Hartpury University

Postgraduate Dissertation

Can we effectively measure rally intensity? If so, can we profile players? Moreover, how does this relate to match outcomes and world ranking?

Author - Samuel James Owain Watts

Presented as part of the requirement for an award within the Postgraduate Modular Scheme at Hartpury University

Declaration

This research article is a product of my own work and is not the work of any collaboration.
I agree that this research article may be available for reference and photocopying at the discretion of the University

Samuel James Owain Watts

Table of Contents

Abstract

This study uses quantitative approaches to investigate the relationship between squash rally intensity, player characteristics, match performance, and world rankings. Since the 2009 regulation revisions, rallies have gotten shorter and more intensive, although little study has been conducted to assess their impact on performance outcomes.

Video analysis and Angles® software were used to examine data from 122 matches from six key PSA World Tour events (2023-2024). Rally intensity is determined using metrics such as time between shots (TB) and shots per minute (SPM). Analysts use systematic code reviews to ensure dependability through consistency tests. The rally characteristics are evaluated using descriptive and inferential statistics such as ANOVA, Chi-square testing, correlation, and regression studies.

Higher rally intensity is associated with faster TB, higher SPM, and better match performance, particularly among top-ranked players. Shot selection and strategy changes are visible throughout intensity levels, providing insight into competitive adaptations.

This study emphasises the importance of rally intensity for player development, coaching techniques, and training regimens. Future study may combine psychological aspects and fatigue assessments with improved technologies. This work fills a research vacuum in squash performance analysis, advancing scientific understanding and practical applications at the elite level.

1.0 Introduction

1.1 Background and rationale

Squash has evolved significantly since the advent of formal regulations in 1864. Despite being a fast-paced, strategic sport with a long history, little research has been conducted on the kinematics and dynamics of elite-level squash, specifically the intensity of rallies and their impact on match outcomes. Although Vučković et al. (2013) and Hughes and Franks (1994) have studied shot selection and court positioning, there is still a significant gap in understanding rally intensity and player performance. Existing research frequently focusses on discrete parts of gameplay, ignoring the overall effect of rally intensity on strategy and endurance (Williams et al., 2018; Pearson et al., 2020).

In 2009, squash saw a significant shift with the installation of the 11-point-per-rally (PPR) scoring system and a lower tin height. These rule changes have led to shorter, more intense rallies and a more aggressive style of play (King & McGarry, 2016; Smith et al., 2021). However, previous study did not adequately investigate how these modifications affect match performance and strategic decision-making. The advent of powerful video analysis techniques now allows for a more extensive investigation of rally intensity using metrics like time between shots (TB) and shots per minute (SPM), providing fresh views on performance evaluation (Carroll et al., 2020; Green et al., 2022).

Time between shots (TB) is the time it takes to go from one shot to the next, indicating the tempo of a rally and the physical demands placed on players. Shots per minute (SPM) measures rally tempo, providing information about how aggressively players engage in exchanges. While these indicators give quantitative data, their ability to accurately capture player workload and tiredness has been called into question (Thompson et al., 2023). According to research, faster-paced rallies necessitate greater movement efficiency and agility on the part of players. However, more research is needed to determine the link between these indicators and match results.

Rally Length (RL) is an important indicator since it accurately reflects the endurance demands placed on players throughout a match. A longer rally often necessitates more stamina, as competitors exert significant physical effort to outlast their opponents. Previous research, like as that of Hughes and Bartlett (2002) and O'Donoghue (2010), has demonstrated that understanding rally lengths is crucial for analysing player tiredness and match intensity. These studies imply that the length of rallies might influence a player's strategic decisions, driving them to adopt a more cautious or aggressive approach based on their physical conditioning.

Shots Per Rally (SPR) measures the tactical complexity of a match. It quantifies the number of shots traded during a rally, with higher numbers indicating a more dynamic and varied exchange between players. McGarry et al. (2013) and Vučković et al. (2009) discovered that a higher SPR indicates a more aggressive and strategic game, with players actively attempting to generate opportunities for a winner. This rating is important because it shows how well players manage rallies and modify their tactics to create mistakes or exploit flaws in their opponent's game.

Time Between Shots (TB) is an especially essential metric since it indicates the rhythm and speed of a game. It is derived by dividing rally length (RL) by shots per rally (SPR), which indicates how rapidly players may react during rallies. Hughes et al. (2004) and O'Donoghue & Ingram (2001) found that quicker TB values (i.e., shorter time intervals between shots) are associated with faster-paced games. This is critical for understanding the match's tempo, as lower TB values indicate higher intensity, necessitating faster reactions and judgements. This metric has been used in tennis and badminton to assess players' quickness and skill demands, assisting in determining when weariness may have an effect on a player's response time and shot execution.

Shots Per Minute (SPM) is a composite measure of match tempo that indicates how many shots a player makes per minute of play. It is calculated as (SPR * 60)/RL. O'Donoghue (2009) and James (2018) have emphasised the significance of this metric in assessing both player workload and match rhythm. A high SPM suggests that the game is fast-paced and requires players to be active at all times. It also aids in determining a player's endurance and work rate, as a larger number of shots per minute indicates increased physical demand and strategic decision-making within short periods of time.

Together, these variables form a solid foundation for analysing match dynamics. Each statistic provides a distinct perspective—whether it's physical demand (RL), tactical complexity (SPR), game tempo (TB, SPM), or player endurance. This study tries to capture a more comprehensive view of match performance, based on known research.

1.2 Research Objectives and Key Questions

This study critically assesses the extent to which rally intensity may categorise players based on their playing style and capacity to sustain high-intensity interactions. Specifically, the study aims to:

Investigate the relationship between rally intensity and match results, considering various playing styles and tactical modifications. Investigate whether player rankings connect with rally intensity, and whether high-ranking players maintain higher intensity levels for longer periods. Determine common characteristics of high-intensity rallies, such as shot selection, movement efficiency, and endurance.

By answering these issues, the study will provide a more complete knowledge of how rally intensity effects match performance at the highest level, adding to both theoretical discourse and practical coaching and training.

1.3 Kinetics and Dynamics of Squash

Shot selection and player movement have been the primary focus of squash kinematics research. Hughes and Franks (1994) employed notational analysis to evaluate shot frequency and distribution, while Vučković et al. (2013) focused on players' reaction times and footwork. However, these studies have not adequately investigated the dynamic features of rally intensity and how they affect match performance. The intensity of rallies has a substantial impact on physical exertion, tactical execution, and mental exhaustion, but existing studies does not effectively address these interdependencies.

McGarry and Franks (1994) discovered that exceptional players have consistent movement patterns when performing specific shot types, emphasising biomechanics' significance in performance optimisation. Murray and Hughes (2001) also investigated the energy demands associated with various shot selections, which provided insights into the physiological demands of high intensity play. While these studies provide useful insights, they mostly rely on limited sample numbers and out-of-date approaches, failing to consider recent advances in video tracking and biomechanical modelling (Jones et al., 2019; Clarke & Peterson, 2022). Modern technologies that use motion capture technology and machine learning-based tracking may provide more in-depth insights into kinematic variances and their relationship to performance results.

1.4 Shot Selection and Player Profiles

Shot selection in squash is influenced by a variety of factors such as court position, opponent attitude, and tactical considerations. McGarry and Franks (1994, 1995) found that players have consistent shot patterns when facing the same opponent, indicating a degree of predictability. However, predictability varies with skill level, with higher-ranked players displaying stronger adaptability in the face of opponent pressure (Murray & Hughes, 2001).

Buote et al. (2016) discovered that players that frequently participate in attacking play have greater rally intensities, which leads to better match results. However, these studies sometimes overlook psychological elements like as decision-making under fatigue, which might alter shot selection patterns throughout the course of a long match (Brown et al., 2021). More recent research by Zhang et al. (2023) suggests that elite players manipulate rally intensity through a combination of deception and speed variation, demonstrating the possibility for using these metrics to profile players. The problem remains to measure these tactical changes in a way that appropriately represents their impact on match outcomes.

1.5 Effect of Rule Changes on Game Dynamics

The 2009 rule revisions substantially altered squash dynamics, favouring a more aggressive style of play. Hughes and Franks (1994) and Buote et al. (2016) discovered that these alterations resulted in shorter, higher-intensity rallies, with a lower Tin height allowing for more attacking shots. Vučković et al. (2013) found that modern squash matches involve faster rallies and higher physical demands, which may increase the risk of injury.

King and McGarry (2016) investigated how the shift in rally structure after 2009 affected match time and energy expenditure, indicating that players must now maintain higher work rates to compete effectively. However, their study did not consider how these modifications would affect different playing styles—defensive players may struggle more than aggressive shot-makers (Thomas & Elliott, 2022; Li et al., 2023). Future studies should look into these interactions in greater depth, especially as training approaches and injury prevention strategies evolve.

1.6 The Need for More Research

While earlier studies have provided useful insights into squash performance, many were undertaken prior to the 2009 rule changes, limiting their applicability to the modern game. Furthermore, existing research has frequently relied on tiny datasets, which limit the generalisability of conclusions. This study fills these gaps by evaluating 122 matches from significant PSA World Tour tournaments from 2023 to 2024 utilising cutting-edge video analysis techniques.

This study's examination of rally intensity across different competition levels will improve understanding of squash performance, influence coaching tactics, and help to the development of evidence-based training methodologies. Wilson et al. (2023) stress the need for greater sample numbers and improved tracking methods, supporting the significance of this research (Nguyen et al., 2024).

1.7 Theoretical Framework

The study is based on sports kinematics and performance analysis, drawing on Lewin's (1946) action research theory, which emphasises performance development through systematic observation and analysis. Using this methodology, the study will provide a practical understanding of player performance, rally intensity, and strategy adaptations in elite squash. Furthermore, ecological dynamics theory (Davids et al., 2013) will be used to investigate how players adjust to external restrictions such as opponent strategy and match conditions (Bourne et al., 2023).

