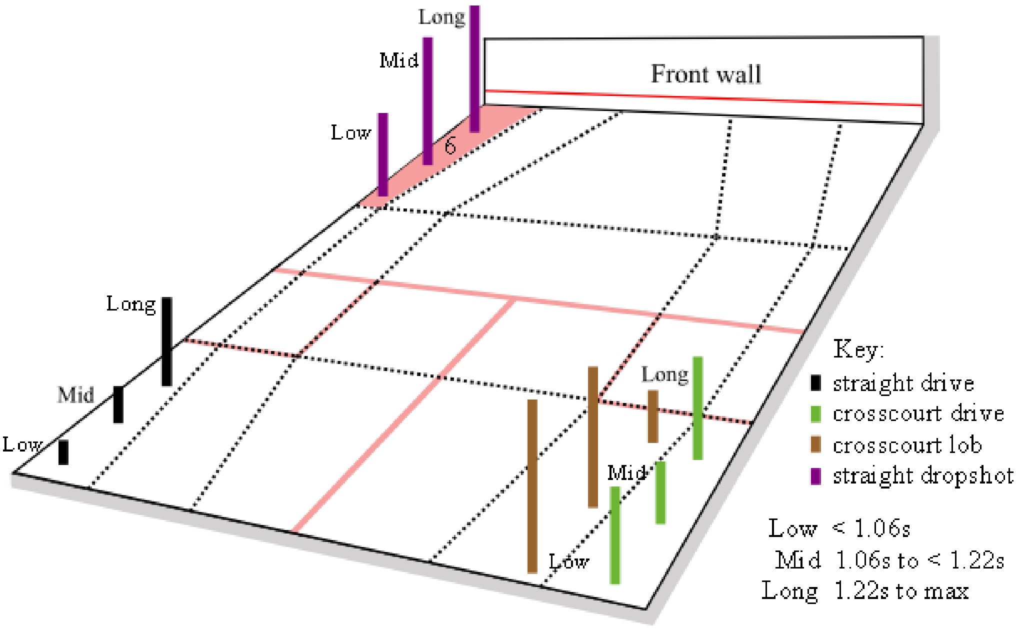


Figure 2.5: Percentage occurrence of shot types for area 6 determined by the time interval in relation to the previous shot (Vučković et al., 2013)

In comparison to the Vučković et al. (2013) paper a similar goal of determining where shots were played to was undertaken by Hughes and Robertson (1998). However, they did not discriminate the availability of time and hence only presented the proportion of shots played (Figure 2.6). Whilst the findings from the two studies were relatively similar, in terms of the overall proportion of shots played to the different areas, the lack of independent variables (in this case availability of time) restricted the usefulness of the findings. Indeed, as suggested by Mackenzie and Cushion (2013), the use of relevant independent variables is crucial if meaningful information is to be derived for use in an applied setting.

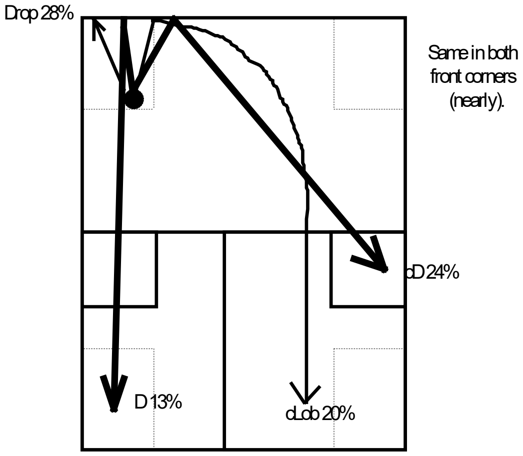


Figure 2.6: Patterns of shots played by elite players (n = 4 matches) at the 1996 British Open (Hughes and Robertson, 1998)

Vučković et al. (2014) determined that shots played more than 1.21s after the preceding shot was struck were excluded from the analysis as these were deemed to give the player so much time that disguise and deception (James and Bradley, 2004) were likely further factors influencing shot selection. Similarly, shots played less than 0.65s after the preceding shot were excluded because they were outliers. Results suggested that in some areas (back corners), two shot responses were most prevalent (in response to straight and crosscourt drives) but as many as seven shot responses were possible (to a volley straight drive played to close to the side wall in the service box). The authors concluded that tight shots (responses from close to the corners of the court) tended to be more predictable (two or three typical shots played) than loose ones (up to seven different shot responses to the same preceding shot when nearer the middle of the court). However there also appeared to be tactical differences based on the side of the court the shot was being played from with a tendency to play the ball towards the back left of the court (to the right-handed player's backhand). This was evident for the back and middle areas of the court only, as shot responses at the front of the court appeared to utilise the diagonal of the court i.e. shots were played to the opposite sides of the front and back of the court. Whilst this paper only considered a small subset of conditions under which shots were played, the results tended to suggest that other conditions should be analysed to determine how well tactics can be determined. The authors suggested that individual player profiles may be determined, which may be opponent specific. Also winning and losing performances could also be analysed to assess tactical profiles in favourable and unfavourable situations.

2.6 The analysis of complex systems

Performance analysis has emerged as a discipline within sports science over the last 40 years or so. Its origins lie within sport related degree programmes where lecturers from varied backgrounds adapted fairly rudimentary techniques to gather sports data, analyse it and present their findings. The notational analysis techniques have changed little over the years although sophisticated data capture techniques have emerged to collect movement data (e.g. GPS, computer vision, sensors) many times per second, physiological data (e.g. heart rate monitors) and performance data (e.g. pattern recognition algorithms).

The abundance of data collected from sporting events has led to the emergence of data analytics which originates in computer science where techniques for analysing large data sets have been utilised and modified for over 50 years. This form of data analysis enables the discovery of complex patterns within these large data sets by considering data as a sequence of related events. This approach contrasts with the many performance analysis research studies which have utilised a reductionist approach (Mackenzie and Cushion, 2013) whereby variables are analysed in isolation. For example, frequency counts of individual actions (e.g. Vučković et al., 2014) can identify general patterns such as the preponderance of shots played in the back corners of the squash court, but they do not allow more complex explanations for why these patterns occur. In squash some decision-making behaviour is required to decide where to hit a shot, generally to one of the four corners of the court. This decision is made on the basis of many different variables such as where the opponent is and the difficulty of playing a shot as well as historical knowledge of previously played similar shots (referred to as situation awareness; Endsley, 1995). It makes sense, therefore, that in order to try to understand a complex sport such as squash better, more advanced and more complex analyses are needed to connect the different variables as they change through the evolution of a rally.

One alternative methodological approach is known as the dynamical systems perspective (e.g. McGarry, Anderson, Wallace, Hughes and Franks, 2002) which has been used to analyse sports including squash (e.g. McGarry, Khan and Franks, 1999). This approach takes a more holistic view and considers the sequential dependency of different events whilst recognising that one action may result in relatively large changes to a system (the interaction between the two players) as a whole. Dynamical systems theory describes how behaviours can deviate through a series of states (stable or unstable) before returning to an original stable state (Kelso, 1995). This theory originated with the work of Haken (1983) and is commonly exemplified in the experiments by Haken, Kelso and Bunz (1985) who showed how the voluntary oscillations of the two index fingers changed from an out of phase to an in-phase relationship simply by increasing the speed of the oscillations. This change in behaviour occurred at a critical oscillation velocity and was suggested to be evidence of self-organising behaviour i.e. the change in the relationship of the oscillations of the two fingers was not seemingly controlled by the brain. McGarry et al. (1999) suggested that squash performance exhibited properties akin to a dynamical system as experts were able to reliably identify stable and unstable periods within squash match play. They further hypothesised that the transition between these two states could be as a consequence of behavioural perturbations. Semantically, they proposed that squash match play could be described as mono stable (stable is the usual situation) but displays variability through transient instability (not long lasting) caused by system perturbations. This approach has the obvious methodological advantage of determining aspects of play which have critical effects on the outcome of rallies i.e. the precedents to winners and errors. However, this methodology is still relatively novel and methods for determining these critical aspects not proven. Whilst McGarry et al. (1999) validated perturbations using experts to agree the point at which a change in game state occurred, there have been few attempts to identify perturbations from game data. One example using data from elite squash matches, Roddy, Lamb and Worsfold (2014) defined a perturbation as occurring when an opponent lost control of the rally and was out of position prior to losing the rally. being different to an unreturnable shot, a let/stroke, an unforced error or a miss-hit. Thus, a rally ending shot would not be considered a perturbation even though, as they concur, the shot perturbed the stability of the rally. Their reasoning was that the winning shot did not coincide with the opponent losing control of the rally i.e. scrambling or over stretching. This seems to suggest that a good shot which makes the opponent struggle to return the ball is deemed a perturbation but a very good shot which is unreturnable is not defined as a perturbation. The underlying premise that forms the basis of this methodology is that a perturbation occurs at the point when the system state changes between stable and unstable rally or match situations. On this basis a fundamental requirement is the definitions of what constitute stable and unstable situations

Hypothetically, this may be possible using player movement variables, shot characteristics or distances between players and between the ball and players. To date this has not been achieved although the methods used to find patterns in data has been used in search for other patterns. For example, Borrie, Jonsson and Magnusson (2002) used software called Theme to identify repeated patterns (called T-patterns; Magnusson, 2000) from within complex data streams. A T-pattern was defined as “a combination of events in which the same events occur in the same order with the real-time differences between consecutive pattern components remaining relatively invariant”. Other techniques for discerning patterns in data utilise data analysis procedures formulated in computer science machine learning or data mining algorithms. These can be supervised where input and output data are used to train the model to predict future events or unsupervised where patterns are discerned through repeated parses of the data e.g. cluster analysis. This thesis will utilise an unsupervised data mining algorithm, namely a two-step cluster analysis using a probability based log-likelihood distance measure, to identify groups within movement and shot type data. It is thought that the use of algorithms designed to find patterns in large data sets will be the next progression in the academic and applied worlds of performance analysis.

2.7 Conclusions

The review of literature has focussed on the analysis of squash match play with little reference to physiological, psychological or biomechanical factors. Whilst these aspects are obviously important for understanding squash performance, the depth and breadth of scientific endeavour under the guise of notational analysis is sufficiently advanced as to make a PhD thesis in its entirety feasible.

Whilst there has been a relatively large number of research studies focussing on the notational analysis of squash much of this has been published in conference proceedings or edited books, meaning that the highest levels of peer review, as for highly rated International journals, has usually not been afforded these studies. As a consequence, many studies have used similar methods and described results in similar, and often flawed ways. This review has focused on critically reflecting on these methods rather than laboriously restating the research findings.

The most pertinent limitation in the extant literature was deemed to be the lack of description of the variability in the data. Often mean values were presented without any reference to any measure of dispersion. A second fundamental weakness was the lack of use of independent variables in some studies. Consequently, factors that were likely to affect performance e.g. time available to play shot, were not accounted for. The third main limitation found in the previous literature pertained to the lack of sophistication in terms of the analysis techniques. Most previous research failed to adequately account for the complexity that is evident in elite squash match play. Whilst this is not easy e.g. computer science algorithms are unintelligible for most non-computer scientists, some progress has been made, including this thesis.

This thesis has therefore sought to develop new methods for analysing squash based on the perceived shortcomings evident in previous research with the explicit aim of publishing the work in the top rated international journals. This has been achieved for studies 2 and 3 with some of the initial reliability testing also published in two high impact journals. The final study is being prepared for journal submission at the time of thesis submission.

Chapter 3: The reliability and accuracy of inputting shot and court area information into SAGIT/Squash and processing the information in Excel and SPSS

This chapter has contributed to the following publications in relation to their reliability and methodology sections: -

5. Vučković, G., James, N., Hughes, M., Murray, S., Milanović, Z., Perš, J. & Sporiš, G. (2014). A New Method for Assessing Squash Tactics Using 15 Court Areas for Ball Locations. *Human Movement Science*, 34, 81-90.
6. Vučković, G., James, N., Hughes, M., Murray, S.R., Sporiš, G. & Perš, J. (2013). The effect of court location and available time on the tactical shot selection of elite squash players. *Journal of Sports Science and Medicine,* 12, 66-73.

3.1 Abstract

Reliability tests are typically undertaken as a way of determining limitations of both the coding structure used to record match information and the understanding and ability for the analyst(s) to use the codes appropriately. This assessment should be at the level of the subsequent analysis (Hughes, Cooper and Nevill, 2002). Of equal importance is the verification of the accuracy of the data prior to undertaking any analysis along with the need to process the data to enable appropriate analyses to take place. Squash matches were filmed using a fixed overhead camera and images processed in SAGIT/Squash to semi-automatically calculate player movement information as well as allow an operator to manually input shot information. The two data streams were synchronised in Matlab before reliability and accuracy testing. Good levels of reliability were found for all court locations and shot information (agreement > 90%) although when an operator coded a long match without a break the variables with the lowest percentage agreement (shot type and foot used) had less than 90% agreement for the second half of coding presumably due to fatigue effects. Error testing, using a series of queries, specific to each data type, following data collection and prior to data analysis, discovered multiple errors in the data due to failing to code events and coding events incorrectly

which had the consequence of producing errors in other variables. This under-reported methodology was deemed critical for the possibility of valid data analysis.

3.2 Introduction

This thesis was concerned with recording and analysing the movements of squash players, where they played shots and the outcomes of these shots. Fundamental to this was the accuracy and reliability of the data collection system, originally devised for indoor team sports (Perš et al., 2002) and later adapted for squash (SAGIT/Squash; Vučković et al., 2003) to record all desired variables (Atkinson and Nevill, 1998) as inaccurate data would lead to erroneous findings when analyses were undertaken. This chapter thus brings together the analyses undertaken within the PhD study to test the accuracy of the data collection methods.

Reliability testing initially took place using data originally collected prior to the start of the thesis, using the same SAGIT/Squash system, which was still being analysed for publication by the supervisors of this Phd study. Some reliability tests had already been carried out on the system prior to the comencement of the PhD and are thus only referred to in this Section. Other new tests were carried out by the PhD student reanalysing the old data and presented here with publications referred to. Finally further tests were carried out using new data, collected for this thesis, and presented in full in this Chapter.

The camera placement and methodology for getting the video images into SAGIT/Squash for automatic processing with operator supervision has been well documented in Vučković et al. (2009) and the reliability associated with the resultant calculations of distance and speed for each player were published by Vučković et al. (2010). The exact camera location for the overhead camera (both vertically and horizontally) was not deemed important, as subsequent calibration for image capture accounted for its location.

This chapter therefore initially focusses on how reliably the area of the court the ball was hit from (published in Vučković et al., 2014) and what shot was played (published in Vučković et al., 2013) could be determined. When these issues were being considered it was clear that one issue that had seemingly not been debated in the literature was the dimension and shape of the cells used for the shot locations. Previous researchers have consistently used 16 rectangular cells of equal dimension with no discussion of their appropriateness (e.g. Murray and Hughes, 2001). Two issues were considered of relevance to the appropriateness of these cell dimensions. Firstly the location of the shot may be more critical near the sidewalls than in the centre of the court. Secondly straight and crosscourt shots tend to have different trajectories, particularly when the sidewalls are hit. Consequently, a case for an alternative method of dividing the court into cells, which takes into consideration squash tactics, was made.

All shot information was manually added to the software at the point when the ball was hit to synchronise shot and player information. The method for inputting this information originated with the operator manually (mouse click) specifiying the ball contact point on the SAGIT/Squash screen, adding approximate ball height information (to assist software algorithm calculation of court cell) and then adding various shot information using drop down selection options. James, Jones and Hollely (2002) suggested three sources of error for this type of manual coding of events; Operational error: where the observer presses the wrong button to label an event, Observational errors: the observer fails to code an event, and Definitional errors: the observer labels an event inappropriately. Observer error in this study was likely due to the wrong shot type or an incorrect ball location being entered into SAGIT/Squash.