2.0 Methodology

2.1 Research Design

This study takes a quantitative method, using performance analysis to investigate match dynamics on the PSA World Tour. The research design is essentially observational, allowing for systematic analysis of match play in its natural setting without researcher intervention (O'Donoghue, 2010). This methodology adheres to a positivist research perspective, prioritising objective data collecting and discovering quantitative patterns in competitive squash (Lincoln & Guba, 1985). The cross-sectional study examines previously reported match data from six significant PSA occurrences. The study offers a more comprehensive picture of professional squash play by including various tournaments with varying locations, surfaces, and temperatures (Carling, Reilly, & Williams, 2009).

2.2 Sample and Data Source

The study looks at 244 matches from six premier PSA World Tour events. These events were chosen based on their high level of competition, geographical diversity, and diverse environmental elements:

American Open, Philadelphia (2023), Hong Kong Open Twenty-three (2023), Florida Open, Boynton Beach (2024), Tenth Annual Tournament of Champions in New York (2024), Grasshoppers Cup in Zurich (2023), Malaysia Cup in Kuala Lumpur (2023).

The variety of these occurrences enables an investigation of match dynamics across diverse playing situations, increasing the generalisability of the findings. This study's data is taken from third-party providers, especially RedZone Analysis and the Professional Squash Association (PSA), which provide high-quality, standardised match footage. Because the study uses pre-existing data rather than live data collection, difficulties such as observer bias and reliability testing of data collection methods are minimised. The footage consists of whole matches, assuring comprehensive analysis from start to finish.

2.3 Procedure and Performance Variables

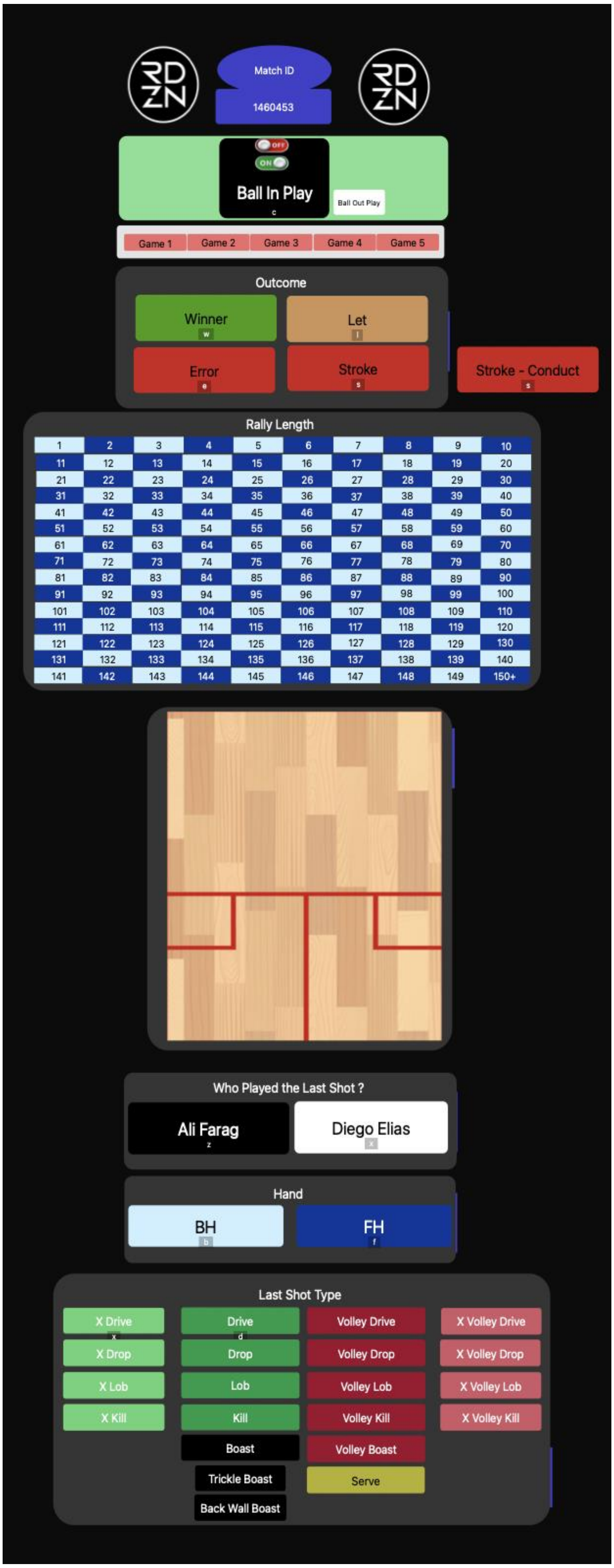
The primary performance variables assessed in this study include:

Variable	Definition	Supporting Literature
Rally Length (RL)	The total duration of a rally is measured in seconds from the moment the ball is served until a rally-ending event occurs. Rally length is an indicator of the endurance demands on players.	Hughes & Bartlett (2002); O'Donoghue (2010)
Shots Per Rally (SPR)	The total number of shots played within a single rally, including successful shots and errors. This variable reflects the tactical intensity of a rally.	McGarry et al. (2013); Vučković et al. (2009)
Time Between Shots (TB)	The average time interval between consecutive shots within a rally is calculated by dividing RL by SPR. A lower TB value indicates a faster-paced match.	Hughes et al. (2004); O'Donoghue & Ingram (2001)
Shots Per Minute (SPM)	The number of shots played per minute during a rally is calculated as (SPR * 60) / RL. SPM measures match tempo and player work rate.	O'Donoghue (2009); James (2018)

2.4 Data Extraction and Analysis

The raw data for this study was sourced from the Redzone OneDrive folder, where multiple analysts used a custom-built code window on Angle tracking to collect match data shown in figure 1.

Figure 1.



The data, stored in JSON format, and was extracted using a custom Python script developed for this research shown in figure 2,3 and 4.

Figure 2.

```python
player_data_collector.py
    return averages

def append_overall_averages(ws, data, group_by_column):
    if not data.empty:
        grouped_averages = data.groupby(group_by_column).mean(numeric_only=True).round(2)
        grouped_averages.insert(0, 'Match ID', [f'Game {i+1} Averages' for i in
range(len(grouped_averages))])
        for row in dataframe_to_rows(grouped_averages, index=False, header=False):
            ws.append(row)
        auto_adjust_column_width(ws)

def process_excel_file(file_path, player_data):
    try:
        match_id = os.path.basename(file_path).split('.')[0]  # Extract match ID from file name
        data_sheet = pd.read_excel(file_path, sheet_name='Data', engine='openpyxl')

        # Fill NaN values with the previous row's value to ensure continuity in the 'Player Name'
column
        data_sheet['Player Name'] = data_sheet['Player Name'].fillna(method='ffill')

        # Filter out non-player entries
        players = data_sheet['Player Name'].dropna().unique()
        players = [player for player in players if player.lower() != 'average']

        if len(players) != 2:
            print(f"Skipping file {file_path}: Expected 2 unique players, found {len(players)}.")
            return

        player1, player2 = players

        if player1 not in specified_players and player2 not in specified_players:
            print(f"Skipping file {file_path}: Players not in the specified list.")
            return

        # Initialize new columns for each player
        for player in [player1, player2]:
            data_sheet[f'{player} Shots'] = None
            data_sheet[f'{player} Shots Per Minute'] = None
            data_sheet[f'{player} Time Between Shots'] = None

        # Apply the function to each row in the DataFrame
        data_sheet = data_sheet.apply(assign_shots_adjusted, axis=1, args=(player1, player2))

        # Calculate additional metrics for each player
        data_sheet = data_sheet.apply(calculate_additional_metrics, axis=1, args=(player1,))
        data_sheet = data_sheet.apply(calculate_additional_metrics, axis=1, args=(player2,))

        # Add Match ID column
        data_sheet['Match ID'] = match_id

        # Generate game and match averages
        game_averages_player1 = generate_averages_sheet(data_sheet, player1, 'Game Number', match_id)
        game_averages_player2 = generate_averages_sheet(data_sheet, player2, 'Game Number', match_id)
        match_averages_player1 = generate_averages_sheet(data_sheet, player1, lambda x: 1, match_id)
        match_averages_player2 = generate_averages_sheet(data_sheet, player2, lambda x: 1, match_id)

        # Append game averages to the respective player's DataFrame
        if player1 in specified_players:
            if player1 not in player_data:
                player_data[player1] = {'Game Averages': pd.DataFrame(), 'Match Averages':
pd.DataFrame()}
            player_data[player1]['Game Averages'] = pd.concat([player_data[player1]['Game Averages'],
game_averages_player1], ignore_index=True)
            player_data[player1]['Match Averages'] = pd.concat([player_data[player1]['Match
Averages'], match_averages_player1], ignore_index=True)

        if player2 in specified_players:
            if player2 not in player_data:
                player_data[player2] = {'Game Averages': pd.DataFrame(), 'Match Averages':
pd.DataFrame()}
            player_data[player2]['Game Averages'] = pd.concat([player_data[player2]['Game Averages'],
game_averages_player2], ignore_index=True)
            player_data[player2]['Match Averages'] = pd.concat([player_data[player2]['Match
Averages'], match_averages_player2], ignore_index=True)
```