Many research papers have identified good practice for conducting reliability tests which can be undertaken using either one analyst coding the same events twice (intra-operator) or two analysts coding the same events (inter-operator). Both tests can give a good indication of analyst accuracy although the inter-observer test has the advantage of detecting operational definition misinterpretations more fully since one analyst (intra-operator) can consistently misapply an operational definition and the test will suggest good accuracy (James, Taylor, & Stanley, 2007). Reliability tests are typically undertaken as a means to determine limitations of both the coding structure used to record match information and the understanding and ability for the analyst(s) to use the codes appropriately. One pertinent issue related to this, and not seemingly considered in many publications (see Pulling and Stenning, 2015 for an exception), is the suggestion that a reliability assessment should be at the level of the subsequent analysis (Hughes, Cooper, & Nevill, 2002). This suggests that a simple value for a reliability test which considers all data collection but fails to discriminate whether any one aspect of the data collection can be reliably coded is insufficient and not fit for purpose. Since this PhD introduced some new shot information previously not used in SAGIT/Squash e.g. foot used referred to the placement of the feet at the time of playing a shot, and as the PhD student was also unfamiliar with the system prior to commencing the study, it was determined that both inter- and intra-operator reliability tests were needed to assess whether each aspect of data collection could be reliably coded.

Whilst reliability tests have been widely used in performance analysis publications since a special edition of the International Journal of Performance Analysis in Sport on reliability (Volume 7, Number 1, January 2007), there has been little mention of checking data sets for errors prior to conducting any analysis. However, complex analysis systems, particularly when lots of data are collected, will inevitably contain errors. However, the extant literature does not tend to either confirm the use of error checking or specify methods for undertaking this procedure. The methods used in this study will therefore be presented.

The aim of this study was to determine the accuracy of event coding using the SAGIT/Squash (updated version named Tracker) software.

3.3 Methods

3.3.1 Sample of matches and participants

Matches were recorded at the World Team Championships (n = 11), the Slovenian National Championships (n = 11) and during a local tournament (n = 15), all played in 2003. One match was randomly selected from each tournament (12 games in total) for a reliability analysis. These games were viewed for a second time and the shot locations and shot types of 2907 shots recorded in an Excel spread sheet. These were compared with those calculated in SAGIT/Squash.

Reliability testing was also undertaken on a selection of matches analysed at the 2010 (n = 14) and 2011 (n = 27) Rowe British Grand Prix, held in Manchester, UK for data analysis in Chapters

4, 5 and 6 in this thesis. Data accuracy and processing procedures were also utilised on this data. Ethics approval for the study was provided by a University's ethics committee (Appendix 3.1).

3.3.2 Procedure

3.3.2.1 Court area

Matches took place on a court set up with a PAL video camera (JBL UTC – A6000H, Korea) attached to the ceiling above the central part of the court and a similar camera located on a tripod at ground level behind the court. The SAGIT/Squash system has a separate input system designed to allow the operator to view the video taken from the overhead camera and the ball could be marked on the court via a touch sensitive interface. To improve the accuracy of this input a separate video recording, taken from behind the court, was used to enable the operator to estimate the height of the ball above the floor at the time of the stroke. This measure was input into SAGIT/Squash so that the software could make a small correction for any perspective distortion caused by the ceiling-mounted camera not always being directly above the ball location. This height information was only used to decrease the perspective error and as such even a relatively coarse estimate by the operator was sufficient to improve, albeit slightly, the software's estimate of ball location. The software then calculated the x and y coordinates of the ball location and then assigned this to one of 15 areas. The logic behind the configuration of these cells was based on two principles. Anecdotally people knowledgeable about squash would suggest that elite players tend to hit the ball closer to the sidewalls more often than less elite players. The data collected using SAGIT/Squash tended to confirm this (Figure 3.1).

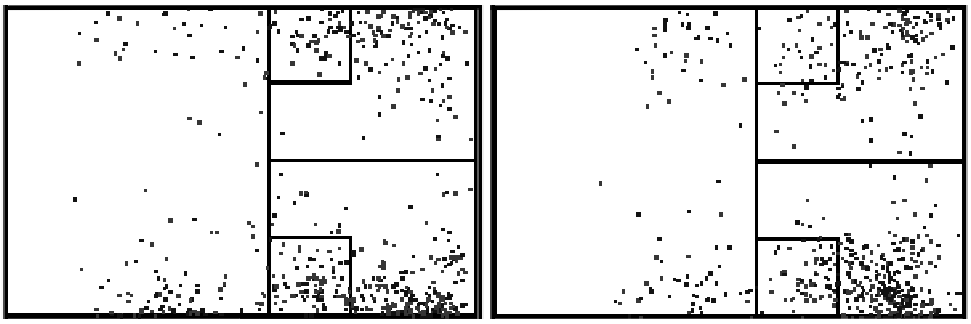


Figure 3.1: Example shot location distributions for an elite (left) and National (right) level game

The reason for this is simple; a ball hit from very close to the sidewall is more difficult than one hit just a short distance away. Elite players are more capable of hitting the ball into the most difficult court areas and tend not to play shots from anywhere near the middle of the court as they are relatively easy to play and thus avoided. On this basis it seems logical that cell dimensions near the sides of the court are far more critical than central areas and hence the area of the cells should reflect this. The second principle considered was the observation that the ball bounces differently when it hits the sidewall and using this sidewall bounce is a deliberate tactic in elite squash. For example, a crosscourt shot from the back of the court to the opposite side of the back of the court is often aimed to hit the sidewall somewhere near to the service box to prevent the opponent volleying the ball. This is tactically astute but means that the trajectory of the ball tends to finish further away from the sidewall the nearer the ball gets to the back wall. A similar observation is made at the front of the court where drop shots are aimed at the nick (the join between the floor and the sidewall) as the ball tends to bounce twice relatively quickly (giving the opponent less time) but also the trajectory changes towards the centre of the court. On this basis we considered that cells should not be rectangular in the front and back of the court but should represent typical ball trajectories for these areas. The consequent areas used are therefore suggested to distinguish between shots that were played close to the sidewalls (1, 5, 6, 10, 11 and 15) shots that were played from similar positions but not close to the sidewalls and from the middle of the court (areas 3, 8 and 15).

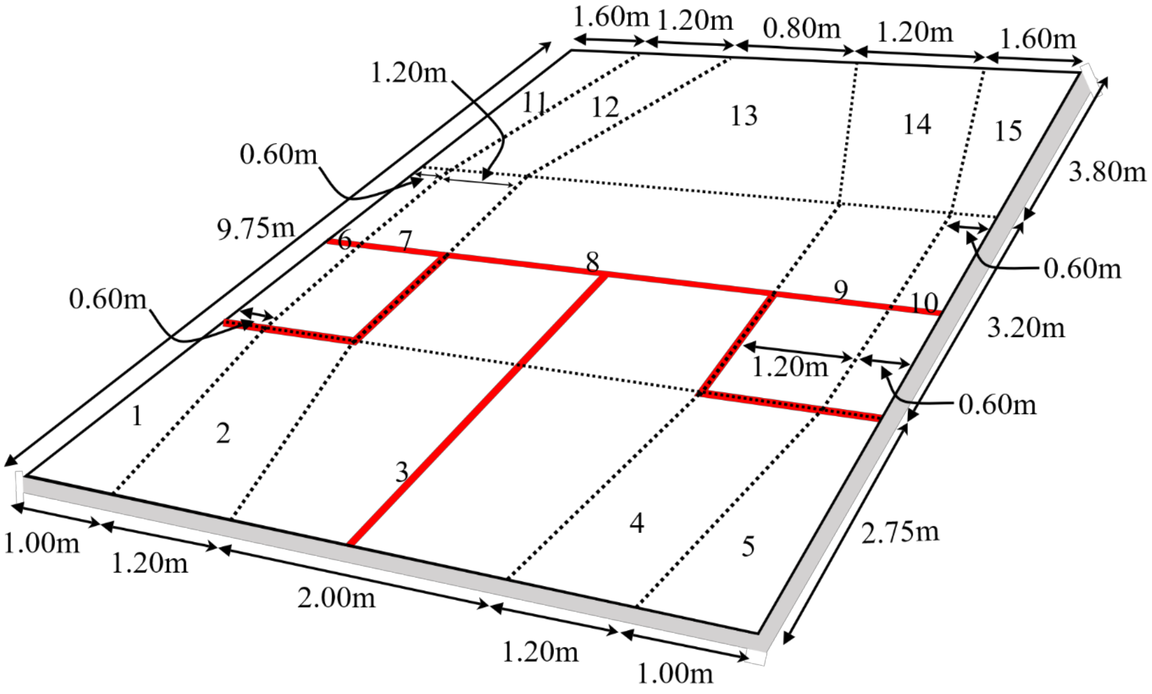


Figure 3.2: Dimensions of the court floor divided into 15 areas

3.3.2.2 Shot information

Once all matches had been coded in SAGIT/Squash each computer file was imported into Matlab (v2011a, The Mathworks Inc.) to filter and summate the player movement data (collected 25 times per second) to synchronise to the time at which each shot was played (input by operator, as described above). This had the effect of reducing the data set to some 4 million data cells (over 40,000 rows i.e. number of shots, and over 100 columns i.e. different information associated with each shot) meaning that visual examination of the raw data to check for errors was impossible. Instead, specific logic was used to identify errors and process the data (create new variables for analysis purposes) for each variable used in each study using spreadsheets (Excel v2010, Microsoft Inc., Redmond, USA) or IBM SPSS Statistics package (v19, IBM Corp., New York, USA).

3.4 Results

3.4.1 Court area

In order to calculate the reliability for determining the ball location accurately two separate measures were compared. The first was the original SAGIT/Squash derived court area (calculated within the software from the input of the ball location at the point of contact) and the second involved the operator viewing the overhead video images and inputting the court locations and shot type for 2907 shots, undertaken 8 years after the initial data entry (operator input). The operator had played squash at an International level, was an International coach and had over 10 years' experience of squash analysis. It was important that the operator was as accurate as possible otherwise a low reliability result could have been achieved due to operator error whereas we wanted to assess the reliability of SAGIT/Squash. The comparison between the two measures of data capture (SAGIT and operator) was thought to be adequate to highlight all possible errors of data input with low errors suggesting that the methodology was appropriate. The two assessments of the correct area (SAGIT and operator) were then compared and 7 obvious errors identified, since the two areas were not adjacent. These errors were attributed to the operator entering the wrong area number as the location had been entered on the wrong side of the court (operational error), not likely when clicking on the image of the ball as during the data input into SAGIT/Squash. These errors were rectified before undertaking Kappa and percentage agreement (number of agreed locations divided by the number of disagreements) calculations. These resulted in an overall Kappa value of 0.869 and percentage agreement of 88.90% for the court location data and a percentage agreement of 99.52% for shot type data. Whilst the reliability value for shot type was deemed acceptable it was felt that an overall reliability statistic for the court areas could potentially hide problematical locations on the court that could yield unreliable information regarding area location. This could be a problem in that the true cell location could be entered incorrectly due to the proximity of the ball being on or close to the cell boundary. Both SAGIT/Squash and the observer on the reliability test can make this type of observational error. For the reliability test when both agree the true location is assumed and when there is a disagreement the reliability is considered compromised. Of course, when a court has lots of cells the frequency of these disagreements increases as the probability of the ball being on a boundary increases. To ascertain the extent to which this was a problem further assessment was carried out to consider areas in pairs. For example, the border between areas 1 and 2 (see Figure 3.3) could be problematical for discerning whether the ball was in area 1 or area 2. The extent to which this, and for all other borders, was a problem was ascertained by carrying out Kappa and percentage agreement calculations for each border where the total number of agreements for both areas were used to compare against the number of times one method recorded one area and the other method recorded the other area (two possible situations), data presented in Appendix 3.2 and results in Figure 3.3.

Figure 3.3: Reliability statistics for between area reliability

The possibility of confusion for two adjacent areas joined by a corner e.g. cells 1 and 7 also existed. Calculations for all 16 of these adjacent areas revealed Kappa values of at least 0.98 and percentage agreement of at least 99.45%.

3.4.2 Shot information (type, side, court side and foot used)

Inter-operator reliability tests compared the codes assigned to 1124 shots (from one randomly selected game) by Murray (PhD student) and Vučković (supervisor) for shot type (97.03 % agreement, Kappa = 0.97; Appendix 3.3.1), shot side (95.90 % agreement, Kappa = 0.96; Appendix 3.3.2), court side (97.07 % agreement, Kappa = 0.97; Appendix 3.3.3) and foot used (90.24 % agreement, Kappa = 0.91; Appendix 3.3.4).

Intra-operator reliability tests compared the codes assigned to 848 shots (from one randomly selected game) by Murray (PhD student) on two occasions, separated by 6 weeks to eliminate memory effects, for shot type (97.44 % agreement, Kappa = 0.97; Appendix 3.4.1), shot side (98.74 % agreement, Kappa = 0.99; Appendix 3.4.2), court side (98.62 % agreement, Kappa = 0.97; Appendix 3.4.3) and foot used (94.76 % agreement, Kappa = 0.92; Appendix 3.4.4).

To assess changes in coding methodology 5 matches spanning the whole coding process, which itself spanned 18 months, were re-coded by Murray (PhD student) 6 months after completing the original coding process. Shot type, shot side and court side has minimum percentage agreement of over 95% whereas foot used had less than 95% agreement for 3 out of 5 matches (Table 3.1).

Table 3.1: Intra-operator percentage error reliability for longitudinal coding methodology changes

	s1 (match 1)	s2 (match 11)	s3 (match 21)	s4 (match 31)	s5 (match 41)	Average	Min	Max	SD
Shot Type Agreement	96.99%	97.37%	95.24%	97.67%	97.69%	96.99%	95.24%	97.69%	0.0091
BH or FH Agreement	97.76%	99.03%	99.20%	98.46%	98.86%	98.66%	97.76%	99.20%	0.0051
Side of Court Agreement	99.26%	98.37%	98.39%	98.46%	98.86%	98.67%	98.37%	99.26%	0.0035
Foot Used Agreement	95.42%	93.52%	94.12%	93.55%	97.69%	94.86%	93.52%	97.69%	0.0157

To assess the effect of coding a long match without a break the longest match (83 minutes) was re-coded by Murray (PhD student) in one session, 6 months after completing the original coding process. Less than 95% agreement were achieved for shot type and foot used for games 4 and 5 (Table 3.2).

Table 3.2: Intra-operator percentage error reliability for longitudinal coding methodology changes

	Game 1	Game 2	Game 3	Game 4	Game 5	Average			
	N=357	N=379	N=185	N=539	N=309	N=1769	Min	Max	SD
Shot Type Agreement	99.44%	98.66%	96.65%	91.75%	93.45%	95.99%	91.75%	99.44%	0.0296
BH or FH Agreement	99.72%	98.93%	97.79%	99.25%	98.69%	98.88%	97.79%	99.72%	0.0064
Side of Court Agreement	99.15%	98.93%	98.91%	99.25%	99.35%	99.12%	98.91%	99.35%	0.0017
Foot Used Agreement	96.82%	96.73%	95.48%	90.43%	92.71%	94.43%	90.43%	96.82%	0.0249
Average	98.78%	98.32%	97.21%	95.17%	96.05%	97.10%			

Shot type and foot used for games 4 and 5 were split into 1st and 2nd halves and re-analysed with both demonstrating fatigue effects (Table 3.3).