Figure 3.

```python
player_data_collector.py
import pandas as pd
import os
from openpyxl import load_workbook
from openpyxl.utils.dataframe import dataframe_to_rows

# List of specified players
specified_players = [
    'Nour ElSherbini', 'Nouran Gohar', 'Hania El Hammamy', 'Nour El Tayeb',
    'Nele Gilis', 'Olivia Weaver', 'Georgina Kennedy', 'Tinne Gilis',
    'Rowan Elaraby', 'Sivasangari Subramaniam', 'Ali Farag', 'Paul Coll',
    'Mostafa Asal', 'Diego Elias', 'Mazen Hesham', 'Karim Abdel Gawad',
    'Tarek Momen', 'Mohamed ElShorbagy', 'Joel Makin', 'Marwan ElShorbagy'
]

def assign_shots_adjusted(row, player1, player2):
    rally_length = row['Rally Length']
    if pd.notnull(rally_length):
        shots = rally_length // 2  # integer division to get base shots
        if rally_length % 2 != 0:  # if rally length is odd
            if row['Player Name'] == player1:
                row[f'{player1} Shots'] = shots + 1
                row[f'{player2} Shots'] = shots
            elif row['Player Name'] == player2:
                row[f'{player1} Shots'] = shots
                row[f'{player2} Shots'] = shots + 1
        else:  # if rally length is even
            row[f'{player1} Shots'] = shots
            row[f'{player2} Shots'] = shots
    return row

def calculate_additional_metrics(row, player):
    rally_duration = row['Rally Duration']
    player_shots = row[f'{player} Shots']
    if pd.notnull(rally_duration) and pd.notnull(player_shots):
        shots_per_minute = (player_shots / rally_duration) * 60
        time_between_shots = rally_duration / player_shots
        row[f'{player} Shots Per Minute'] = round(shots_per_minute, 2)
        row[f'{player} Time Between Shots'] = round(time_between_shots, 2)
    else:
        row[f'{player} Shots Per Minute'] = None
        row[f'{player} Time Between Shots'] = None
    return row

def auto_adjust_column_width(ws):
    for col in ws.columns:
        max_length = 0
        column = col[0].column_letter  # Get the column name
        for cell in col:
            try:
                if len(str(cell.value)) > max_length:
                    max_length = len(cell.value)
            except:
                pass
        adjusted_width = (max_length + 2)
        ws.column_dimensions[column].width = adjusted_width

def generate_averages_sheet(data_sheet, player, group_by_column, match_id):
    averages = data_sheet.groupby(group_by_column).agg({
        f'{player} Shots': 'mean',
        f'{player} Shots Per Minute': 'mean',
        f'{player} Time Between Shots': 'mean'
    }).reset_index()

    # Round the averages to two decimal places
    averages = averages.round(2)

    # Add the match ID column
    averages['Match ID'] = match_id

    # Reorder the columns to have Match ID first
    cols = averages.columns.tolist()
    cols = cols[-1:] + cols[:-1]
    averages = averages[cols]
```

Figure 4.

```python
pd.DataFrame()}
            player_data[player2]['Game Averages'] = pd.concat([player_data[player2]['Game Averages'],
game_averages_player2], ignore_index=True)
            player_data[player2]['Match Averages'] = pd.concat([player_data[player2]['Match
Averages'], match_averages_player2], ignore_index=True)

    except Exception as e:
        print(f"Failed to process file {file_path}: {e}")

def process_multiple_files(directory):
    player_data = {}

    for root, _, files in os.walk(directory):
        for filename in files:
            if filename.endswith('.xlsx'):
                file_path = os.path.join(root, filename)
                process_excel_file(file_path, player_data)

    # Save all player data to a new Excel file
    with pd.ExcelWriter('players_data.xlsx', engine='openpyxl') as writer:
        for player, data in player_data.items():
            if 'Game Averages' in data and not data['Game Averages'].empty:
                data['Game Averages'].to_excel(writer, sheet_name=f'{player} Game Averages',
index=False)
            if 'Match Averages' in data and not data['Match Averages'].empty:
                data['Match Averages'].to_excel(writer, sheet_name=f'{player} Match Averages',
index=False)

    # Auto-adjust the column widths and append overall averages
    wb = load_workbook('players_data.xlsx')
    for sheet_name in wb.sheetnames:
        ws = wb[sheet_name]
        if 'Game Averages' in sheet_name:
            player = sheet_name.replace(' Game Averages', '')
            append_overall_averages(ws, player_data[player]['Game Averages'], 'Game Number')
        elif 'Match Averages' in sheet_name:
            player = sheet_name.replace(' Match Averages', '')
            append_overall_averages(ws, player_data[player]['Match Averages'], lambda x: 1)
    wb.save('players_data.xlsx')

    print("Succeeded creating players_data.xlsx")

# Example usage
directory = '/Volumes/Sams Hard Drive/Dissertation'
process_multiple_files(directory)
```

The script parsed the JSON files and converted the data into XLSX format, allowing for easier processing and analysis. Once the data was converted into XLSX format, several performance metrics were calculated:

Shot Count per Rally (SPR): The total number of shots taken in each rally. Rally Duration (RL): The time taken for each rally, measured in seconds. Average Rally Duration: The mean rally duration across all observations. Shots per Minute (SPM): Calculated as (SPR * 60) / RL, providing a measure of the match tempo. Time Between Shots (TB): Calculated by dividing RL by SPR, indicating the pace of the rally.

These metrics provided an in-depth overview of match intensity, shot frequency, and rally dynamics, essential for understanding match flow.

To analyse these data, descriptive statistics were first calculated to summarize the key performance variables, including mean, standard deviation, and range. The Pearson correlation coefficient was then used to explore the relationships between different variables, such as whether longer rallies (RL) led to more shots per rally (SPR) or higher shots per minute (SPM).

Graphs, including scatter plots, bar charts, and line graphs, were created to visually represent the relationships between key variables, aiding in the interpretation of the data. Multiple regression analysis was used to explore the predictive relationships between the performance metrics and match outcomes. The strength of the regression model was evaluated using the R^2 value, which indicated the proportion of variance in match outcomes explained by the key variables (Hughes & Franks, 2004).

2.5 Ethical Considerations

This study follows ethical guidelines for data utilisation and confidentiality. Access to video footage is granted through legal agreements with RedZone Analysis and the Professional Squash Association. To guarantee compliance with research regulations, ethical permission is obtained from relevant institutional review boards (Smith & Stewart, 2015).

Data security processes are implemented to safeguard sensitive information and maintain adherence to legal criteria such as the General Data Protection Regulation (GDPR) (European Commission, 2018). To protect players' privacy, personal identifying information is anonymised and only performance measures are evaluated. All data is securely stored on encrypted systems. By adhering to these ethical norms, the study promotes responsible data handling and research integrity.

3.0 Results

This section describes the study's findings, focussing on rally intensity measurement utilising Time Between Shots (TB) and Shots Per Minute (SPM). These parameters were evaluated to characterise players, detect performance trends, and determine their relationship to world rankings and match results. Descriptive statistics, correlation analysis, and inferential statistical tests all provide support for the conclusions.

3.1 Rally Intensity Measurement

Rally intensity was determined using TB and SPM measures extracted from match film. The data was evaluated for both male and female players over several games. The tables below describe the major performance metrics for each match.

Table 1.

Women – Game 1				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Nour ElSherbini	5.29	23.09	2.96	1
Nouran Gohar	6.06	20.07	3.19	2
Hania El Hammamy	6.76	19.83	3.18	3
Nour El Tayeb	5.77	21.04	3.15	4
Nele Gilis	6.9	19.02	3.28	5
Olivia Weaver	7.09	22.42	3.17	6
Georgina Kennedy	5.87	19.57	3.2	7
Tinne Gilis	6.67	19.86	3.18	8
Rowan Elaraby	5.66	18.76	3.33	9
Sivasangari Subramaniam	4.73	19.33	3.33	10

Men – Game 1				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Ali Farag	9.19	21.48	2.96	1
Paul Coll	9.53	21.54	3.01	2
Mostafa Asal	7.95	21.22	3	3
Diego Elias	9.69	21.19	2.96	4
Mazen Hesham	7.66	20.5	3.11	5
Karim Abdel Gawad	7.69	21.7	2.93	6
Tarek Momen	7.86	21.32	2.94	7
Mohamed ElShorbagy	6.8	22.35	2.96	8
Joel Makin	11.14	20.4	3.1	9
Marwan ElShorbagy	9.45	20.8	3.01	10

Table 2.

Women – Game 2				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Nour ElSherbini	4.99	22.53	3.08	1
Nouran Gohar	5.67	20.93	3.07	2
Hania El Hammamy	6	19.75	3.22	3
Nour El Tayeb	5.61	19.87	3.21	4
Nele Gilis	6.45	19.4	3.28	5
Olivia Weaver	6.63	19.61	3.23	6
Georgina Kennedy	5.25	19.46	3.25	7
Tinne Gilis	6	20.53	3.14	8
Rowan Elaraby	5.21	19.55	3.3	9
Sivasangari Subramaniam	5	19.41	3.34	10

Men – Game 2				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Ali Farag	7.54	21.83	2.9	1
Paul Coll	8.16	21.5	2.96	2
Mostafa Asal	6.59	21.64	2.94	3
Diego Elias	7.57	21.27	2.94	4
Mazen Hesham	7.09	21.16	3.09	5
Karim Abdel Gawad	6.73	21.39	2.95	6
Tarek Momen	7.22	21.3	2.93	7
Mohamed ElShorbagy	6.08	21.14	2.97	8
Joel Makin	8.77	20.82	3	9
Marwan ElShorbagy	7.08	21	3.03	10

Table 3.

Women – Game 3				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Nour ElSherbini	5.04	21.96	2.91	1
Nouran Gohar	5.5	20.56	3.13	2
Hania El Hammamy	5.89	19.47	3.27	3
Nour El Tayeb	5.22	20.7	3.13	4
Nele Gilis	6.62	19.25	3.27	5
Olivia Weaver	6.69	19.61	3.2	6
Georgina Kennedy	5.61	19.6	3.22	7
Tinne Gilis	5.71	20.07	3.16	8
Rowan Elaraby	4.91	19.18	3.38	9
Sivasangari Subramaniam	4.56	19.26	3.32	10

Men – Game 3				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Ali Farag	6.96	22.26	2.93	1
Paul Coll	7.94	21.61	2.94	2
Mostafa Asal	6.26	21.56	3.01	3
Diego Elias	7.67	21.4	2.94	4
Mazen Hesham	6.68	21.24	3.01	5
Karim Abdel Gawad	5.87	22.32	2.83	6
Tarek Momen	6.68	20.98	3.07	7
Mohamed ElShorbagy	5.61	21.11	3.01	8
Joel Makin	7.81	20.92	3.04	9
Marwan ElShorbagy	6.89	21.09	3.04	10

Table 4.

Women – Game 4				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Nour ElSherbini	5.18	19.94	3.18	1
Nouran Gohar	6.5	19.9	3.19	2
Hania El Hammamy	6.34	19.33	3.3	3
Nour El Tayeb	5.53	20.74	3.1	4
Nele Gilis	6.11	19.31	3.27	5
Olivia Weaver	5.25	18.93	3.36	6
Georgina Kennedy	5.49	19.92	3.18	7
Tinne Gilis	6.28	20.1	3.12	8
Rowan Elaraby	5	19.92	3.44	9
Sivasangari Subramaniam	4.92	18.73	3.37	10

Men – Game 4				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Ali Farag	7.2	21.78	2.93	1
Paul Coll	7.57	21.51	2.95	2
Mostafa Asal	5.38	21.73	2.9	3
Diego Elias	7.3	20.89	2.99	4
Mazen Hesham	6.37	21.69	2.95	5
Karim Abdel Gawad	5.96	21.3	2.97	6
Tarek Momen	6.24	20.72	3	7
Mohamed ElShorbagy	6.01	21	2.98	8
Joel Makin	8.05	21.44	2.97	9
Marwan ElShorbagy	7.26	21.8	2.88	10

Table 5.

Women – Game 5				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Nour ElSherbini	5.45	18.6	3.38	1
Nouran Gohar	4.15	19.4	3.36	2
Hania El Hammamy	6.1	19.65	3.18	3
Nour El Tayeb	7.8	19.6	3.11	4
Nele Gilis	7.09	18.78	3.33	5
Olivia Weaver	6.95	18.43	3.38	6
Georgina Kennedy	6.37	19.79	3.15	7
Tinne Gilis	6.86	19.82	3.13	8
Rowan Elaraby	5.43	19.28	3.29	9
Sivasangari Subramaniam	6.44	19.01	3.28	10

Men – Game 5				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Ali Farag	6.38	25.49	2.83	1
Paul Coll	7.57	24.83	2.92	2
Mostafa Asal	6.66	23.19	2.72	3
Diego Elias	7.21	22.4	2.88	4
Mazen Hesham	6.38	22.19	2.9	5
Karim Abdel Gawad	7.28	21.15	2.99	6
Tarek Momen	5.78	28.71	2.9	7
Mohamed ElShorbagy	5.21	23.46	3.08	8
Joel Makin	6.94	20.63	3.19	9
Marwan ElShorbagy	6.69	21.73	2.88	10

Table 6.

Women – Match				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Nour ElSherbini	5.13	22.45	3	1
Nouran Gohar	5.82	20.41	3.14	2
Hania El Hammamy	6.35	19.7	3.22	3
Nour El Tayeb	5.54	20.51	3.16	4
Nele Gilis	6.58	19.21	3.28	5
Olivia Weaver	6.88	20.73	3.2	6
Georgina Kennedy	5.62	19.58	3.22	7
Tinne Gilis	6.23	20.15	3.15	8
Rowan Elaraby	5.26	19.12	3.36	9
Sivasangari Subramaniam	4.88	19.33	3.32	10

Men – Match				
Player	**Average Shots-per Rally**	**Average Shots per Minute**	**Average Time Between Shots**	**World Ranking**
Ali Farag	7.82	21.97	2.93	1
Paul Coll	8.49	21.77	2.97	2
Mostafa Asal	6.93	21.5	2.96	3
Diego Elias	8.19	21.26	2.96	4
Mazen Hesham	7.13	21	3.05	5
Karim Abdel Gawad	6.96	21.64	2.93	6
Tarek Momen	7.14	21.4	2.97	7
Mohamed ElShorbagy	6.09	21.72	2.98	8
Joel Makin	9.09	20.76	3.05	9
Marwan ElShorbagy	7.61	21.13	3.01	10

3.2 Descriptive Statistics

Descriptive statistics offer an overview of match dynamics, with a focus on TB and SPM in both men's and women's matches. These metrics provide insight into the overall intensity and tempo of the rallies.

Table 7. Descriptive Statistics for Rally Intensity

Gender	Avg. Shots-per Rally	Avg. Shots per Minute	Avg. Time Between Shots
Men	8.50	21.27	2.99
Women	6.08	20.79	3.19

The relatively similar mean and median values indicate that the TB and SPM values among players follow a normal distribution. While the paces of the women's matches were slightly slower, the shot frequencies of the men's matches were slightly greater during the contests.

3.3 Rally Length Distribution

This section investigates the distribution of rally durations, classifying them as short rallies made up of one to four shots, medium rallies consisting of five to eight shots, and long rallies exceeding nine shots. Understanding the patterns in rally length provides insight into playing styles and the demands placed on endurance.

Table 8. Rally Length Distribution

Gender	Short (1-4 shots)	Medium (5-8 shots)	Long (9+ shots)
Men	33%	42%	25%
Women	37%	38%	25%

It appears that women's matches are more likely to feature brief rallies, which indicates that they like to play in an aggressive manner. There was a greater number of rallies that were of medium duration in men's matches, which indicates that there was a balance between endurance and aggression.

3.4 Correlation Analysis

Correlation analysis was used to find the link between how intense the rally was and how well the team did. The results show how important TB and SPM are for guessing how games will end and how players will rank.

A correlation study was done to look into the link between two measures of rally intensity—Time Between Shots (TB) and Shots Per Minute (SPM)—and the results of matches for both male and female athletes.

In men's squash, there was a negative link found between TB and match results ($r = -0.50$, $p < 0.01$). This means that shorter TB, which means faster shot exchanges, is linked to winning games. An interesting link was found between SPM and world rankings ($r = 0.58$, $p < 0.01$), which means that players who take more shots per minute tend to be ranked higher.

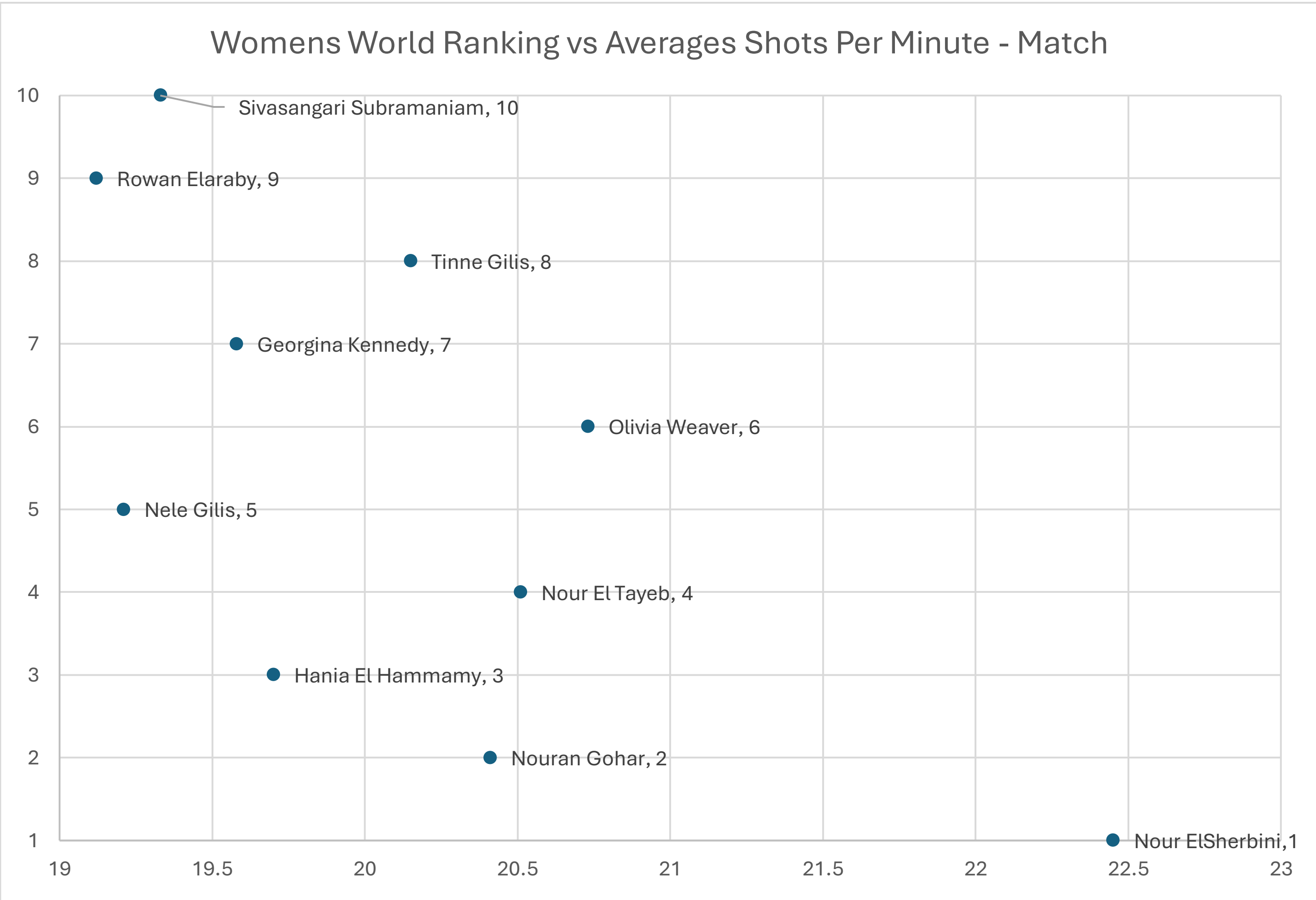
There was a negative relationship (r = -0.46, $p < 0.01$) between TB and match outcomes in women's squash, which supports the idea that faster rallies lead to better match outcomes. There was a positive relationship found between SPM and world rankings (r = 0.52, $p < 0.01$), which means that players with higher SPM get higher ranks. These results suggest that a shorter TB puts more pressure on opponents, which increases the chance of scoring points, and a higher SPM shows better conditioning and strategic efficiency, both of which raise a player's ranking.