Table 3.3: Intra-operator percentage error reliability for longitudinal coding methodology changes

Game 4 (n=538)	1st Half (n=269)	2nd Half (n=269)	Chi Square
Shot Type	93.68 % diff.	89.75% diff.	p = 2.14
Foot Used	97.33 % diff.	82.53% diff.	p = 25.39

3.4.3 Error checking and data processing prior to analysis

Error checking performance data had to be specific to each data type. For example, distances moved by players between shots could not have values less than 0 or greater than 7 metres. Visual examination using boxplots could identify if these values were present, and if so could be identified as SPSS specifies the case number for each outlier or unusual value (Figure 3.4).

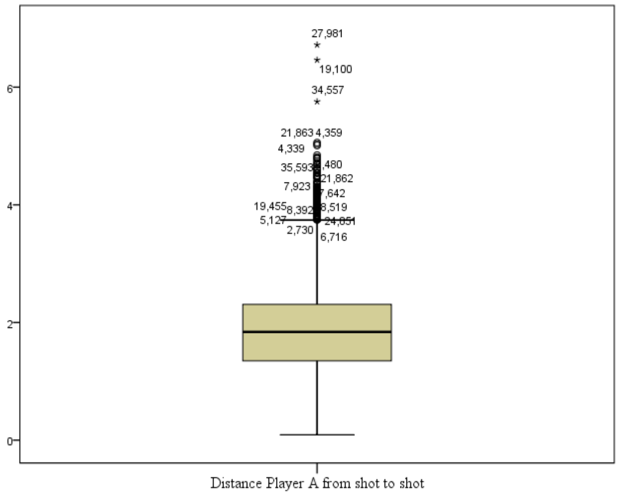


Figure 3.4: Box plot to identify errors for player distance covered

For rally ending outcomes (stroke, let, winner, error) the following shot type had to be a serve. Consequently, crosstabs in SPSS (select cases used to only select serves for demonstration purposes) produced the following that identified two errors (serve followed play continues, Table 3.4).

Table 3.4: Frequency of serves following different shot outcomes

		Following shot type Serve	Total
Preceding shot outcome	Error	1017	1017
	Let	300	300
	Play Continues	2	2
	Stroke	7	7
	Winner location	1078	1078
Total		2404	2404

The most commonly used data processing procedure was the use of functions in Excel to amend the data into a format suitable for specific analysis tests. For example, the output from Matlab produced the time between shots but not rally time. This had to be calculated in Excel using shot time and rally number using an IF function to add up the values correctly.

3.5 Discussion

Many performance analysis researchers refer to reliability testing, although even in good quality journal publications, the details of such testing are often not given. Perhaps, this is because, for some, reliability testing is a necessary means to an end, and of little interest to either the researchers or the readers of the papers. This is understandable given that the real purpose of most studies, except ones that explicitly aim to determine the reliability and accuracy of a measurement device, have little to do with the reliability of the measurements.

Reliability in this Chapter was deemed to be both informative and essential for the training of the researcher and the evolution of SAGIT/Squash. Ultimately everything in this Chapter had the simple goal of ensuring that the data capture and handling procedures produced accurate data and that all subsequent analyses were appropriate to the accuracy of that data. The first major finding was the appropriateness of the 15 court cells used in SAGIT/Squash. The initial system had used 21 cells (the areas closest to the side walls were split into 2 equal size areas, with the division parallel to the side wall). Reliability testing proved these 12 areas to be unreliable and hence adjustments to the court cells made resulting in the good reliability presented here. The other reliability tests confirmed the ability of the researchers to use the system appropriately with the caveat that trying to code for hours on end was irrational. This was because the frequency of mistakes increased and hence made the next stage of error checking more arduous than necessary. However, even under the most favourable conditions, it was recognised that mistakes were inevitable since the error checking procedures discovered multiple errors in the data due to failing to code events or coding events incorrectly. These errors had the consequence of producing errors in other variables and for this reason, the error checking procedures, briefly mentioned here, were deemed crucial for eliminating the majority of errors from the data.

3.6 Conclusion

Reliability testing, whilst relatively unexciting in comparison to the findings of the research, proved to be a vital component of determining the limitations of the data collection process. When these had been identified by the reliability results amendments were made prior to data collection for the thesis. However, even following this procedure error checking post data collection and processing revealed multiple errors in the data. A series of queries, specific to each data type, were implemented to find the errors and rectify them. This under-reported methodology was deemed critical for the possibility of valid data analysis.

Chapter 4: Effects of rule changes on physical demands and shot characteristics of elite-standard men's squash and implications for training

This chapter has been published with the following reference:-

Murray, S., James, N., Hughes, M.D., Perš, J., Mandeljc, R. & Vučković, G. (2016). Effects of rule changes on physical demands and shot characteristics of elite-standard men's squash and implications for training. *Journal of Sports Sciences*, 10, 2, 129-140.

4.1 Abstract

The physical demands and rally characteristics of elite-standard men's squash have not been well documented since recent rule changes (scoring and tin height). This information is needed to design optimal training drills for physical conditioning, provided here based on an analysis of movement and shot information. Matches at the 2010 (n = 14) and 2011 (n = 27) Rowe British Grand Prix were analysed. Rallies were split into four ball-in-play duration categories using the 25th (short), 75th (medium), 95th percentiles (long) and maximum values. Cohen's d and Chi squared tests of independence evaluated effects of rally and rule changes on patterns of play. The proportion of long, middle and short shots was related to the duration of the rally with more shots played in the middle and front of the court in short rallies (phi = 0.12). The frequencies of shots played from different areas of the court have not changed after the adoption of new rules but there is less time available to return shots that reflects the attacking nature of match play for elite-standard men players. Aspiring and current elite-standard players need to condition themselves to improve their ability to cope with these demands using the ghosting patterns presented that mimic demands of modern match play.

4.2 Introduction

Specific training and practice is necessary to condition athletes optimally for performance (Reilly, Morris, & Whyte, 2009). Hence, there is a need to improve understanding of match characteristics of a sport at the standard of participation (Murray & Hughes, 2001). In squash, previous research has identified demands of match play at different playing standards, although changes to the scoring system and tin height could have altered patterns of play.

Squash was first analysed by Sanderson and Way (1977) using hand notation to record the frequency and distribution of winning shots and errors. Hughes (1985) computerised this system and identified tactical differences among club-, county-, and national-standard players, partially attributed to different movement capabilities. For elite-standard men's squash, Hughes and Robertson (1998) described typical match characteristics (e.g. rallies had a mean duration of 21 s), using a sample of five matches that involved players ranked in the world's top 20. While this provided detailed information of the matches analysed, the usefulness of simple means for training purposes was limited, particularly since these types of data tend to be non-normal in their distributions.

Girard, Chevalier, Habrard and Millet (2007) presented rally durations in 3 s intervals up to 24 s, then 6 and 10 s intervals followed by all other rallies grouped for durations over 40 s. These time intervals were selected from a physiological perspective, but this might not be ideal from tactical and training perspectives. Similarly, Vučković and James (2010) used four categories (0 to 3.9 s, 4 to 11.9, 12 to 24.9 and 25 and over) but, for training, the first category was too short and the last was too long.

Player movements were first analysed using a manual tracking system on a computerised digitisation pad to assess speed, accelerations, and distances (Hughes & Franks, 1994). More recently, a reliable semi-automated computer vision tracking system for squash (Vučković, Perš, James, & Hughes, 2010) was developed. The SAGIT/Squash system was initially used to assess movement in the 'T' area of the court (Vučković, Perš, James, & Hughes, 2009). Winning players spent a greater proportion of total playing duration in the T area than losers. However individual match analysis has been suggested as not being the most appropriate measure for determining differences in performance between winners and losers because it is often the case, particularly in close matches, that the losing player wins a high proportion (nearly 50%) of the rallies (Vučković, Dežman, Erčulj, Kovačič, & Perš, 2004).

Vučković et al. (2014) used a new squash-specific method for categorising court locations in which the ball was played to present typical shots responses for elite-standard players. These responses depended on position on court and the duration between shots (Vučković et al., 2013). The studies used squash matches played under the 9 point-on-serve (POS to 9) rules with a 48.3 cm high tin (line on front wall that is out of play). In 2009, the World Squash Federation and the Professional Squash Association aligned to standardise all professional men's squash matches to play to 11 point-per-rally (PPR to 11) with a 43.2 cm tin. A comparison of elite-standard squash matches played under the two systems found the number of rallies had reduced from a median of 34 (IQR = 15) to 20 (IQR = 8) although the duration of rallies had not changed (Murray, James, Dineen, Hughes, & Vučković, 2013). Mean match duration, distance covered and speed had also reduced under the new system although these results were based on a small sample size (10 matches under the new rules).

The aim of this paper was to present general match and physical characteristics for PPR to 11 squash (43.2 cm tin) and more detailed rally information such that specific training could be devised. This included presenting more informative descriptive statistics than just measures of central tendency and dispersion, to improve the specification of appropriate training.

4.3 Methods

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 (Perš, Kristan, Perše, & Kovačič, 2008) that is a newer version of the SAGIT/Squash software (Vučković et al. 2009). 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. A further 11 matches with players ranked in the top 16 in the world were analysed to obtain POS to 9 comparison data (as used in Vučković et al. 2009). Ethics approval for the study was provided by the health studies sub-committee of Middlesex University's ethics committee (Appendix 3.1). No external agencies were involved with data collection, analysis or interpretation and have no rights regarding the publication of this research.

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 to make all of the floor plus some of the walls visible. 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 (Section 3.3.2).

General match information and shot distributions were calculated to facilitate comparison with previous research on matches that used the old scoring system (data from Vučković et al., 2013; Vučković et al., 2014). This included both game and rally information to improve a common training routine called ghosting i.e. players imitate rally movements without striking a ball as a solo drill. Rallies needed to be categorised according to duration, movement locations and physiological demand e.g. exercise-to-rest ratios. Previously presented rally duration intervals of 3 (Girard et al., 2007) and 4 s (Vučković & James, 2010) were considered too short for training purposes and longer rally durations had not been considered fully (Vučković & James, 2010). The distribution for rally durations was positively skewed so rallies were split into four categories using the 25th (short), 75th (medium), 95th (long) percentiles and maximum values (very long) as the upper values for each rally duration category. Match characteristics were then calculated to inform the prescription of ghosting schedules. This analysis resulted in matches being categorised according to the World rankings of the players as this was related to match duration.

Statistical analysis was performed using IBM SPSS software (version 21.0; SPSS Inc., IL). The data were assessed for normality (Shapiro-Wilks' test) and the skewed distributions specified that the median and interquartile range were used to describe them. Game and rally duration, number of rallies, player distance and speed and shot distributions were calculated for rallies categorised by their duration. Cohen's d (Cohen, 1988) was used to assess the magnitude of differences between distances covered by rally winners and losers, evaluated as trivial (0-0.19), small (0.20-0.49), medium (0.50-0.79) and large (0.80 and greater) (Winter, Abt, & Nevill, 2014). Chi squared tests of independence tested whether the proportion of long (to the back of the court), middle and short (front) shots were related to the duration of the rally and whether the distribution of shots had changed under the new rules. Statistical significance was set at $p < 0.05$.

4.4 Results

Games played under the PPR to 11 rules have reduced in length (median = 11 min 37 s) compared with POS to 9 (Table 4.1) as there were typically fewer rallies per game (median = 21) and hence less distance covered by players. However, game duration varied between 4 and 32 min. Individual rally characteristics have changed slightly with more shots being played in shorter duration than POS to 9.

Table 4.1: Game and rally statistics for World ranked male squash

	POS to 9		PPR to 11	
	Median	*IQR*	Median	*IQR*
Game				
duration	15min 45 s	*8min 26 s*	11min 37 s	*6min 51 s*
ball in play	54.4%	*9.0%*	51.4%	*11.4%*
distance travelled	1054 m	*543*	496.3 m	*292.6*
rallies per game	34	*15*	21	*8*
Rally				
duration	15.0 s	*5.4*	13.2 s	*15.7*
shots	11	*16*	13	*19*
distance travelled	22.1 m	*31.2*	18.8 m	*24.2*
speed	1.5 m/s	*0.1*	1.4 m/s	*0.3*

The frequencies of shots played from the different areas of the court were trivially different between POS to 9 and PPR to 11 (chi-square = 269.98, df = 14, p < .001; phi = 0.08; Figure 4.1).

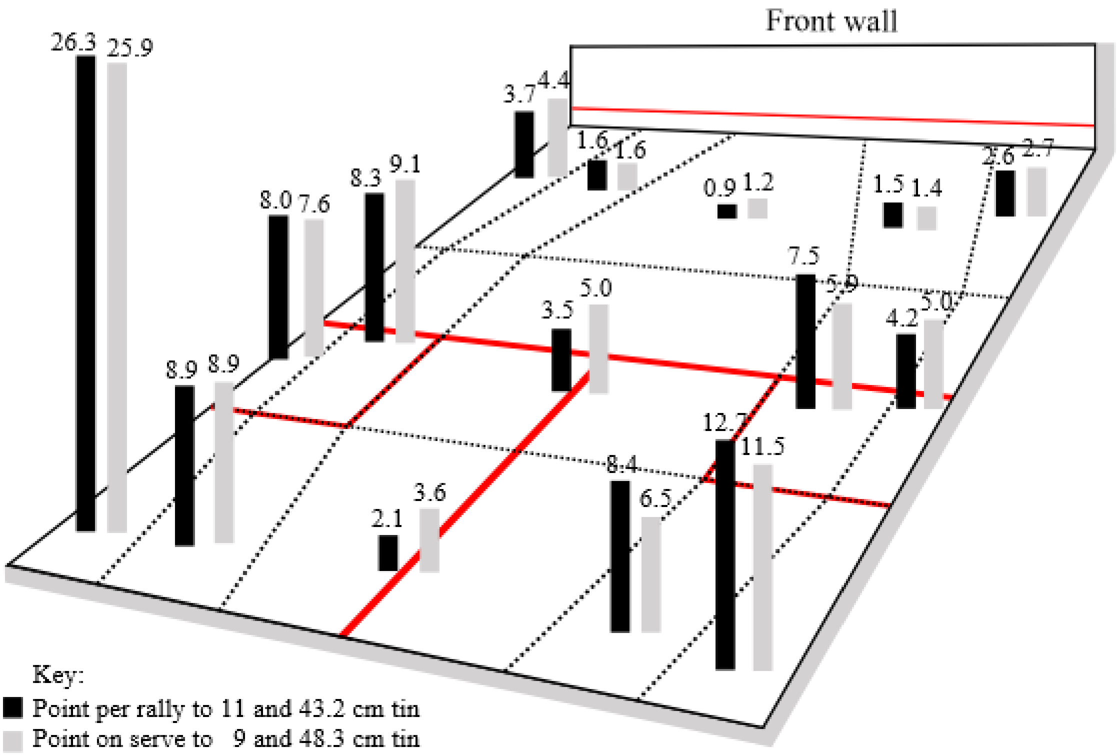


Figure 4.1: Shot distributions played under point-per-rally to 11 and point-on-serve to 9 rules

The variability in rally characteristics for all matches were presented as medians (for comparisons with previous studies) along with percentiles i.e. upper values for each category (Table 4.2) to better present the variability (and skewness) for prescribing specific training routines (Table 4). Rally losers covered trivially more distance than winners in 54.4% of the rallies for short (d = 0.09), medium (59.0%; d = 0.09) long (53.7%; d = 0.07) and very long rallies (54.9%; d = 0.04). The proportion of long, middle and short shots was related to the duration of the rally (chi-square = 440.0, df = 6, p < .001; phi = 0.12; Table 4.2) with fewer shots played in the middle and front of the court as rally duration increased.

Table 4.2: Descriptive statistics for rallies categorised by duration

		Median	Short 25th percentile	Medium 75th percentile	Long 95th percentile	Very long Maximum
Duration of rally (max for categories)		13.2 s	7.0 s	22.7 s	46.4 s	146.5 s
Shots per rally (max. both players)		13	6	25	42	157
Distance (max. per player)		18.8 m	9.6 m	33.5 m	68.3 m	200.0 m
N rallies per game (max inc. lets)		21	17	25	34	41
Game time		11 min 37 s	9 min 6 s	15 min 57 s	23 min 49 s	32 min 6 s
Mean speed in rally	Winner	1.4 m/s	1.2 m/s *(0.4)*	1.4 m/s *(0.3)*	1.4 m/s *(0.2)*	1.4 m/s *(0.2)*
	Loser	1.4 m/s	1.2 m/s *(0.4)*	1.4 m/s *(0.3)*	1.5 m/s *(0.2)*	1.4 m/s *(0.2)*
Mean distance in rally	Winner	18.7 m	5.5 m *(4.1)*	19.6 m *(11.4)*	44.7 m *(13.5)*	90.5 m *(28.7)*
	Loser	19.4 m	5.8 m *(4.8)*	20.3 m *(11.5)*	45.5 m *(14.7)*	91.6 m *(26.9)*
Shots played from	Front	10.3%	16.7%	12.5%	9.2%	6.7%
	Middle	31.3%	44.2%	32.4%	30.2%	29.3%
	Back	58.4%	39.1%	55.1%	60.6%	64.0%

Match duration was related to difference in World rankings between the two players ($r = -0.65$, Figure 4.2) and usually lasted (including breaks between games and rallies) between 35 and 85 minutes except when played between players with dissimilar World rankings (around 40 or more) where much shorter durations occurred (Table 4.3 and Figure 4.2).

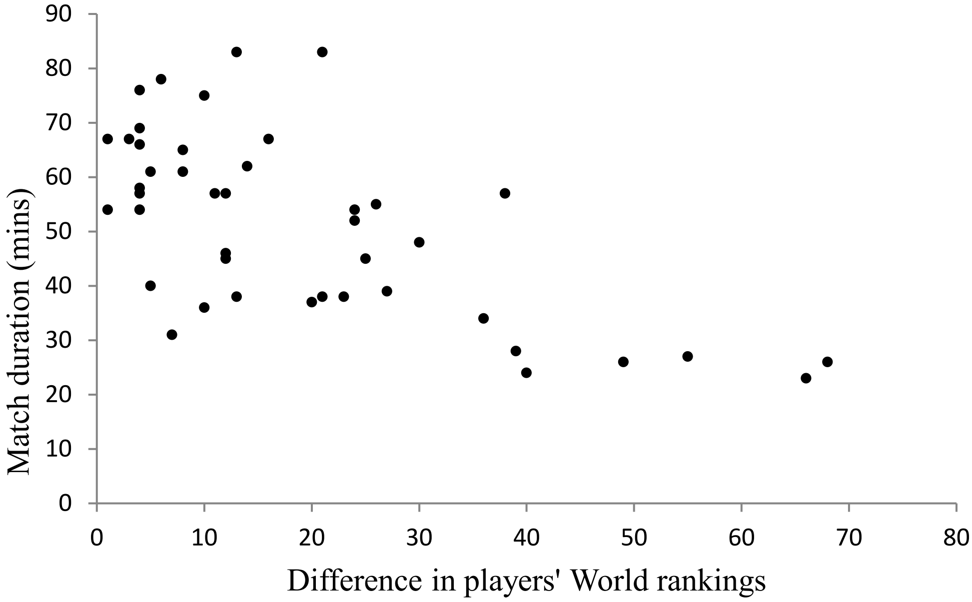


Figure 4.2: Match duration against difference in player ranks

For similarly ranked players (less than 40 ranking points difference), players tended to move a median of 2 km in about 23 minutes ball-in-play duration, split into 80 rallies each lasting 17 s.

Table 4.3: Match statistics (median and *IQR*) for different levels of World ranked male squash players

		Difference between players' World ranking			
	All matches	0 to 10	11 to30	0 to 39	40 or more
N	41	16	17	36	5
Match duration	54 min *(25.5)*	61 min *(13.0)*	50 min *(16.5)*	56 min (24.5)	26 min *(2.0)*
Ball in play	22 min *(10.3)*	25 min 12 s *(8.1)*	23 min 18 s *(7.2)*	23 min 48 s *(8.6)*	11 min 30 s *(1.7)*
Distance (ball in play)	1848.7 m *(1045.7)*	2218.1 m *(796.0)*	1848.7 m *(788.8)*	1995.7 m *(829.4)*	953.1 m *(120.3)*
Number of rallies	77 *(29)*	85 *(22)*	80 *(23)*	82 *(25)*	51 *(2)*
Rally duration	13.2 s *(15.7)*	13.1 s *(15.7)*	13.6 s *(16.8)*	13.3 s *(15.9)*	11.4 s *(13.7)*

Table 4.4 presents the ghosting patterns for the different rally and game durations presented in Table 2. Rest periods of 14 s were used (except after short rallies where 5 s was allowed) to replicate normal between-rally durations (median = 13.7 s).

Table 4.4: Number of ghosting repetitions required to mimic frequency, duration and number of shots for rallies in elite male squash.

	Short rally (7 s)	Medium rally (23 s)	Long rally (47 s)	Very long rally (160 s)	Number of repetitions
Ghosting pattern	1 x front 2 x side 2 x back	2 x front 4 x middle 6 x back	2 x front 6 x side 12 x back	6 x front 24 x side 48 x back	
Short game (9 minutes)	4	7	4	0	15
Medium game (16 minutes)	6	11	5	1	23
Long game (24 minutes)	9	18	8	1	36
Very long game (32 minutes)	12	24	9	2	47

4.5 Discussion

The new PPR to 11 rules (scoring and tin height) have reduced the possibility of rallies not resulting in a point (Lets are still possible), hence, the number of rallies and distance covered have reduced considerably. These shorter game durations, with reduced tin height, indicate that players have changed their shot strategies to take advantage of these easier (physical and environmental) conditions. However, this research revealed that elite-standard men players were hitting the ball to similar areas of the court under the new rules compared to the old but more to the front of the court in shorter rallies than longer ones. These short shots are symptomatic of an attacking strategy as the duration available to return this type of shot is typically less than for shots played to the back of the court. It is not clear if this is a consequence of the new rules, but rally durations were slightly shorter (median = 13.2 s) for PPR to 11 than for POS to 9 (median = 15.0 s) while the number of shots had increased to a median of 13 from 11. Aspiring and current elite-standard players need to condition themselves to cope with the physical demands associated with these rally characteristics. However future research also needs to assess these changes in greater detail to determine the duration between shots for different types of shot, as it is likely that some types will force an opponent to play quicker and thus have less time. This is an important consideration for training to ensure that match play intensities are correctly replicated in training.

While specificity of training is commonly regarded as essential for the conditioning of elite athletes (Reilly et al., 2009), there has been a lack of direction from the scientific literature in some sports. In squash, most research papers have presented mean values for shot (Murray & Hughes, 2001; Hughes & Robertson, 1998), movement (Hughes & Franks, 1994) and match (Murray & Hughes, 2001) characteristics that provide descriptions, but do little to help players devise appropriate training programmes. This study found that distances travelled were mainly a consequence of rally duration (very large effect size), although rally outcome had a trivial effect (partial eta squared = 0.02), with rally winners travelling less distance than losers. On this basis, rally durations were categorised as short, medium, long and very long using 25th, 75th, 95th percentiles and the maximum value obtained in the sample. These four categories were selected so that ghosting routines could be prescribed in a similar ratio as they tended to occur i.e. 5:10:4:1 (up to 25th percentile, up to 75th percentile, 90th percentile and the final 10%).

The first shot for each player requires little movement (return of serve player is stationary and server walks to T after serving). This has more effect on players' speed for short rallies but the influence diminishes as the number of shots in a rally increases. Hence short rallies had lower speeds than the other rally categories, but for training, this is unimportant. Similarly, differences in speed and distance between winning and losing players were small, and less apparent as rally durations increased, and trivial for training

Rally duration had only a small effect on the proportion of shots to the front, middle and back of the court with the clearest difference being for short rallies, which had a greater proportion of shots in the front and middle of the court than other rally categories. This suggested an increased proportion of volleys and the need for a slightly different movement pattern when replicating these rallies. On this basis, movement patterns were presented for short, medium, long and very long rallies with the number of repetitions calculated such that short, medium, long and very long games could be replicated. It is envisaged that players interested or currently playing at elite standard can use these protocols to replicate match durations of their choice (using information from Table III).

4.6 Conclusion

The new rules (scoring and tin height) have reduced the time elite-standard men have to perform shots. Aspiring and current players of this standard need to condition themselves to cope with these demands. This paper has presented a ghosting protocol that replicates the movement patterns for short, medium, long and very long rallies with the number of repetitions calculated such that short, medium, long and very long games could be replicated. Future studies should determine differences in rally characteristics with greater resolution e.g. duration between shots for different types of shot and for different players based on world ranking or playing style.

Chapter 5: Using a Situation Awareness approach to determine decision-making behaviour in squash

This chapter has been published with the following reference:-

Murray, S., James, N., Perš, J., Mandeljc, R. & Vučković, G. (2018). Using a situation awareness approach to determine decision-making behaviour in squash. *Journal of Sports Sciences*, 36, 12, 1415-1422.

5.1 Abstract

Situation awareness (SA) refers to the awareness of all relevant sources of information, an ability to synthesise this information using domain knowledge gained from past experiences and the ability to physically respond to a situation. Expert-novice differences have been widely reported in decision-making in complex situations although determining the small differences in expert behaviour are more elusive. This study considered how expert squash players use SA to decide on what shot to play. Matches at the 2010 (n = 14) and 2011 (n = 27) Rowe British Grand Prix were recorded and processed using Tracker software. Shot type, ball location, players' positions on court and movement parameters between the time an opponent played a shot prior to the player's shot to the time of the opponent's following shot were captured 25 times per second. Six SA clusters were named to relate to the outcome of a shot ranging from a defensive shot played under pressure to create time to an attempted winner played under no pressure with the opponent out of position. This new methodology found fine-grained SA differences in expert behaviour, even for the same shot type played from the same court area, beyond the usual expert-novice differences.

5.2 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.

The aim of this study was to describe shots in squash according to their objective of placing the opponent under pressure using a novel situation awareness (SA) approach.

5.3 Methods

This study used the same data capture procedures with the same matches sampled (see section 4.3) as the previous study.

5.3.1 Data used

The shot type (n=30) and ball location (cell, Figure 3.2) 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 5.1). Information collected following player A's shot (variables related to player B's movments) were used as measures of the action (shot type) selcted 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.

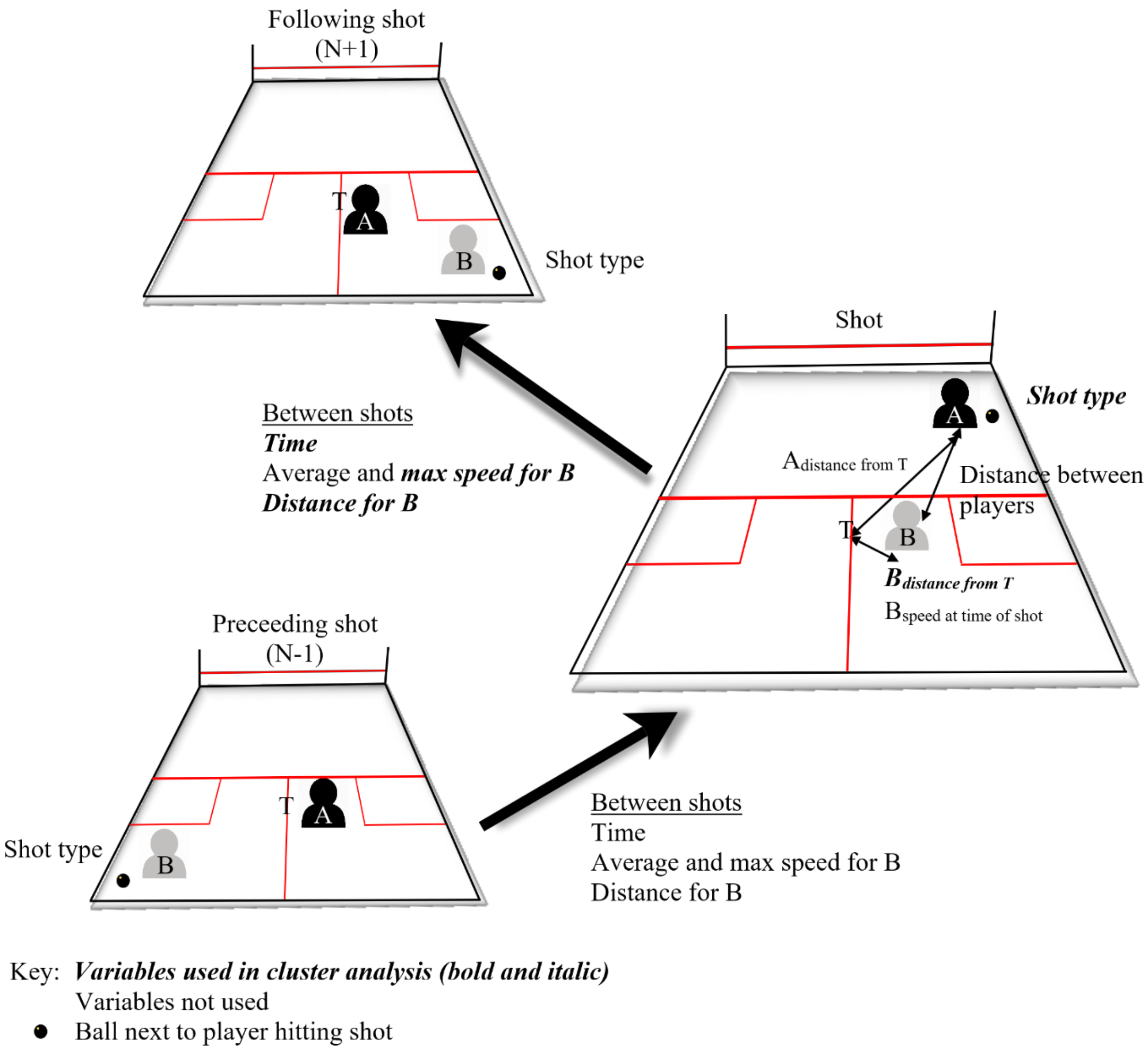


Figure 5.1: Variables collected and those found to discriminate shot types in cluster analysis

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 5.1). 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.