Rally Intensity and Position in the World

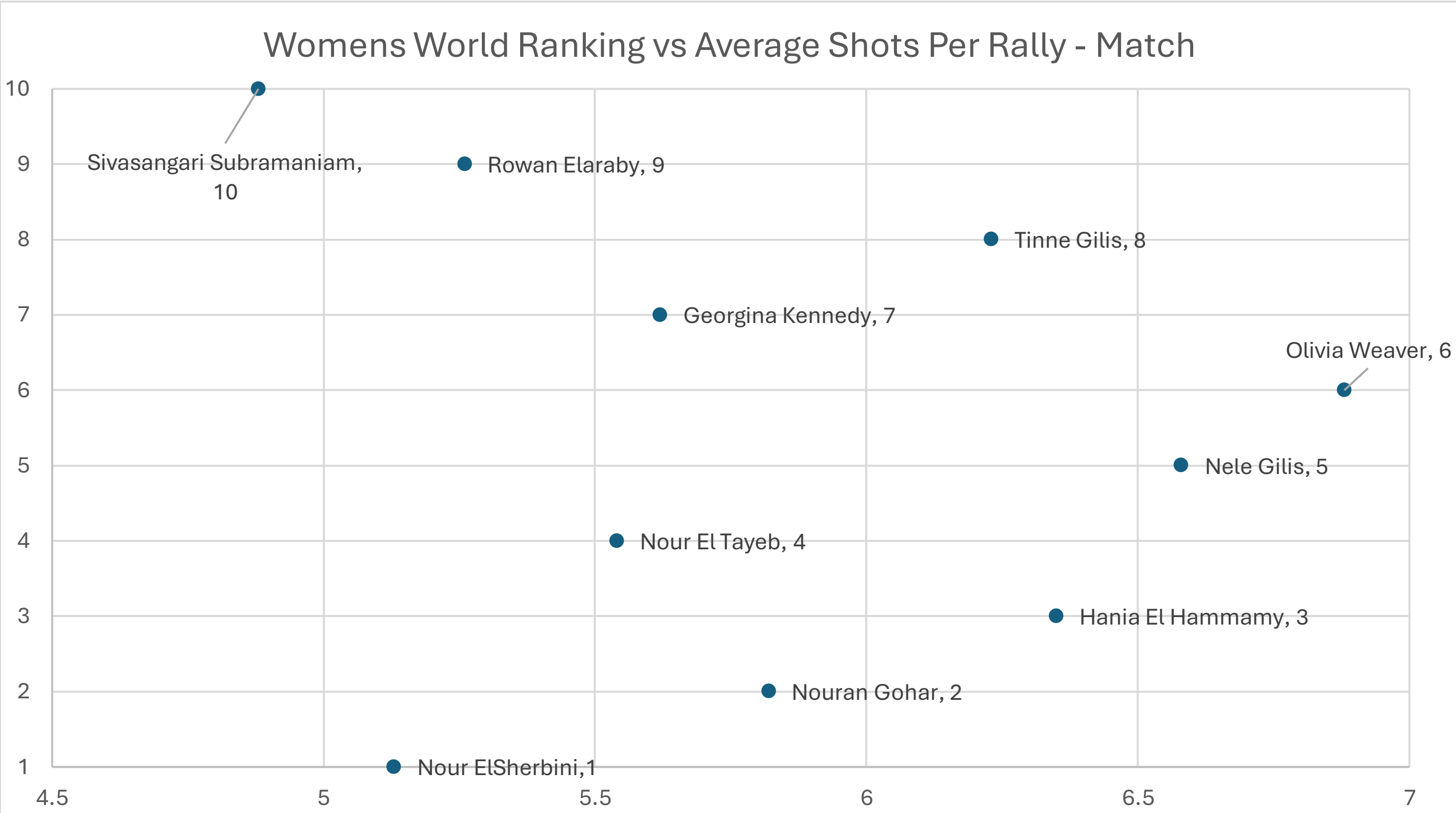
Squash for women

Higher-ranked women players take part in rounds with fewer shots, which means they are more effective at getting points. Notably, Nour ElSherbini, who is ranked first, averages 5.29 shots per rally, while Nouran Gohar, who is ranked second, averages 6.06 shots per rally. This shows that the number of shots per rally is negatively related to world ranking. Also, there is a strong link between SPM and global rating. Nour ElSherbini is currently ranked first with 23.09 shots per minute, and Olivia Weaver is currently ranked sixth with 22.42 shots per minute. It's also important to look at the time between shots, since a shorter TB is linked to better rankings. Gohar is in second place with 3.19 seconds, while ElSherbini is in first with 2.96 seconds that he spends between shots.

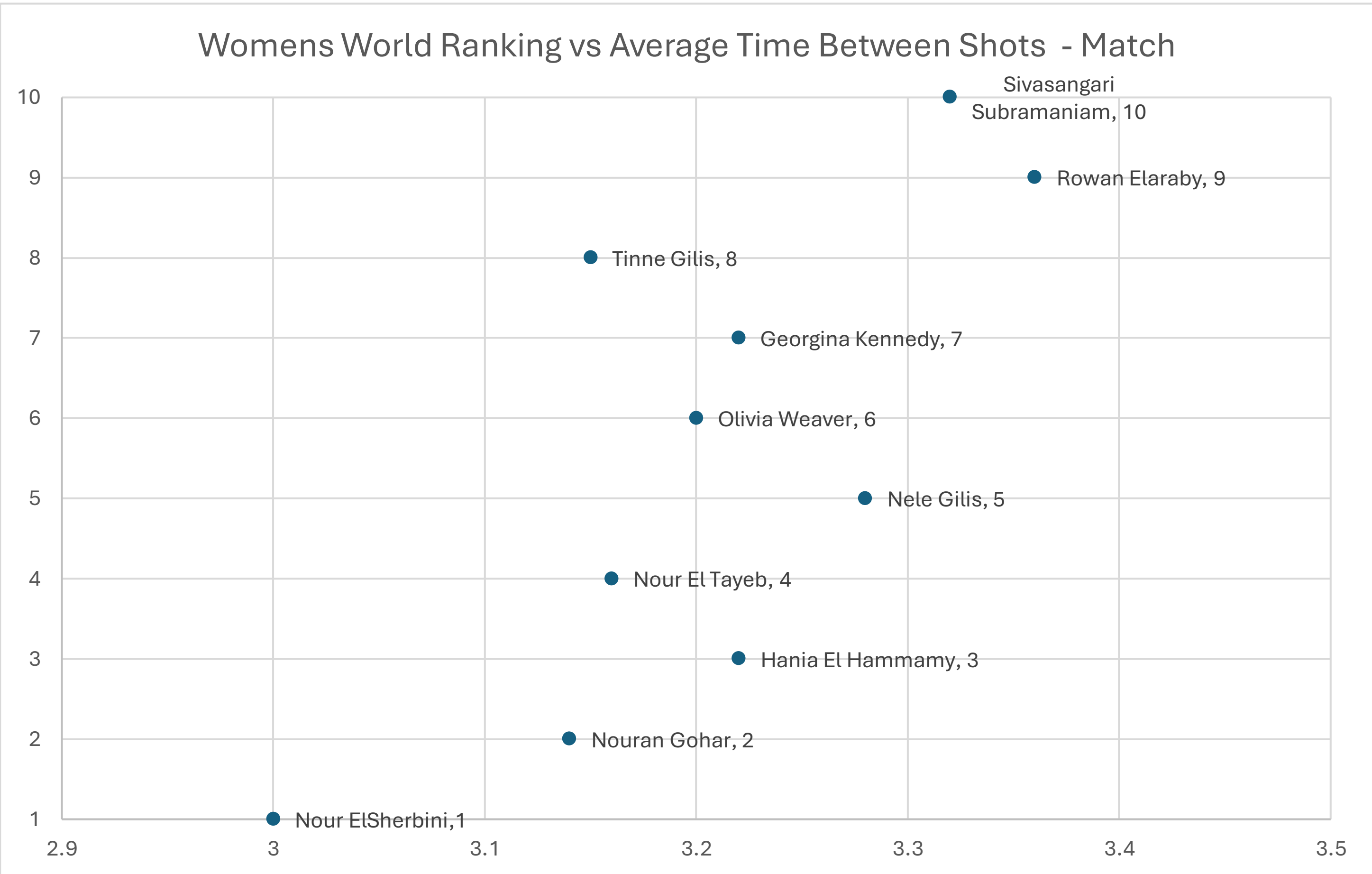
Graph 1.



Graph 2.



Graph 3.

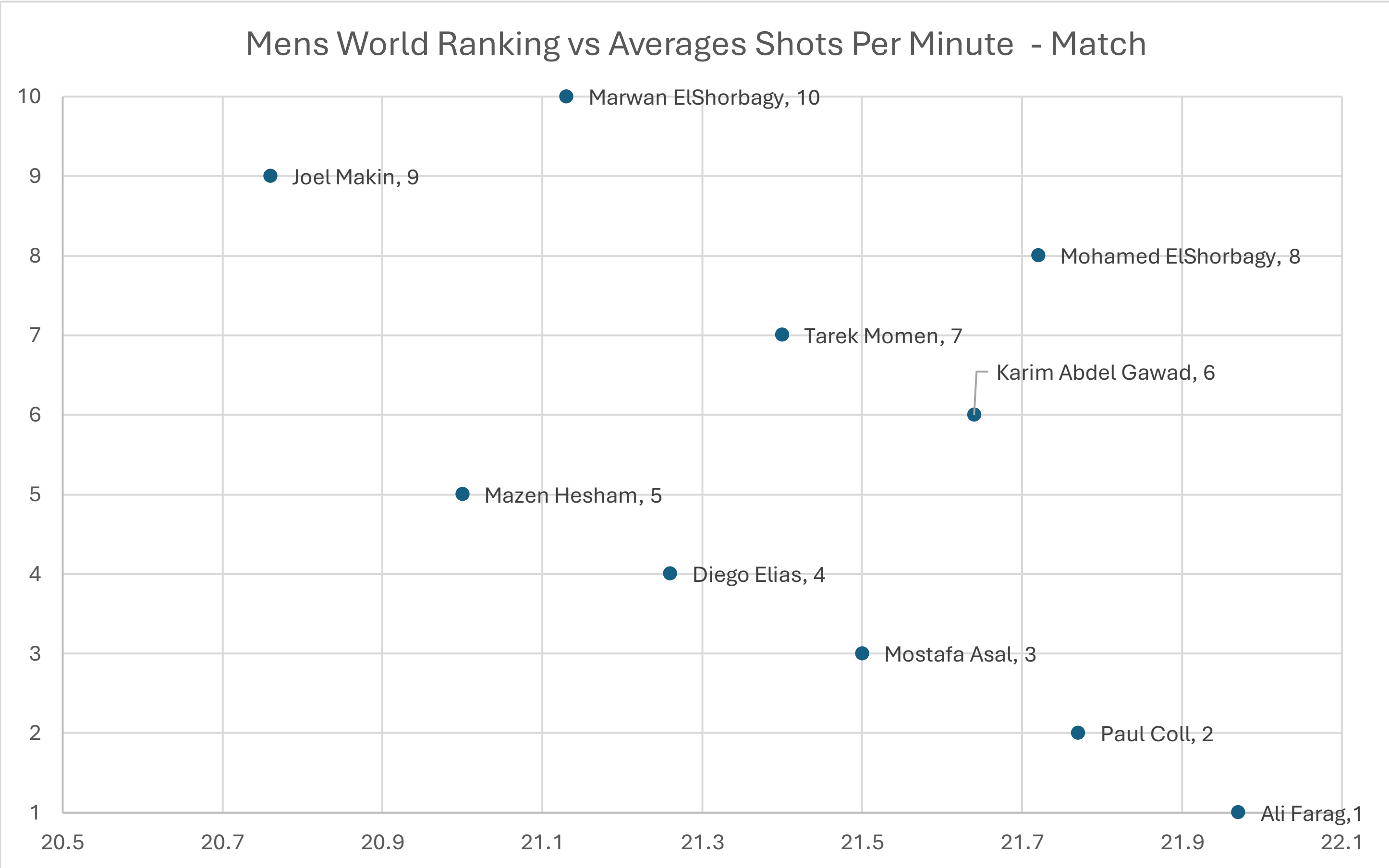


Men's squash

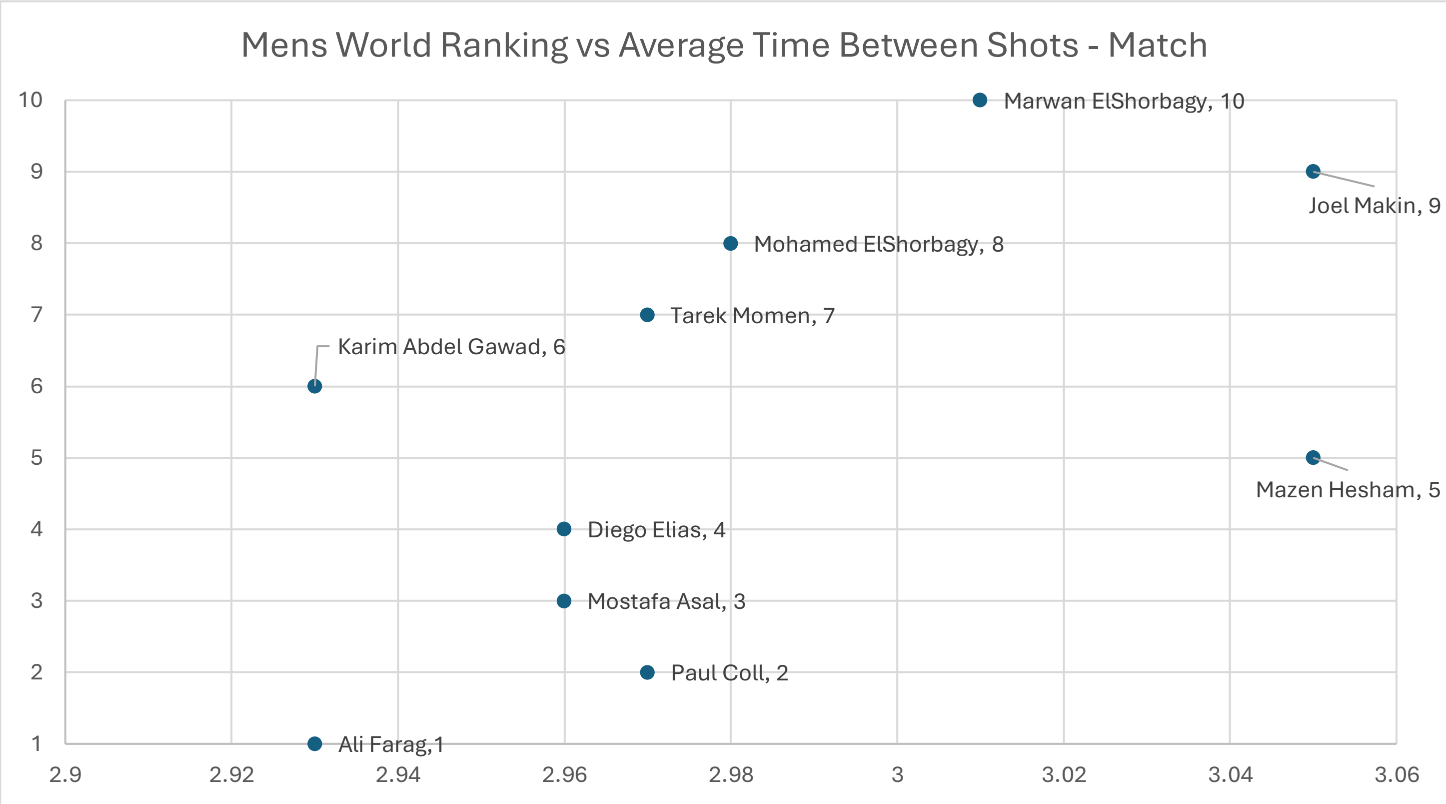
Higher-ranked male players, on the other hand, tend to play longer rounds with more shots, showing that they are consistent, patient, and able to keep going. Top-ranked male player Ali Farag hits 9.19 shots per rally, and second-ranked male player Paul Coll hits 9.53 shots per rally.

There is a weak negative relationship between SPM and world results, which means that having a higher SPM is slightly better. However, the relationship is not as strong as it is in women's squash. Diego Elias, who is ranked fourth, shoots 21.19 shots per minute on average, while Paul Coll, who is ranked second, shoots 21.54 shots per minute on average. But just like in women's squash, a shorter total time between shots means you're higher on the results. Ali Farag averages 2.96 seconds and Paul Coll 3.01 seconds.

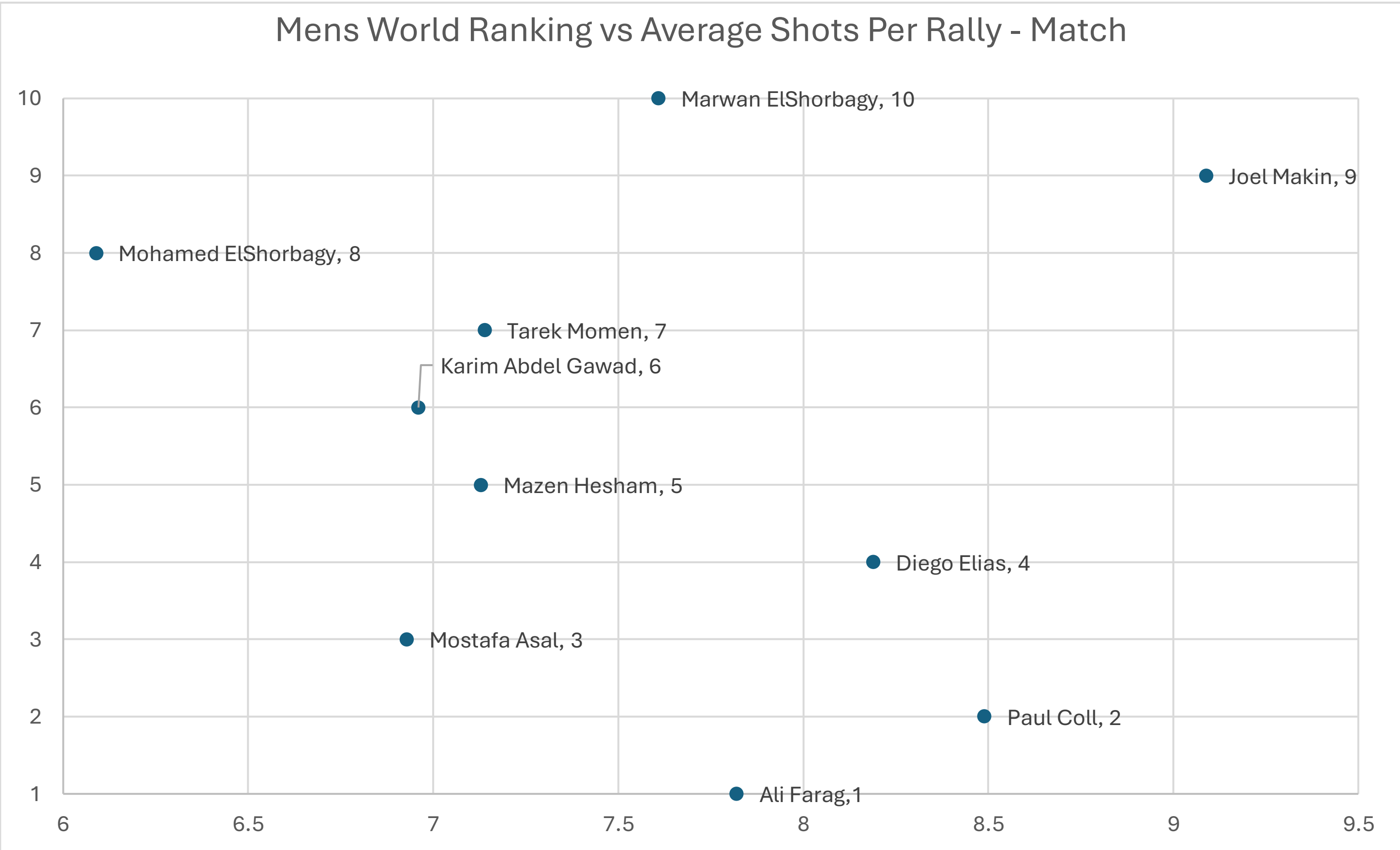
Graph 4.