5.3.2 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 5.1: Shot types selected based on court cell played from and cell returned from

Shot type	Cell played from	Cell ball returned from
Drive[a] and Drive wall[b]	11 & 12 or 14 & 15 (front) 1 & 2 or 4 & 5 (back)	1 & 2 or 4 & 5
Drive crosscourt[a] and Drive crosscourt wall[b]	11 & 12 or 14 & 15 (front) 1 & 2 or 4 & 5 (back)	4 & 5 or 1 & 2
Drop	6, 7 & 8 or 8, 9 & 10 (middle) 11 & 12 or 14 & 15 (front)	6, 7, 11 & 12 or 9, 10, 14 & 15
Drop crosscourt	6, 7 & 8 or 8, 9 & 10 (middle) 11 & 12 or 14 & 15 (front)	9, 10, 14 & 15 or 6, 7, 11 & 12
Boast attack and Boast defence	1 & 2 or 4 & 5 (back) 6 & 7 or 9 & 10 (middle)	14 & 15 or 11 & 12
Lob	11 & 12 or 14 & 15	1 & 2 or 4 & 5
Lob crosscourt	11 & 12 or 14 & 15	4 & 5 or 9 & 10 1 & 2 or 6 & 7
Kill	1, 2, 6 & 7 or 4, 5, 9 & 10	6, 7, 11 & 12 or 9, 10, 14 & 15
Drive (v)[a] and Drive wall (v)[b]	6, 7 & 8 or 8, 9 & 10	1 & 2 or 4 & 5
Drive crosscourt (v)[a] and Drive crosscourt wall (v)[b]	6, 7 & 8 or 8, 9 & 10	4 & 5 & 2 or 1 & 2
Drop (v)	6, 7 & 8 or 8, 9 & 10	6, 7, 11 & 12 or 9, 10, 14 & 15
Drop crosscourt (v)	6, 7 & 8 or 8, 9 & 10	9, 10, 14 & 15 or 6, 7, 11 & 12
Boast attack (v) and Boast defence (v)	6, 7 & 8 or 8, 9 & 10	14 & 15 or 11 & 12
Kill (v)	1, 2, 6, 7 & 8 or 4, 5, 8, 9 & 10	6, 7, 11 & 12 or 9, 10, 14 & 15
Kill crosscourt (v)	1, 2, 6, 7 & 8 or 4, 5, 8, 9 & 10	9, 10, 14 & 15 or 6, 7, 11 & 12

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 volleyed (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 (***Time***)
- Distance player B moved to return player A's shot (***Distance for B***)
- Maximum velocity of player B from the moment player A hit the shot to player B returning the shot (***max speed for B***)
- Distance player B was from the T at the moment player A hit the shot ($\boldsymbol{B}_{\textit{distance from T}}$)

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).

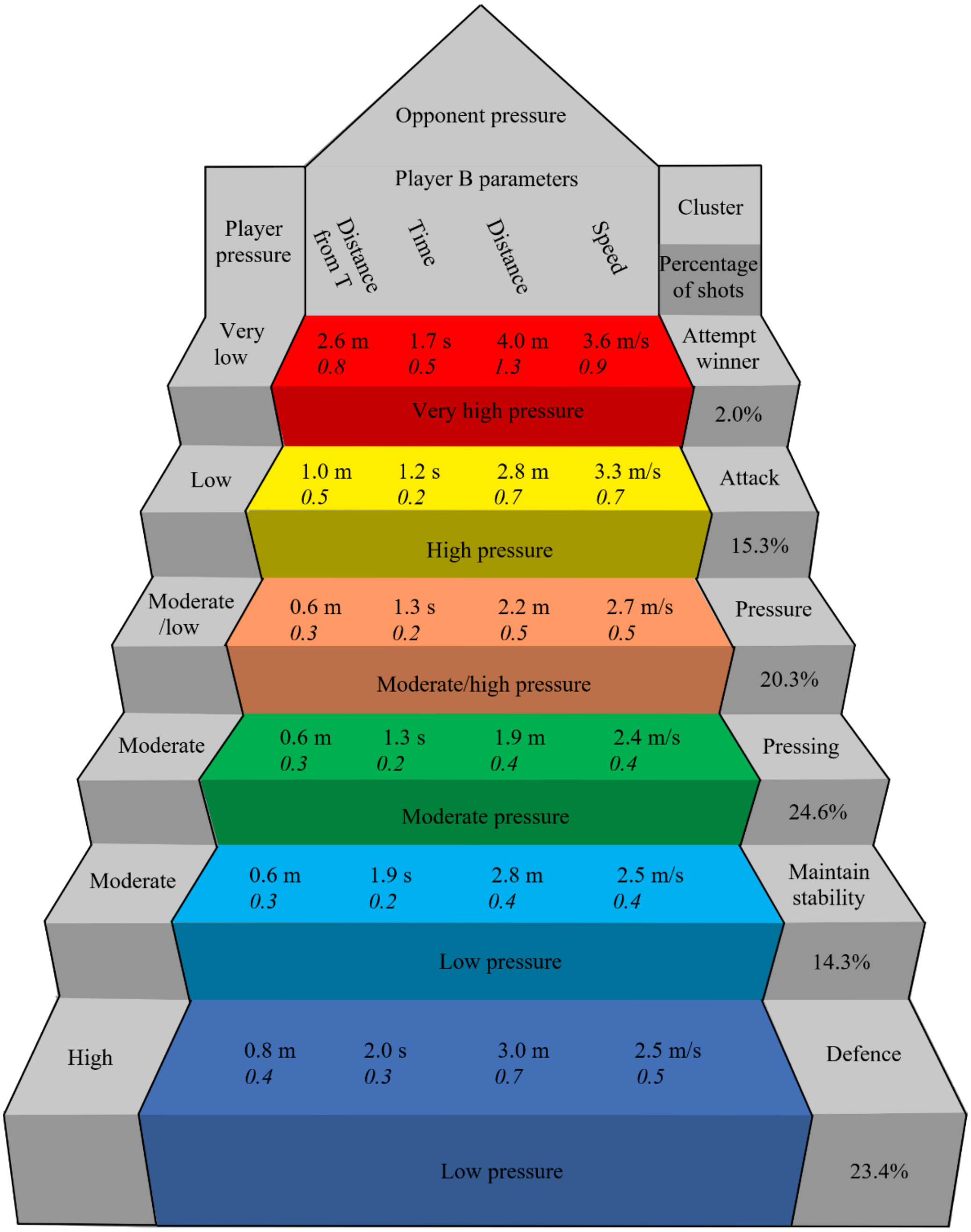
5.3.2.1 Determining the number of clusters

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.

5.4 Results

Six SA clusters were named to relate to the outcome of a shot (Figure 5.2). 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 5.3).



Player pressure	Distance from T	Time	Distance	Speed	Opponent pressure	Cluster	Percentage of shots
Very low	2.6 m *0.8*	1.7 s *0.5*	4.0 m *1.3*	3.6 m/s *0.9*	Very high pressure	Attempt winner	2.0%
Low	1.0 m *0.5*	1.2 s *0.2*	2.8 m *0.7*	3.3 m/s *0.7*	High pressure	Attack	15.3%
Moderate /low	0.6 m *0.3*	1.3 s *0.2*	2.2 m *0.5*	2.7 m/s *0.5*	Moderate/high pressure	Pressure	20.3%
Moderate	0.6 m *0.3*	1.3 s *0.2*	1.9 m *0.4*	2.4 m/s *0.4*	Moderate pressure	Pressing	24.6%
Moderate	0.6 m *0.3*	1.9 s *0.2*	2.8 m *0.4*	2.5 m/s *0.4*	Low pressure	Maintain stability	14.3%
High	0.8 m *0.4*	2.0 s *0.3*	3.0 m *0.7*	2.5 m/s *0.5*	Low pressure	Defence	23.4%

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

Figure 5.2: Time, distance and speed parameters for the six SA clusters

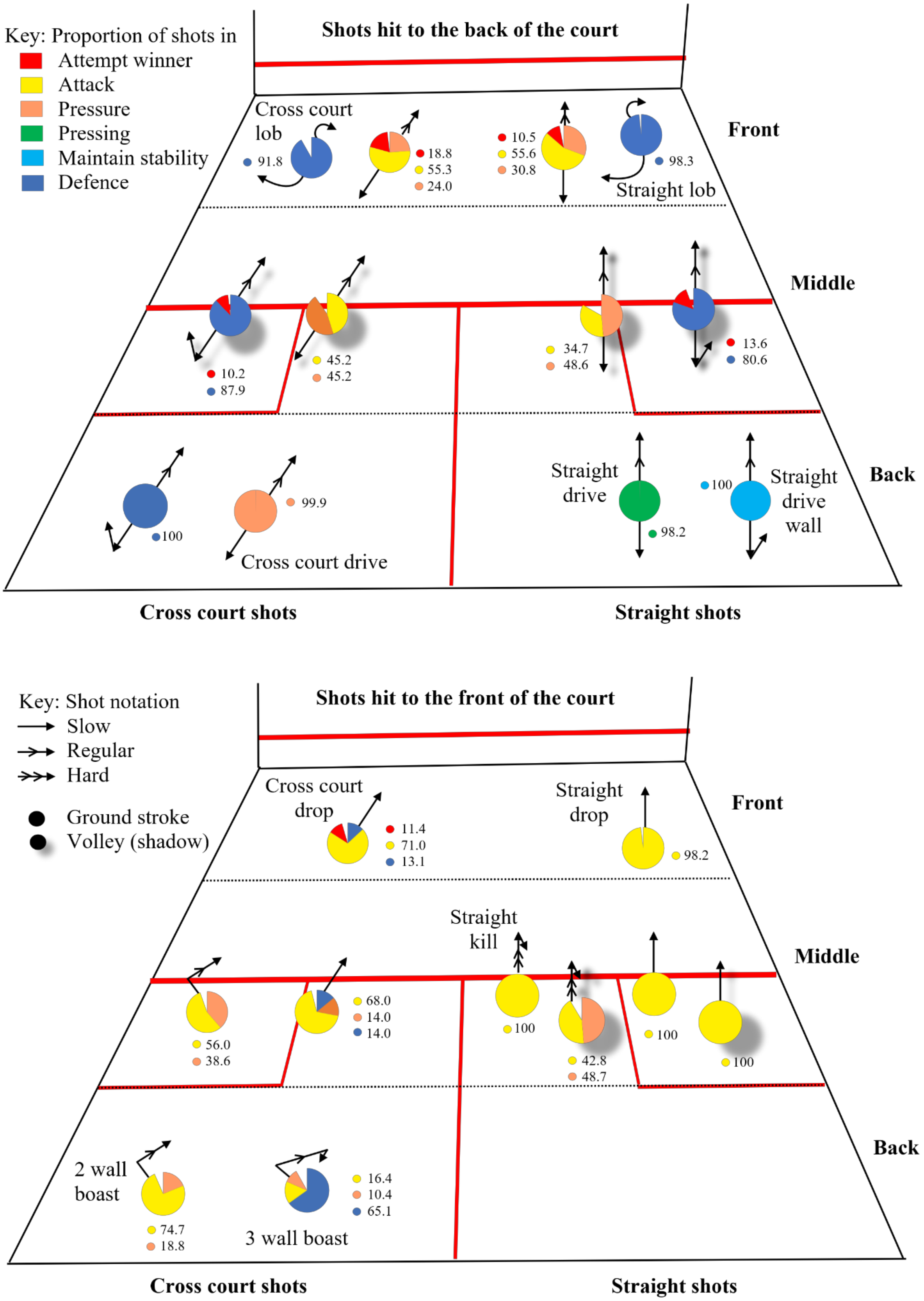


Figure 5.3: Proportion of shot types classified in six SA clusters

5.5 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.

5.6 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).

Chapter 6: Using a Situation Awareness approach to identify differences in decision-making behaviour between the World's top two squash players and their opponents

6.1 Abstract

The pressure exerted on a squash player is a consequence of the quality of a shot coupled with the ability of the player to return the ball i.e. the coupling of the two players' situation awareness (SA) abilities. SA refers to an awareness of all relevant sources of information, the ability to synthesise this information using domain knowledge and the ability to physically respond to a situation. Matches at the 2011 (n = 9) Rowe British Grand Prix, held in Manchester, UK were recorded and processed using Tracker software. Shot type, ball location, players' positions on court and movement parameters between the time an opponent played a shot prior to the player's shot to the time of the opponent's following shot were captured 25 times per second. All shots (excluding serves and rally ending shots) produced five main SA clusters, similar to those presented by Murray et al. (2018), except a greater proportion of shots were categorised in the greater pressure clusters and less in the lower pressure ones. Individual matches were presented using cluster profile infographics which demonstrated how individual players played differently in different matches. This approach should be further modified to determine within match changes in performance.

6.2 Introduction

In squash, like all racket sports, the main objective of any shot is to minimise the amount of time available to the opponent to hit their shot. This is optimally achieved by hitting the shot accurately and early e.g. a volley, forcing the opponent to move quickly over a maximal distance. To counter this pressure, expert players can anticipate where the ball will go (e.g. Abernethy, 1990; Triolet, Benguigui, Le Runigo, & Williams, 2013) and move very efficiently by utilising a split step to initiate movement (James, & Bradley, 2004) in the correct direction, follow a well-defined path before lunging to hit the ball. This also allows a very efficient return to the T area of the court, where winning players have been shown to spend a greater proportion of total playing duration than losers (Vučković, Perš, James, & Hughes, 2009). This means that two factors determine the amount of pressure exerted on a player, one, as a consequence of the quality of a shot, and two, the ability of the player to move to return the ball which involves knowing where the ball will go as soon as possible, potentially some anticipatory behaviour. Triolet et al. (2013) estimated that elite tennis players only demonstrated anticipation behaviours in between 6.14% and 13.42% of the situations analysed. The suggestion being that in most situations, tennis players don't need to anticipate, because waiting for sufficient ball flight information will enable them to return the ball without any risk. James and Bradley (2004) also found limited use of anticipation in expert squash players as they initiated their first movement towards the ball 270ms (±0.09s) after ball contact, factoring in reaction time this suggested they utilised ball flight information before moving. However, only relatively easy shots were sampled, to prevent situational probabilities from being used (suggested as a potential confounding variable by Abernethy, Gill, Parks, & Packer (2001). Whilst these studies suggested that anticipatory behaviours were not as prevalent as perhaps assumed, an alternative hypothesis is that players could anticipate more than suggested by the data. Either they chose not to act, possibly because overuse of anticipation would be detected by their opponent and over anticipating could end up counterproductive, or anticipatory behaviour simply enables the response to be planned and executed more effectively, often without the need for either an early movement or unnecessary speed.

A fundamental question, albeit a difficult one to answer, relates to which shot should be played in any situation. If an optimal shot was always available, it would be likely that expert players would usually select this shot. This would mean that discernible patterns of play i.e. consistent shots played in certain situations, would be evident. McGarry and Franks (1996) found that invariant patterns of play were difficult to ascertain but suggested that the complexity of discriminating the situation in which the shot was played was a crucial factor. They suggested that the preceding shot alone was likely to be insufficient to predict the subsequent shot. In response, Vučković, James, Hughes, Murray, Milanović, Perš, & Sporiš (2014) controlled for previous shot type, time between shots, court location and the handedness of the players. They found that tight shots (played from close to the corners of the court) tending to be more predictable (two or three typical shots played) compared to loose ones (up to seven different shot responses to the same preceding shot when nearer the middle of the court).

Murray, James, Perš, Mandeljc, & Vučković, (2018) described shot selection in squash from a situation awareness (SA) perspective (Endsley, 1995). SA refers to the awareness of relevant sources of information, the synthesis of this information using domain knowledge gained from past experiences (Abernethy et al., 2001) and the ability to physically respond to the situation. Murray et al. (2018) suggested the relevant sources of information were 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…..". This perspective demonstrates the complexity involved in deciding which shot to play and raises the question as to what extent individual differences affect this decision-making process. Within this SA perspective the final task of actually playing the shot is important since an inaccurate shot would likely offset any advantage gained from having successfully accomplished the first two tasks.

Previous research has tended to analyse relatively large data sets which group individual players according to their level of expertise (e.g. Vučković, et al., 2014). This aproach fails to consider individual differences, potentially falling into what Mackenzie and Cushion (2013) identified as a 'theory-practice gap', where research findings were suggested to have a lack of transferability and had little or no relevance to practitioners in sport. They advocated that performance analysis research should be for practitioners to utilise the results to improve performance. To address this issue, more discriminating information relating to, processes rather than just outcome measures (James, 2009), and in relation to individual, rather than multiple, players or teams are required.

Murray et al. (2018) presented six shot type clusters, referred to as SA clusters, named to relate to the outcome of a shot ranging from a "defensive" shot played under pressure to create time to an “attempted winner” played under no pressure with the opponent out of position. They used a two-step cluster analysis using two distance parameters (how far the player moved to return the shot and the distance the player was from the T at the moment the shot was hit) as well as the time and maximum velocity of the player returning the shot (between the shot and the returning shot). They only used shots that were played from selected areas of the court (front, middle or back) that had achieved their objective, namely the ball was returned from the area of the court aimed for. The logic for this decision being that shots that didn’t achieve their objective would have, potentially significant, different movement parameters e.g. when an opponent anticipated a shot and was able to volley the ball or the shot was played badly enough to allow this type of interception. By only analysing shots that achieved their objective, the authors were able to differentiate different SA clusters for the same shot type from the same court area, suggested as being consistent with players changing the pace and trajectory of the shot because of different objectives (SA tasks). However, this may also reflect the accuracy of a shot, since less accurate shots would likely place less pressure on an opponent compared to a more accurate one. This logic also meant that they differentiated the same shot when played from different areas of the court. Whilst this allowed the fine discrimination sought in this paper, the methodology had the effect of removing around 50% of shots from the analysis. This meant that the ecological validity of the study was reduced since information regarding a lot of shots was missing.

Previous papers have grouped players according to their world ranking (e.g. Hughes and Robertson, 1998; Murray et al., 2016) but we argue that players are always moving up or down the ranking list and their current world ranking may not be an accurate reflection of their ability at the time a match is played. This is particularly obvious for young emerging players or older players moving down the ranking list. Similarly, players may have different strengths and weaknesses that mean that they play with somewhat different approaches e.g. high tempo risky versus defensive attrition. Grouping these different players together will therefore reduce the accuracy of the analysis. It is the aim of this paper, therefore, to compare the shot selections, and shot effectiveness, of two elite players, ranked as the top two players in the world at the time of data collection, playing against each other and against different ranked opponents using shots that both achieved and did not achieve their objective i.e. where the return shot was played from was not a factor check except for lobs which were returned from the front of the court as this very unusual situation was removed from the analysis. This approach will provide a more detailed analysis of the differences evident between players of very similar ability.

6.3 Methods

This study used the same data capture procedures with the same matches sampled (see section 4.3) as the previous two studies.

6.3.1 Data used

The shot type (n=25) and ball location (cell, Figure 3.2) 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 (see Murray et al., 2018 for original methods and justification). This study did not differentiate the same shot when played from different areas of the court as did Murray et al. (2018), rather shots were classified by type e.g. straight drive, irrespective of whether it was played from the front or back of the court. This procedure was considered more appropriate since all rally continuing shots were analysed, rather than only the ones that achieved their objective, as Murray et al. (2018) did. This meant that the variability associated with the variables collected was far greater and this complexity prompted the simplification of the shot classification. One shot was removed from the data (lob from front of the court that was volleyed in the front of the court) as the variables collected suggested this was an attacking shot. This was, however, either a poorly executed defensive shot or a very unusual interceptive movement by the opponent. For elite players, both situations are rare and were hence deemed outliers and removed.

6.3.2 Statistical analyses

The same cluster analysis as used by Murray et al. (2018) was used. This 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). 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).

The two-step cluster analysis, using a probability-based log-likelihood distance measure (SPSS) enabled the same continuous (two distance parameters, time and maximum velocity) and categorical (shot type) variables to be used. However, when running a cluster analysis on different data, we used all shots rather than Murray et al.'s constrained shots, different clusters were found from those reported by Murray et al. (2018). The cluster parameters (all players, all shots) were however, similar, hence we used the same names for the new clusters. Differences became more marked, however, when individual players were analysed necessitating the need to quantify cluster similarity and the magnitude of the difference.

6.3.2.1 Determining which was the most similar cluster

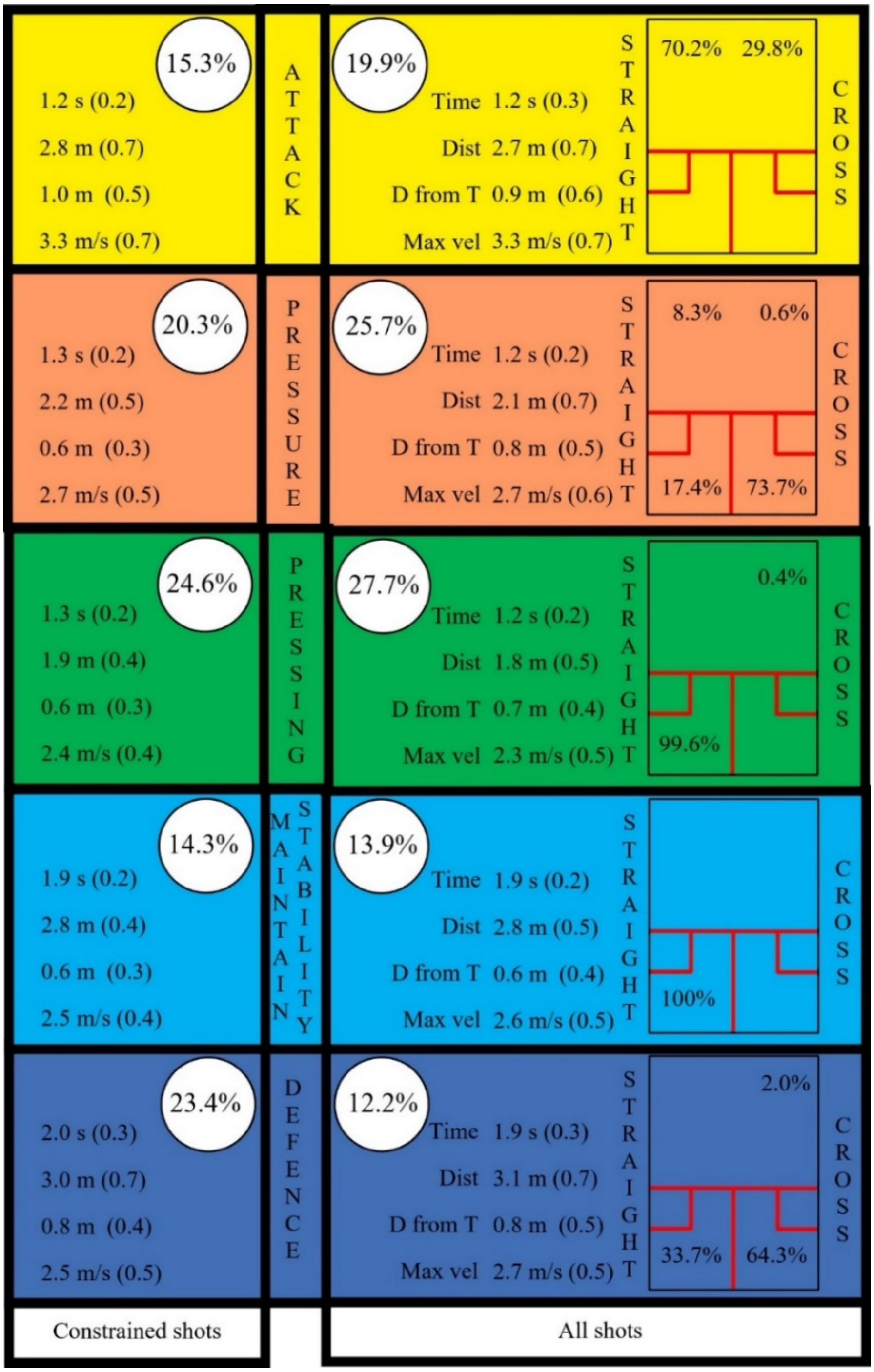
Each cluster was determined by the group mean (cluster centroid) based on the four continuous (two distance parameters, time and maximum velocity) and one categorical (shot type) variable. To determine which cluster (all players, all shots) each individual player cluster most resembled, the absolute differences, between the means for each continuous variable for one individual player cluster and the same variable for all clusters (all players, all shots) were calculated. The cluster which had the lowest sum, of the four absolute differences, was hence deemed the most similar.

6.3.2.2 Determining the degree of difference between the clusters

This was achieved by finding how far the mean for each parameter, for one individual player, was from the mean of each parameter, for all players, in terms of standard deviations i.e. the z score. The four z scores were then summed, not averaged because scores could be both negative and positive, to give an overall deviation value.

6.4 Results

Five SA clusters were named, the same as for Murray et al.'s (2018) constrained shot approach, to relate to the outcome of a shot (Figure 6.1). When all shots from Murray et al.'s (2018) data set were used, the proportion of shots creating the most pressure on the opponent, increased in comparison to the previously used, constrained shot approach.



Key: All movement values are means (standard deviation in brackets)

Dist is distance opponent moved, Time and Max vel (maximum speed of opponent) all from shot being played to return shot

D from T is opponent distance from T when shot played

Court floor indicates where shots within cluster were played to (back or front of court) and whether played straight (shown on left side of court) or crosscourt (right)

Figure 6.1: Time, distance and speed parameters for five SA clusters using constrained shots (Murray et al., 2018) compared to all shots (all data from Murray et al., 2018)

In order to present the different clusters relative to each other, whilst also presenting all four variables, an infographic was created (Figure 6.2). The centre of each circle (cluster) is located according to the mean value for time (x) and distance (y), between the shot being played and the return shot. The distance the opponent was from the T at the time of the shot is represented by the length of the T which is drawn relative to the x axis. Finally, the diameter of the circle is proportional to the maximum speed the opponent ran to return the shot. This infographic depicts three attacking (attack, pressure and pressing) and two defensive (defence and maintain stability) clusters.

Figure 6.2: Shot clusters for all shots (all data from Murray et al., 2018)

The infographic (Figure 6.2) was then used for all 9 matches involving the World number 1 and 2 players (their performances in the middle and their opponents outside) culminating in the final played between them (Figure 6.3).

Each match demonstrated different cluster patterns with matches involving the World number 1 displaying a tendency for greater pressure to be exerted as the standard of the opponent increased, 30.9% defensive shots (Ashour, 38% his opponent) against the opponent ranked outside the World’s top 24 compared to 22.1% (24.7% opponent) against his top 8 ranked opponents.

Figure 6.3: Shots clusters for matches involving Ramy Ashour (World ranked number 1) and Nick Matthew (World ranked number 2)

In the final the World number 2 forced his opponent to move slightly further with more accurate shots to the back (volleys straight, crosscourt drives) and front of the court (straight kills, volley straight kills) categorised in the pressure cluster whereas these shots were categorised in the pressing cluster for the world number 1. In contrast, the World number 1 gave his opponent less time on shots usually associated with the defence cluster (as they were for world number 2). Hence 54% of his crosscourt drives that reached the back wall, 89.5% of volleys straight that reached the back wall, 37.5% of 3 wall boasts and 100% of crosscourt lobs were categorised in an attacking cluster.

6.5 Discussion

The amount of pressure exerted on a squash player is a consequence of the quality of a shot coupled with the ability of the player to move to return the ball. This paper has categorised shots according to this pressure, therefore the categorisation of shot types according to four variables associated with opponent movement encapsulates both the quality of the shot and the opponent's ability to offset the pressure. Murray et al. (2018) focussed more on the former part of this pressure, namely the pressure exerted by the shot, as they only selected shots that achieved their objective. They removed shots where the opponent volleyed the ball in the middle of the court for example, often a consequence of anticipating the ball trajectory. This approach was deemed to discriminate decision-making where the same shot type played from the same court area produced different outcomes (SA clusters) as this was suggested as consistent with players changing the pace and trajectory of the shot because of different objectives (SA tasks).

This study adopted an alternative approach and included shots that did not achieve their objective, in other words shots which were played less accurately or where the opponent was able to anticipate and return the ball early. This approach complicates the analysis as more factors are likely to determine the amount of pressure a player is under but clearly has greater ecological validity in that this is a more accurate reflection of elite squash match play.

Utilising the approach of using all shots (excluding rally ending shots as these require a separate analysis; see also Murray et al., 2018) five main SA clusters were found to be very similar to those presented by Murray et al. (2018). The attempt winner cluster only accounted for 0.6% of shots in this study and was thus not presented. The clear impact of using all shots was that a greater proportion of shots were categorised in the greater pressure clusters (pressing, pressure and attack) and less in the lower pressure ones (maintain stability and defence).

The relationship between the four movement variables that defined each SA cluster was not clearly presented by Murray et al. (2018) prompting the creation of an infographic in this paper. The challenge of presenting four dimensions was alleviated by using just two dimensions (time and distance) with the other two represented by the size of the circle and length of T. This clearly differentiated two low and three high pressure clusters when all players and all shots were used. However, this overview of multiple players lacks the transferability in relation to individual players, the so called 'theory-practice gap' (Mackenzie and Cushion, 2013).

Individual matches were presented to highlight how individual players exhibited different cluster formations in different matches. One observation was that the World number 1 tended to increase the pressure on opponents as the opponent quality increased. This is suggestive of a strategy of playing within himself when the opponent threat was minimal but when necessary his performance levels increased. This supports the finding of McGarry and Franks (1996) who found consistent patterns of play elusive. However, their work comprised a sample of 8 elite players taken from 10 matches where invariant patterns of play would be less likely than for one player in one match as presented here. The degree of difference, both between and within players, found here suggests that many researchers have previously underestimated the extent that individual differences play in decision-making processes, in this case deciding which shot to play.

A fine-grained analysis of the final, played between the two top players in the World at the time, indicated subtle difference of relevance to practice (Mackenzie and Cushion, 2013). For example, the World number 1 gave his opponent less time on crosscourt drives and volleys straight, shots which reached the back wall, and usually associated with the defence cluster. This was because he was hitting these shots very hard, hence reducing the time and causing the shots to be categorised in an attacking cluster. This is predominately a consequence of the quality of a shot since the opponent was unable to return the ball early. Another of his shots, the crosscourt lob, was similarly categorised in an attacking cluster even though this is usually a very defensive shot. In this case it is likely that the opponent was forced to volley the ball because of the accuracy of the shot i.e. reducing the time but also if the opponent had to move forwards to cover a drop shot then the distance covered, and maximum speed would also be higher and hence reflect the attacking cluster categorisation.

6.6 Conclusion

This paper has further demonstrated the usefulness of analysing squash from a SA approach but has also demonstrated the inherent variability associated with squash match play. The dynamic between the player trying to put pressure on an opponent by playing accurate shots is offset to some extent by an opponent who move efficiently thanks to an awareness of relevant sources of information and the synthesis of this information using domain knowledge gained from past experiences (Abernethy et al., 2001). This sensitivity of this approach can be further modified to determine within match changes in performance.

Chapter 7: General discussion

7.1 Summary findings from thesis

This thesis aimed to better understand elite male squash match play from a more general perspective, due to rule changes having potentially changed the way the game is played, to more specific information related to shot selection. In attempting to achieve these goals new methods were presented that answered the criticisms of previous research posed in the review of literature as well as creating novel ways to present performance analysis data using infographics.

At the outset of this thesis it was recognised that reliability tests were necessary to determine limitations of the data collection methods and the understanding and ability for the analyst(s) to use the notation codes appropriately. This was crucial, as without accurate data, the validity of the whole thesis would be jeopardised. Consequently, significant training on using the software (O1a) was undertaken before data could be assessed, at the level of the subsequent analysis (Hughes, Cooper and Nevill, 2002). This included a novel approach for determining court areas where the ball was hit using Kappa and percentage agreement calculations, as well as associated shot information e.g. shot type and the effects of prolonged notating matches (O1b). Equally important, but seldom mentioned in research papers, was the verification of the accuracy of the data prior to undertaking any analysis along with the need to process the data to enable appropriate analyses to take place. A series of queries, specific to each data type, following data collection and prior to data analysis, discovered multiple errors in the data, due to failing to code events and coding events incorrectly, which had to be corrected (O1c).