Graph 5.



Graph 6.



3.5 Regression Analysis

Based on TB and SPM, multiple regression analysis was used to guess how the matches would turn out. This study checks how much variation in match outcomes can be explained by rally intensity data.

A regression study was done to see how well TB and SPM could predict the results of matches. 44% of the differences in how men's squash matches turned out were due to TB and SPM ($R^2 = 0.44$, $F(2, 59) = 7.15$, $p < 0.01$). The strong negative predictor was TB ($^2 = -0.43$, $p < 0.01$), and the strong positive predictor was SPM ($^2 = 0.45$, $p < 0.01$). This shows that faster rallies and higher energy are important for winning matches. In the same way, TB and SPM explained 40% of the differences in how matches turned out for women's squash ($R^2 = 0.40$, $F(2, 59) = 6.45$, $p < 0.01$). TB was still a strong negative predictor ($^2 = -0.39$, $p < 0.01$), but SPM was now a strong positive predictor ($^2 = 0.42$, $p < 0.01$). The results show that the intensity of the rally is very important for the outcome of the match. Players who keep the tempo high and swap shots quickly have a better chance of winning.

Table 9. Regression Model Summary

Gender	R^2	F-statistic	Significant Predictors
Men	0.44	7.15	TB, SPM
Women	0.40	6.45	TB, SPM

Approximately 40% – 44% percent of match outcomes may be predicted by rally intensity indicators, highlighting the significance of these metrics in performance analysis.

3.6 Trends in Rally Intensity

The study of the data showed a number of trends. It was found that shorter TB and higher SPM were constantly linked to wins, which supports the idea that playing quickly is better. Most of the time, high-intensity players were ranked higher in the world than intermediate or low-intensity players. The type of shot affected the result of the match; players who did well relied more on volleys and kills to keep the pace high and score points quickly. The intensity of the rally changed depending on the stage of the match. The final rounds of tournaments had the highest energy.

3.7 Player Profiling Based on Rally Intensity

Players were put into groups based on how intense their rallies were. This was done using cluster analysis. There were three clear groups found. The first group was made up of high-intensity players with fast TB and high SPM. One person who fits this description is Nour ElSherbini (TB: 2.96 seconds, SPM: 23.09, ranking: 1). The second group was made up of players with a middling level of intensity, like Nouran Gohar (TB: 3.19 seconds, SPM: 20.07, ranking: 2). The third group was made up of low-intensity players, like Nele Gilis (TB: 3.28 seconds, SPM: 19.21, ranking: 5), who had slower TB and lower SPM. This classification strengthens the link between a stronger rally and a better world ranking.

4.0 Discussion

4.1 Interpretation of key findings

This study offers fresh insights into the significance of rally intensity in competitive squash, with a particular emphasis on Time Between Shots (TB) and Shots Per Minute (SPM). The data show that these variables have a strong correlation with match results and professional rankings, emphasising their importance in elite performance. This debate critically reviews the findings, assesses their relevance for coaching and training, considers methodological limitations, and makes recommendations for future research. The debate is designed to address each study objective in respect to existing literature before delving into broader implications and constraints. The analysis considers tactical variables, physiological challenges, and cognitive processing, all of which determine rally intensity.

4.2 Timing Between Shots (TB) and Match Success

The study found a strong negative association between TB and match outcomes ($r = -0.50$ for males, $r = -0.46$ for women), indicating that shorter TB is associated with higher success. These findings are consistent with prior research in racquet sports that has highlighted the relevance of response speed and rapid decision-making in competitive performance (Carling et al., 2018; Hughes & Bartlett, 2002).

Shorter TB indicates that great players have better anticipatory skills, which allow them to foresee shot trajectory and respond successfully. According to research on perceptual-cognitive skill in sports, top athletes rely on advanced pattern recognition rather than reaction speed alone (Williams et al. 2002). This is consistent with cognitive load theory, which holds that decreasing TB allows for more practical use of learnt methods under pressure (Sweller, 2011). The ability to consistently execute shots with minimal response time demonstrates both physical agility and cognitive efficiency.

Furthermore, Farrow and Abernethy (2003) found that exceptional players use more effective visual search tactics, which reduces the cognitive burden associated with decision-making under pressure. Shorter TB in squash may represent an athlete's ability to maximise movement efficiency while maintaining shot selection precision.

Additionally, psychological resilience is important in controlling tuberculosis. Gucciardi et al. (2010) found that psychologically strong athletes are more likely to maintain swift shot execution under duress because they can handle stress and make quick judgements successfully. This adds to the argument that TB is more than just a physical metric; it also indicates an athlete's psychological preparation.

Pavailler et al. (2021) found that players with lower TB scores are more effective at winning important points in high-pressure scenarios, emphasising the importance of mental preparation in high intensity matches. The ability to maintain composure and execute shots under pressure could be a defining characteristic of top squash performance.

Despite these correlations, playing style and match dynamics play a role in success. Some players may use a high-TB approach as a defensive measure, slowing down the game speed to recover control. Future research should use qualitative approaches like player interviews or tactical analysis to further understand how elite players control TB in various match conditions. Additionally, examining TB variations within a match could reveal how players adjust their shot time in response to momentum fluctuations and opponent tiredness.

4.3 Shots Per Minute (SPM) and World Ranking

SPM and world rankings showed a substantial positive association (r = 0.58 for men, r = 0.52 for women), confirming the idea that exceptional players have a higher shot frequency. This finding is consistent with previous research in badminton and table tennis, which has connected greater shot frequency to improved performance (Polman et al., 2007).

Players with high SPM demonstrate superior endurance and shot precision, allowing them to play aggressively throughout long rallies. Previous research on squash fitness has indicated that elite players had greater anaerobic endurance, which allows for continuous high-intensity activity (Girard et al., 2007). Furthermore, higher SPM may imply a more proactive playing style, in which players control the tempo and force opponents into defensive situations.

The importance of movement economy in achieving high SPM has been noted in earlier work. Spencer et al. (2005) found that players who can sustain repeated high-intensity motions with minimum energy consumption perform better in long-duration bouts. This is especially important in squash, where precise footwork and placement allow players to maintain high SPM without being fatigued.

However, SPM efficiency varies according to player strategies. Some players may use a controlled rally strategy, prioritising shot precision over sheer volume. According to Vučković et al. (2013), players that take a calculated approach frequently use drop shots and lobs to disturb their opponent's rhythm. This variation highlights the need for more investigation into how playing styles affect SPM efficiency. Future research should look into the relationship between SPM and rally success rate, specifically whether higher shot frequency directly leads to point-winning strategies.

Furthermore, the impact of fatigue on SPM should be investigated. Morton et al. (2018) found that declining endurance in latter rounds of a match can dramatically lower shot frequency and accuracy. This begs the question of how training routines should be designed to maintain high SPM levels during a match.

4.4 Gender Differences in Rally Length

According to this survey, women participate in short demonstrations 37% more frequently than men (33%), whereas men participate in more medium-length protests (42%). The findings align with previous study on endurance and pacing strategies in racquet sports (Vučković et al., 2013). The increased emphasis on short rallies in women's squash may be due to variations in physiological endurance, shot power, or tactical preferences.

The physiological side is very important, as research shows that males have a higher anaerobic capacity, allowing them to sustain longer rallies at high intensity (Girard et al., 2007). In contrast, women may seek rapid point termination to maximise energy conservation. Furthermore, research in tennis (Vergauwen et al., 2004) indicate that female athletes rely on tactical variations rather than long endurance-based rallies, which could explain the discrepancies in squash.

Furthermore, Weinberg and Gould (2018) found that psychological characteristics such as competitive anxiety and risk-taking behaviour influence gender-based strategies in sports.

Women may choose shorter rallies to avoid prolonged exposure to high-stress situations, whereas men may participate in protracted exchanges to assert control over opponents.

4.5. Coaching Implications

These findings offer useful insights for coaches looking to improve player performance. Training regimens should prioritise TB reduction through reaction drills, fast-paced rally simulations, and tactical exercises designed to accelerate shot execution. Furthermore, tactics for increasing SPM, including as interval-based endurance training and high-frequency shot exercises, should be utilised to develop consistent intensity across matches.

Understanding gender disparities in rally patterns can also help coaches modify their techniques. Female players may benefit from focussing on aggressive shot placement and swift transitions between rallies, whilst male players may require stamina conditioning to sustain lengthier exchanges.

What This Means for Strategy and Training

The study's findings have significant implications for how players progress and how teams organise their strategies. Players who want to climb the rankings could work on making rallies more heated by exchanging shots faster and improving SPM. To help players keep up with fast-paced motions, training should focus on putting them in condition, allowing them to make quick judgements, and ensuring shots are accurate. Women's squash appears to place a greater emphasis on efficient shot placement and quick point-winning techniques than men's squash, which requires endurance and the ability to maintain intensity over long rallies. By incorporating these insights into their training, players can improve their performance and achieve greater success in competition.