The literature review identified that previous research in squash had tended to use less than optimal methods to portray findings e.g. using mean values without any measure of dispersion (e.g. Hughes and Robertson, 1998). Chapter 4 presented match and game statistics using percentiles, median and maximum values (O2b) to better portray the extent to which elite male squash rallies vary, to some extent due to the degree of similarity of the two playing standards (O2d). The recently advocated use of more meaningful inferential statistics i.e. effect sizes (Winter et al., 2014) to meaningfully compare differences(O2c), were used to compare rally variables between point on serve to 9 and point per rally to 11 (O2a). Novel infographics were also designed in Chapter 5 to present the amount of pressure placed on an opponent by different shot types played from different areas of the court (O3c) and in Chapter 6 to display the four variables used to determine SA clusters and this was used to present the decision-making behaviours evident in different matches (O4c).

The literature review also presented the significant work in performance profiling of squash (e.g. Murray and Hughes, 2001). This research attempted to determine the number of matches required to adequately represent what a player tended to do in a match (Hughes, Evans and Wells, 2001) although questions were asked about the validity of this (O'Donoghue, 2005), the lack of consistency in some performances (McGarry and Franks, 1996) and the number of variables that needed to be controlled for to find invariance of some degree (Vučković et al., 2014). The suggestion for this thesis was therefore to use more complex analyses to find patterns on the data. This was achieved using a data mining technique, cluster analysis, to determine the variables (time, distance, speed, shot type, court area) that best discriminated the amount of pressure an opponent was put under following a shot that achieved its goal (O3b). This work was based on a novel situation awareness (SA) approach to understanding decision-making behaviours (shot selections) and extended in Chapter 6 to include shots that did not achieve their goal (O4a). This work derived slightly new SA clusters and prompted the development of a methodology for comparing SA clusters derived from different data sets (O4b), in this instance the top two squash players in the world playing against each other and against different ranked opponents.

7.2 Key limitations identified in thesis

- This thesis only analysed elite male squash players and hence the findings are unlikely to be relevant to either female players or sub-elite players.
- Mackenzie and Cushion (2013) identified the use of relevant independent variables to be crucial if meaningful information is to be derived for use in an applied setting. This thesis used a number of independent variables, Chapter 4 used games rules, rally outcome, rally duration and World ranking points difference whilst Chapter 6 used playing standard.

However, other independent variables such as fatigue, injury and time within match may have affected the results. To data no research has considered how tactics may evolve through a match.

7.3 Summary of practical implications for applied squash performance

One of the explicit aims of this thesis was to address the comments of Mackenzie and Cushion (2013) who advocated that performance analysis research findings need to be of use in an applied setting. As someone who has worked with elite squash players over many years, utilising the methods presented in many of the research papers referred to here, this comment resonates deeply.

Previous researchers had consistently used 16 rectangular cells of equal dimension with no discussion of their appropriateness (e.g. Murray and Hughes, 2001). This thesis considered that the location of the shot would be more critical near the sidewalls than in the centre of the court. Also, straight and crosscourt shots tend to have different trajectories, particularly when the sidewalls are hit. Consequently, a new method of dividing the court into cells, which took into consideration squash tactics, was presented in Chapter 3 and used throughout the thesis. This method has greater logical validity and is suggested as an imperative evolution for future squash analysis.

Rule changes brought in to shorten games (score on every rally) and make the game more exciting (lower tin to encourage attacking shots) were assessed by comparing matches played under the old POS to 9 with the new PPR to 11. The number of rallies and distance covered were shown to have reduced considerably although these elite players were hitting the ball to similar areas of the court under the new rules compared to the old but more to the front of the court in shorter rallies compared to longer ones. Rally durations were slightly shorter (median = 13.2 s, previously 15.0 s) while the number of shots had increased to a median of 13 from 11. Distances travelled were shown to be mainly a consequence of rally duration and rally outcome had a trivial effect, with rally winners travelling marginally less distance than losers. On this basis, rally durations were categorised as short, medium, long and very long using 25th, 75th, 95th percentiles and the maximum value obtained in the sample. These four categories were selected so that ghosting routines could be prescribed in a similar ratio as they tended to occur. Hence aspiring and current elite-standard players need to condition themselves to cope with the physical demands associated with these rally characteristics, ghosting routines were developed in Chapter 4 to mimic match play demands (O2e).

7.4 Original contributions to knowledge of squash performance

Reliability testing, whilst relatively unexciting in comparison to the findings of the research, proved to be a vital component of determining the limitations of the data collection process. When these had been identified by the reliability results amendments were made prior to data collection for the thesis. However, even following this procedure error checking post data collection and processing revealed multiple errors in the data. A series of queries, specific to each data type, were implemented to find the errors and rectify them. This under-reported methodology was deemed critical for the possibility of valid data analysis.

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). Experts undertake these tasks faster and exhibit better performance on each of these SA tasks. Chapter 5 demonstrated that 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 suggesting that players change the pace and trajectory of a shot because of different objectives (SA tasks). The discrimination of these differences was unique to this study and offer a new insight into the very subtle differences between players of quite similar expertise. 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.

7.5 Future research directions

The use of movement data in this thesis has allowed a more complex analysis of squash than most previous research. Careful consideration of each aspect of the analysis has also produced new techniques and outputs. These have been framed within the evolution in performance analysis from the use of sports science techniques primarily developed by the disciplines of statistics, the sciences and psychology to the more recent use of computer science techniques. These have become more popular and have been incorporated in this thesis. However, in performance analysis the use of pattern recognition e.g. machine learning algorithms, have been limited. This would seem to be a logical way forward in trying to understand how tactical changes may occur within matches. To date, this has not been possible as current analysis techniques have been unable to discriminate tactical changes. However, elite players are undoubtedly able to alter their tactics when things are going wrong for example. New methods to discriminate these subtle changes would help performance analysis to better identify tactics, and hence player profiles, leading to better information

7.6 Conclusions

At the outset of this thesis the prevailing view of performance analysts working in squash was the belief in stable playing profiles. Little empirical work supported this view even though in the applied world belief of individual styles of play were abundant. Clarification of what a playing profile was, the extent to which players play in the same way or differently, was largely unknown. This thesis sought to analyse elite male squash match play with a view to at least clarify how the game was played, the degree of invariance associated with shots selected and whether individual playing profiles were stable. These aspirations led to the presentation of match data using percentiles to show the variability inherent in squash match play, a method that was thought superior to the misleading use of mean values. An alternative method for categorising shot types was devised, to include movement data, which ultimately assessed shot outcomes in terms of the Situation Awareness of both players i.e. the ability of one player to hit a good shot coupled with his opponent's ability to return the shot efficiently. This approach demonstrated through the use of a novel cluster profile infographic that individual players change tactics between matches and, although not explicitly tested here, within matches.

This thesis has provided updated performance data, since new rules were adopted, that can assist aspiring elite players to train effectively. Match play has become more intense with the time between shots reducing because of improvements in the SA of players. New methods of analysis and presentation of results have been presented in this thesis which have been published in International journals with the wish for some of these advances to be further developed by other researchers in other sports.

Chapter 8: References

Abernethy, B. (1990). Anticipation in squash: Differences in advance cue utilization between expert and novice players. *Journal of Sports Sciences*, 8(1), 17-34.

Abernethy, B., Gill, D.P., Parks, S.L., & Packer, S.T. (2001). Expertise and the perception of kinematic and situational probability information, *Perception*, 30, 233-252.

Araujo, D., Davids, K.W., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of Sport and Exercise*, 7(6), 653-676.

Atkinson, G., & Nevill, A. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26, 217-238.

Borrie, A., Jonsson, G.K., & Magnusson, M.S. (2002). Temporal pattern analysis and its applicability in sport: an explanation and exemplar data. *Journal of Sports Sciences*, 20, 845-852.

Chabris, C.F., & Hearst, E.S. (2003). Visualization, pattern recognition, and forward search: effects of playing speed and sight of the position on grand master chess errors, *Cognitive Science*, 27, 637-648.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices, *Cognitive Science*, 5(2), 121-52.

Cohen, J. (1988). Statistical power analysis for the social sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Downey, J.C. (1973). *The Singles Game.* London: E.P. Publications.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems, *Human Factors*. 37(1), 32-64.

Ericsson, K.A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains, *Academic Medicine*, 79(10 Suppl), S70-S81.

Ericsson, K.A. (2006). The Influence of experience and deliberate practice on the development of superior expert performance. In K.A. Ericsson, N. Charness, P.J. Feltovich & R.R.

Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683-704). New York, NY: Cambridge University Press.

Ericsson, K.A., Krampe, R.T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance, *Psychological Review*, 100, 363-406.

Fullerton, H.S. (1912). The inside game: the science of baseball. *The American Magazine*, *LXX*, 2-13.

Girard, O., Chevalier, R. Habrard, M., & Millet, G.P. (2007). Game analysis and energy requirements of elite squash. *Journal of Strength and Conditioning Research*, *21*(3), 909-914.

Gréhaigne, J.K., & Godbout, P. (1995) Tactical knowledge in team sports from a constructivist and cognitivist perspective. *Quest,* 47, 490-505.

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1995). Multivariate data analysis with readings (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Haken, H. (1983). *Advanced synergetics: instability hierarchies of self-organizing systems and devices.* New York: Springer-Verlag.

Haken, H., Kelso, J.A.S., & Bunz, H.A. (1985). Theoretical model of phase transitions in human hand movements. *Biological Cybernetics,* 51, 347-356.

Hong, Y., Robinson, P.D., Chan, W.K., Clark, C.R., & Choi, T. (1996). Notational analysis on game strategy used by the world's top male squash players in international competition. *The Australian Journal of Science and Medicine in Sport,* 28(1), 18-23.

Hughes, M.D. (1985). A comparison of the patterns of play in squash. In: I.D. Brown, R. Goldsmith, K. Coobes & M.A. Sinclair (Eds.), *International Ergonomics,* (pp. 139-141), London: Taylor & Francis.

Hughes, M., & Franks, I.M. (1994). Dynamic patterns of movement of squash players of different standards in winning and losing rallies. *Ergonomics*, *37*(1), 23-29.

Hughes, M., & Robertson, C. (1998). Using computerised notational analysis to create a template for elite squash and its subsequent use in designing hand notation systems for player development. In M. Hughes & I. Maynard (Eds) *Science and Racket Sports II,* (pp. 227-234), London: E & FN Spon.

Hughes, M., Wells, J., & Mathews, K. (2000). Performance profiles at recreational, county and elite levels of woman's squash. *Journal of Human Movement Studies,* 39(2), 85-104.

Hughes, M.D., Evans, S., & Wells, J. (2001). Establishing normative profiles in performance analysis. *International Journal of Performance Analysis in Sport*, 1, 1-26.

Hughes, M.D., Cooper, S., & Nevill, A. (2002). Analysis procedures for non-parametric data from performance analysis, *International Journal of Performance Analysis in Sport*, 2, 6-20.

Hughes, M., Watts, A., White, C., & Hughes, M. T. (2006). Game structures of elite male squash under different rules. In: A. Lees, J-F. Khan & I. Maynard (Eds.) *Science and Racket Sports IV*, (pp. 227-231), London: E & FN Spon.

Hughes, M., Hughes, M. T., & Behan, H. (2007). The evolution of computerised notational analysis through the example of racket sports. *International Journal of Sports Science and Engineering*, 1, 3-28.

James, N. (2009). Performance analysis of golf: Reflections on the past and a vision of the future. *International Journal of Performance Analysis in Sport,* **9**, 2, 188-209.

James, N., Jones, N., & Hollely, C. (2002). Reliability of Selected Performance Analysis Systems in Football and Rugby, The Netherlands, Amsterdam. In *Proceedings of the 4 th International Conference on Methods and Techniques in Behavioural Research* (pp. 116-118).

James, N., & Bradley, C. (2004). Disguising ones intentions: The availability of visual cues and situational probabilities when playing against an international level squash player. In A. Lees, J.-F. Kahn & I.W. Maynard (Eds.) *Science and Racket Sports III*, (pp. 247-252), Abingdon, Oxon: Routledge.

James, N., & Patrick, J. (2004). The role of Situation Awareness in sport. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to Situation Awareness: Theory and Application*, (pp. 297-316), Aldershot: Ashgate Publishing Limited.

James, N., Taylor, J., & Stanley, S. (2007). Reliability procedures for categorical data in Performance Analysis. *International Journal of Performance Analysis in Sport,* 7(1), 1-11.

Kelso, J.A.S. (1995) *Dynamic patterns: The self-organisation of brain and behaviour.* Cambridge, MA: MIT Press.

Kermarrec, G., & Bossard, C. (2014). Defensive soccer players' decision making: A naturalistic study, *Journal of Cognitive Engineering and Decision Making*, 8(2), 187-199.

Kingsley, M., James, N., Kilduff, L.P., Dietzig, R.E., & Dietzig, B. (2006). An exercise protocol that simulates the activity patterns of elite junior squash, *Journal of Sports Sciences*, 24(12), 1291-1296.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. *Decision making in action: Models and methods*, *5*(4), 138-147.

Laban, R. (1963). *Modern Educational Dance*. 2nd edition revised by L. Ullmann. London: MacDonald and Evans. (First published 1948).

Locke, S., Colquhoun, D., Briner, M., Ellis, L., O'Brien, M., Wollstein, L., & Allen, G. (1997). Squash rackets: A review of physiology and medicine, *Sports Medicine*, 23(2), 130-138.

Mackenzie, R., & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research, Journal of Sports Sciences, 31(6), 639-676.

Macquet, A.C. (2009). Recognition within the decision-making process: A case study of expert volleyball players, *Journal of Applied Sport Psychology*, 21(1), 64-79.

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

McGarry, T., & Franks, I.M., (1996). In search of invariant athletic behaviour in competitive sport systems: an example from championship squash match-play, *Journal of Sports Sciences*, 14, 445-456.

McGarry, T., Khan, M.A., & Franks, I.M. (1999). On the presence and absence of behavioural traits in sport: An example from championship squash match-play. *Journal of Sports Sciences*, 17, 297-311.

McGarry, T., Anderson, D.I., Wallace, S.A., Hughes, M., & Franks, I.M. (2002) Sport competition as a dynamical self-organizing system. *Journal of Sports Sciences,* 20**,** 771-781.

Messersmith, L.L., & Corey, S.M. (1931). Distance Traversed by a Basketball Player. *Research Quarterly*, 2**(**2), 57-60.

Murray, S., & Hughes, M. (2001). Tactical performance profiling in elite level senior squash, In M. Hughes, I.M. Franks (eds.), *Pass.com*, (pp. 185-194). Cardiff: CPA.

Murray, S., James, N., Dineen, P., Hughes, M., & Vučković, G. (2013). The effect of changing the scoring system on game related activity in squash. In D.M. Peters & P. O'Donoghue (Eds.), *Performance Analysis of Sport IX,* (pp. 151-155), Oxford: Routledge.

Murray, S., James, N., Hughes, M.D., Perš, J., Mandeljc, R., & Vučković, G. (2016). Effects of rule changes on physical demands and shot characteristics of elite-standard men's squash and implications for training, *Journal of Sports Sciences*, 10(2), 129-140.

Murray, S., James, N., Perš, J., Mandeljc, R., & Vučković, G. (2018). Using a situation awareness approach to determine decision-making behaviour in squash. *Journal of Sports Sciences*, 36(12), 1415-1422.

Neville, T.J., & Salmon, P.M. (2016). Never blame the umpire - a review of Situation Awareness models and methods for examining the performance of officials in sport, *Ergonomics*, 59(7), 962-975.

Norušis, M.J. (2011). IBM SPSS Statistics 19 Statistical Procedures Companion. Prentice hall: Pearson.

O'Donoghue, P.G. (2004). Sources of variability in time-motion data; measurement error and within player variability in work-rate. *International Journal of Performance Analysis of Sport*, 4(2), 42-49.

O'Donoghue, P.G. (2005). Normative profiles of sport performance. *International Journal of Performance Analysis of Sport,* 5(1), 104-119.

Ofoghi, B., Zeleznikow, J., MacMahon, C., & Raab, M. (2013). Data mining in elite sports: A review and a framework, *Measurement in Physical Education and Exercise Science*, 17(3), 171-186.

Perš, J., Bon, M., Kovačič, S., Šibila, M., & Dežman, B. (2002). Observation and analysis of large-scale human motion. *Human Movement Science*, 21, 295-311.

Perš, J., Kristan, M., Perše, M., & Kovačič, S. (2008). Analysis of Player Motion in Sport Matches. In A. Baca, M. Lames, K. Lyons, B. Nebel, & J. Wiemeyer (Eds.), *Computer Science in Sport - Mission and Methods,* http://drops.dagstuhl.de/opus/volltexte/2008/1689/pdf/08372.PersJanez.Paper.1689.pdf. Accessed on 12 December 2016.

Pollard, R., (2002). Charles Reep (1904-2002): pioneer of notational and performance analysis in football. *Journal of Sports Sciences*, 20(10), 853-855.

Pulling, C., & Stenning, M. (2015). Offloads in rugby union: Northern and Southern hemisphere teams. *International Journal of Performance Analysis in Sport*, 15, 217-228.

Raab, M., & Johnson, J.G. (2007). Expertise-based differences in search and option generation strategies, *Journal of Experimental Psychology: Applied*, 13, 158-170.

Reilly, T., & Thomas, V. (1976). A motion analysis of work-rate in different positional roles in professional football match-play. *Journal of Human Movement Studies*, 2, 87-97.

Reilly, T., Morris, T., & Whyte, G. (2009). The specificity of training prescription and physiological assessment: A review. *Journal of Sports Sciences, 27*(6), 575-589.

Sanderson, F.H., & Way, K.I.M. (1977). The development of objective methods of game analysis in squash rackets. *British Journal of Sports Medicine,* 11(4), 188.

Schumaker, R.P., Solieman, O.K., & Chen, H. (2010). Sports data mining. *Integrated series in information systems*, Vol. 26. New York: Springer.

Triolet, C., Benguigui, N., Le Runigo, C., & Williams, A.M. (2013). Quantifying the nature of anticipation in professional tennis, *Journal of Sports Sciences*, 31(8), 820-830.

Vaz, L., Rooyen, M.V., & Sampaio, J. (2010). Rugby game-related statistics that discriminate between winning and losing teams in IRB and super twelve close games, *Journal of Sports Science and Medicine*, 9, 51-55.

Vučković, G., Dežman, B., Erčulj, F., Kovačič, S., & Perš, J. (2003). Comparative movement analysis of winning and losing players in men's elite squash. *Kinesiologia Slovenica*, 9(2), 74-84.

Vučković, G., Dežman, B., Erčulj, F., Kovačič S., & Perš, J. (2004). Differences between the winning and the losing players in a squash game in terms of distance covered. In A. Lees, J.F. Khan & I. Maynard (Eds). *Science and Racket Sports III*, (pp. 208-213). London: Routledge.

Vučković, G., Perš, J., James, N., & Hughes, M. (2009). Tactical use of the T area in squash by players of differing standard, *Journal of Sports Sciences*, 27(8), 863-871.

Vučković, G., Perš, J., James, N., & Hughes, M. (2010). Measurement error associated with the Sagit/squash computer tracking software, *European Journal of Sports Science*, 10(2), 129-140.

Vučković, G., & James, N. (2010). The distance covered by winning and losing players in elite squash matches. *Kinesiologia Slovenica*, *16*(1/2), 44-50.

Vučković, G., James, N., Hughes, M., Murray, S.R., Sporiš, G., & Perš, J. (2013). The effect of court location and available time on the tactical shot selection of elite squash players, *Journal of Sports Science and Medicine*, 12, 66-73.

Vučković, G., James, N., Hughes, M., Murray, S., Milanović, Z., Perš, J., & Sporiš, G. (2014). A New Method for Assessing Squash Tactics Using 15 Court Areas for Ball Locations. *Human Movement Science*, *34*, 81-90.

Williams, A.M., Davids, K., & Williams, J.G. (1999). Visual perception and action in sport. London: Spon press.

Winter, E.M., Abt, G.A., & Nevill, A.M. (2014). Metrics of meaningfulness as opposed to sleights of significance, *Journal of Sports Sciences*, *32*(10), 901-902.

APPENDICES

Appendix 1.1 Ethical approval Letter from Professional Squash Association

Date 14.07.2009

Re: Ethical Approval Application for PhD study

Dear Stafford,

I am pleased to inform you that on behalf of the PSA we are happy for you to utilise video footage of our players captured during the British Grand Prix Squash Championships (during the seasons 2010 and 2011) for research and data capture towards your PhD studies in squash profiling.

The PSA wishes you luck with your endeavours

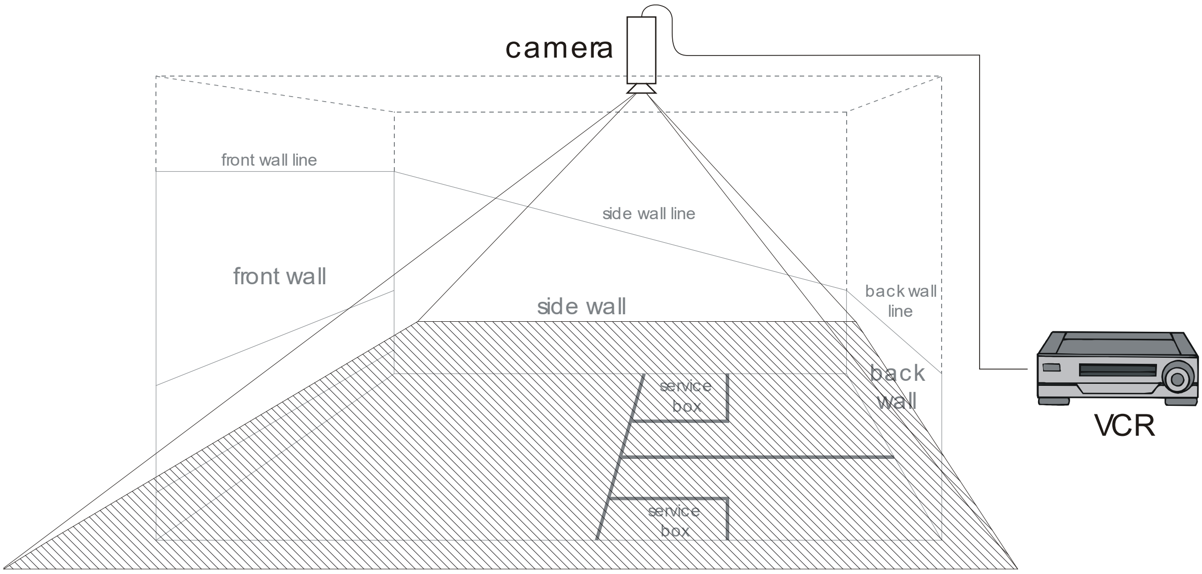
Best Regards,

Lee Beachill

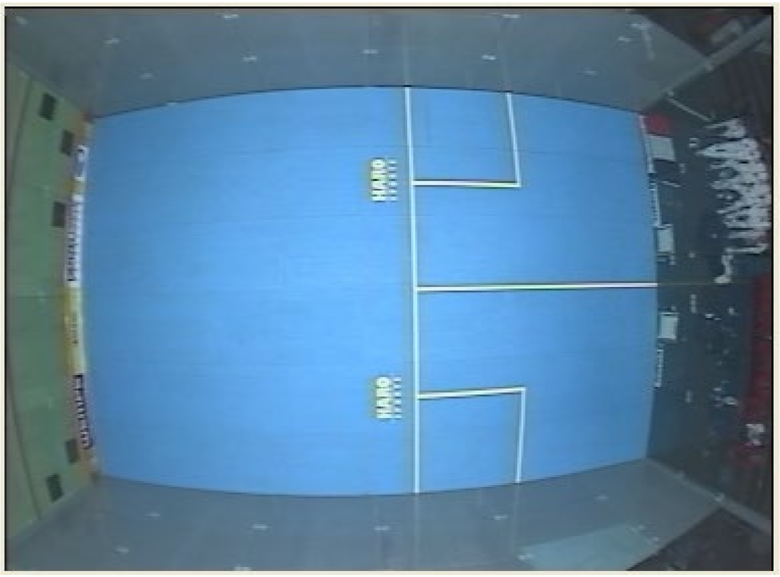
Lee Beachill
Chief Operating Officer
Professional Squash Association

Appendix 1.2 Visualisations and screenshots of SAGIT/Squash

Appendix 1.2.1 Camera placement



Appendix 1.2.2 View of court from overhead camera



Appendix 1.2.3 Screenshot of tracking screen (operator manually adjusts tracking if automatic tracking makes errors)

Video project: D:\00656981\SagitSquash\BritishGP2011\MatthewShabana\MatthewShabana1...

Video streams and players overview

1 - Matthew Nick, Winner
2 - Shabana Amr, Loser

PLAYER STATUS:

○ Track
● Don't touch
○ Stay there

00h:05m:08s:00f 7700

Tracker is ready.

Appendix 1.2.4 Screenshot of annotation screen (input shot information)

Video project: D:\00656981\SagitSquash\BritishGP2011\MatthewShabana\MatthewShabana1...

Video streams and players overview

1 - Matthew Nick, Winner
2 - Shabana Amr, Loser

■ Active phase

Team_state: Pasive phase

Height: 1.50 m

Shot: Do not touch

Effect: Do not touch

Side: Do not touch

00h:01m:00s:07f 1507

Annotator is ready.

Appendix 3.1 Ethical Approval Letter for all studies

Middlesex
University

School of Health and Social Sciences
The Archway Campus
Furnival Building
10 Highgate Hill
London N19 5LW
Tel: +44 (0)20 8411 5000
www.mdx.ac.uk

To: Stafford Murray
PhD

Ref: HSESC/10/792 Date: 8th August 2011

Dear Stafford

Re: Stafford Murray - Ethics Application 792. *'Enhancing Performance Profiling Methodologies in Elite Male Squash using automated tracking technology (SAGIT system).* Supervisors, Nic James, Joe Taylor & Goran Vuckovic

Thank you for the information which you submitted to the ethics sub-committee (health studies) regarding the above project.

I can confirm that since your project was categorised as A1 and does not formally require ethical approval, your application will be logged on our database for information only.

Yours sincerely

Ms Dympna Crowley

Chair of Ethics Sub-committee (Health Studies)

Appendix 3.2 Data for reliability tests to assess accuracy of determination of court cell ball played from

		SAGIT/Squash															
	Area	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total
Operator	1	814	20	0	0	1	17	3	2	0	0	0	0	0	0	0	857
	2	56	276	11	0	0	0	11	2	0	0	0	0	0	0	0	356
	3	0	0	80	0	0	0	0	2	0	0	0	0	0	0	0	82
	4	0	1	15	170	10	0	0	0	6	0	0	0	0	0	0	202
	5	1	0	0	14	279	0	0	1	1	12	0	0	0	0	0	308
	6	0	0	0	0	0	168	17	0	0	0	2	0	0	0	0	187
	7	2	2	0	0	0	15	191	15	1	0	0	1	0	0	0	227
	8	0	0	2	1	0	0	1	178	2	0	0	0	0	0	0	184
	9	1	0	0	9	1	0	0	6	129	9	0	0	0	2	0	157
	10	0	0	0	0	6	0	0	0	14	100	0	0	0	0	5	125
	11	0	0	0	0	0	0	0	0	0	0	69	1	0	0	0	70
	12	0	0	0	0	0	2	0	0	0	0	7	27	4	0	0	40
	13	0	0	0	0	0	0	0	1	0	0	0	0	34	0	0	35
	14	0	0	0	0	0	0	0	0	0	0	0	0	4	27	6	37
	15	0	0	0	0	0	1	0	0	0	0	0	0	0	3	36	40
	Total	874	299	108	194	297	203	223	207	153	121	78	29	42	32	47	2907

Appendix 3.3 Data for inter-operator reliability tests to determine shot information

Appendix 3.3.1 Data for shot type

		Operator 2																							
Shot Type (Abrev.)		BA	BD	C	CD	CK	CS	CW	D	DW	LB	LC	SD	SK	VBA	VBD	VC	VCD	VCK	VCS	VCW	VS	VSD	VSK	VSW	Total
Operator 1	BA	45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45
	BD	0	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	35
	C	1	0	137	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	138
	CD	0	0	0	18	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	20
	CK	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
	CS	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17
	CW	0	0	2	0	0	0	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54
	D	0	0	0	0	0	0	0	228	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	233
	DW	0	0	0	0	0	0	0	3	147	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	151
	LB	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23
	LC	0	0	0	0	0	0	0	0	0	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	48
	SD	0	0	0	0	0	0	0	0	0	0	0	130	1	0	0	0	0	0	0	0	0	0	0	0	131
	SK	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	9
	VBA	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
	VBD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	1	0	4
	VC	0	0	4	0	0	0	0	0	0	1	0	0	0	0	0	31	0	0	0	0	0	0	0	0	36
	VCD	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	9
	VCK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	6	0	0	0	0	0	0	7
	VCS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	5
	VCW	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	27	0	0	0	0	29
	VS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39	0	0	0	39
	VSD	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	49	0	0	50
	VSK	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	15	0	17
	VSW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	18
Total		46	35	143	18	4	18	52	231	149	26	48	131	13	4	4	32	8	6	6	28	39	49	16	18	1124

Appendix 3.3.2 Data for shot side

Shot Side	BH	FH	Total
BH	609	25	634
FH	19	471	490
Total	496	388	1124

Appendix 3.3.3 Data for court side

Court Side	L	R	Total
L	647	12	659
R	20	445	465
Total	540	344	1124

Appendix 3.3.4 Data for foot used

Foot Used	B	L	R	Total
B	276	17	33	326
L	22	245	6	273
R	15	7	503	525
Total	224	211	449	1124

Appendix 3.4 Data for intra-operator reliability tests to determine shot information

Appendix 3.4.1 Data for shot type

		T2																								
Shot Type (Abrev.)		BA	BD	C	CD	CK	CS	CW	D	DW	LB	LC	SD	SK	VBA	VBD	VC	VCD	VCK	VCS	VCW	VS	VSD	VSK	VSW	Total
T1	BA	29	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30
	BD	1	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	29
	C	0	0	93	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94
	CD	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
	CK	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	CS	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	CW	0	0	2	0	0	0	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34
	D	0	0	0	0	0	0	0	203	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	209
	DW	0	0	0	0	0	0	0	3	125	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	128
	LB	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
	LC	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	33
	SD	0	0	0	0	0	0	0	0	0	0	0	107	1	0	0	0	0	0	0	0	0	0	0	0	108
	SK	0	0	0	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	5
	VBA	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
	VBD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	VC	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	35
	VCD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	5
	VCK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	0	0	0	0	0	0	5
	VCS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	4
	VCW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	23
	VS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	24
	VSD	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	28	0	0	29
	VSK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	8
	VSW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	11
Total		30	29	99	11	1	7	32	207	127	12	33	108	8	2	1	32	5	4	4	23	24	28	8	11	848

Appendix 3.4.2 Data for shot side

Shot Side	BH	FH	Total
BH	493	8	501
FH	3	380	383
Total	496	388	884

Appendix 3.4.3 Data for court side

Court Side	L	R	Total
L	532	4	536
R	8	340	348
Total	540	344	884

Appendix 3.4.4 Data for foot used

Foot Used	B	L	R	Total
B	199	7	10	216
L	14	203	1	218
R	11	1	438	450
Total	224	211	449	884