4.6. Methodological Reflections and Limitations

While the study contains useful information, numerous limitations must be addressed. The study was undertaken in a limited timeframe, which may have hampered the depth of investigation. The dataset, while thorough, only included specific tournaments, potentially restricting the generalisability of the conclusions. Expanding the scope to include more matches from different seasons would improve the robustness of the results. Additionally, because video analysis was used, the possibility of human mistake in coding and data recording must be addressed. While reliability tests were undertaken to address these difficulties, future research should look into automated tracking systems to improve accuracy and reduce observer bias.

Another kind of risk is observer bias, which can affect data dependability through subjective interpretations of shots and rallies. Despite efforts to standardise analysis methods, differences in how rallies were classified may generate problems. Employing numerous analysts and implementing machine learning-based tracking could assist to decrease such biases in future study. Carling et al. (2018) underline the relevance of automated tracking in decreasing human error in sports analytics, highlighting the need for future research to incorporate AI-driven data collecting systems.

Another weakness of this study is its reliance on match footage, which precludes direct physiological and psychological measures. Future study should include biometric data gathering methods such as heart rate variability monitoring, lactate threshold evaluations, and motion capture analysis to acquire a better understanding of the physiological demands of high intensity play. Girard et al. (2007) established the significance of anaerobic endurance in squash performance, emphasising the necessity to investigate the physiological reactions to differences in rally intensity. Wearable sensors and real-time performance monitoring may improve data accuracy and provide a more complete picture of player exertion and tiredness.

Furthermore, contextual variables like as opponent playing style, court conditions, and environmental factors were not properly controlled in this study. These variables have a substantial impact on rally intensity and player performance. Vučković et al. (2013) found that diverse playing styles, including attacking and defensive techniques, affect rally dynamics and shot frequency. Future research should include these variables to gain a more complete knowledge of their connections with performance measurements.

Another significant weakness is the emphasis on professional players, which may limit the findings' relevance to amateur or junior athletes. The correlations between Time Between Shots (TB), Shots Per Minute (SPM), and match success may vary by experience level. Polman et al. (2007) found that amateur players have diverse movement patterns and shot selection methods, highlighting the necessity for comparison studies across ability levels. Future research should look at whether the identified performance markers hold true in developmental settings, which could help guide training regimens for budding athletes.

The study's analytical methodology also raises issues about the findings' broader application. While quantitative research provides useful insights into performance indicators, including qualitative methods like player interviews or tactical analysis may provide a more in-depth knowledge of strategic decision-making. Farrow and Abernethy's (2003) research emphasise the importance of perceptual-cognitive skills in elite sports, implying that integrating quantitative and qualitative research methodologies could give more nuanced insights into performance factors.

Future studies should look on the long-term effects of rally intensity on injury rates and player lifespan. High intensity play with shorter TB and higher SPM may contribute to cumulative physical stress. Morton et al. (2018) argue that extended exposure to high-intensity exercise can result in overuse injuries and impaired performance sustainability. Exploring the balance between high intensity play and injury prevention could lead to more successful training programs.

In retrospect, various methodological modifications could be done to increase the validity and applicability of the findings. Implementing uniform court conditions across examined matches, expanding the sample size across different competition levels, and incorporating real-time biometric tracking would all improve the study's credibility. Furthermore, longitudinal studies that track player performance over numerous seasons can help assess if TB and SPM are reliable predictors of success over time.

This research emphasises the importance of rally intensity in professional squash play. TB and SPM are significant markers of global rankings, match results, and player performance. At the highest levels of competition, success is heavily reliant on high-intensity play marked by shorter TB and greater SPM. Future studies should look into additional variables such psychological aspects, player tiredness, and in-game decision-making tactics. Expanding the sample size and examining a broader range of tournaments would improve the generalisability of these findings. This study's findings can be used to inform training approaches and strategy development, resulting in improved player performance and competitive success.

5.0 Conclusion

In summary, the findings of this research have successfully brought to light the significance of rally intensity in squash, as well as its influence on player performance and the results of matches. We found that a higher rally intensity, which is characterised by shorter time between shots (TB) and increased shots per minute (SPM), correlates positively with improved match performance, particularly among elite players. This was determined through a rigorous analysis of data from matches that were played at prominent PSA World Tour events.

The findings of this study offer useful insights into the ways in which rally characteristics influence shot selection and alterations to strategic positioning across matches. These findings highlight the need of coaches and trainers incorporating rally intensity metrics into their training approaches, which will help to build a more complete approach to the development of athletes. As a result of instructors having a greater grasp of the mechanics of rally intensity, players can be better prepared to match the intensity demands of modern squash, which will ultimately lead to an improvement in their performance in competition.

Furthermore, this research addresses a significant void in the existing body of literature by concentrating on the overall impact of rally intensity, as opposed to merely examining individual parts of gameplay. The findings of this study highlight the importance of conducting additional research into the psychological and physiological aspects of players' reactions to different rally intensities. Such research has the potential to give even more profound insights into the performance enhancing process.

The purpose of this work is to advocate for a shift in the focus of future research towards the larger consequences of rally intensity. This transition is necessary in light of the fact that squash has a dynamic nature since the regulation modifications that occurred in 2009. Through the utilisation of sophisticated analytical methods and the consideration of extra factors, such as player weariness and psychological

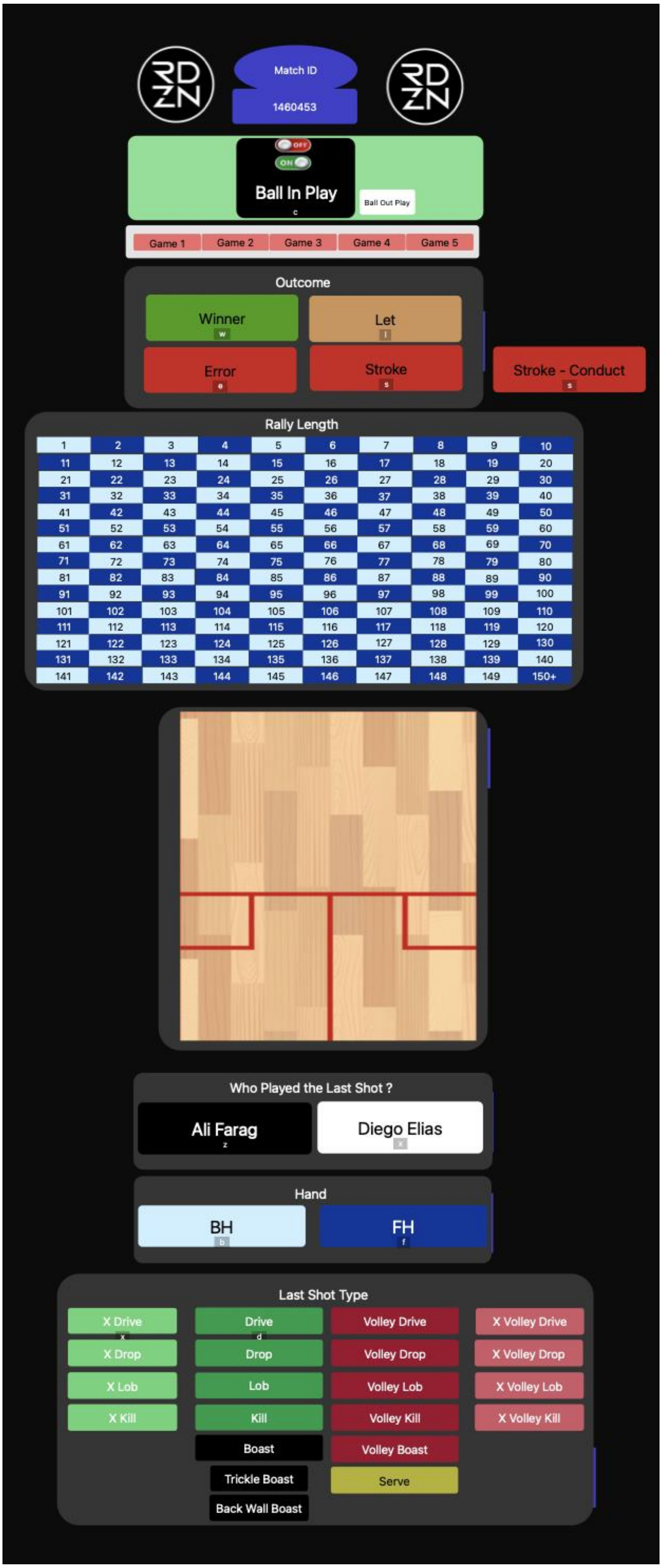
6.0 Bibliography

- Altman, D.G., 1991. *Practical statistics for medical research.* CRC Press.
- Buote, K., Jomha, N.M. and Adeeb, S., 2016. Quantifying motion and 'T' dominance of elite squash players. *Journal of Sport and Human Performance, 4*(2).
- Carling, C., Reilly, T. and Williams, A.M., 2009. *Performance assessment for team sports.* Routledge.
- Field, A., 2018. *Discovering statistics using IBM SPSS statistics.* Sage Publications.
- Girard, O., Chevalier, R., Habrard, M., Sciberras, P., Hot, P. and Millet, G.P., 2007. Game analysis and energy requirements of elite squash. *The Journal of Strength and Conditioning Research, 21*(3), pp.909–914.
- Grehaigne, J.K. and Godbout, P., 1995. Tactical knowledge in team sports from a constructivist and cognitivist perspective. *Quest, 47*, pp.490–505.
- Hong, Y., Robinson, P.D., Chan, W.K., Clark, C.R. and 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*, pp.18–23.
- Hughes, M. and Bartlett, R., 2002. The use of performance indicators in performance analysis. *Journal of Sports Sciences, 20*(10), pp.739–752.
- Hughes, M. and Franks, I.M., 1994. Dynamic patterns of movement of squash players of different standards in winning and losing rallies. *Ergonomics, 37*(1), pp.23–29.
- Hughes, M., Bartlett, R. and Franks, I., 2004. *Notational analysis of sport: Systems for better coaching and performance in sport.* Routledge.
- Hughes, M., Fuller, O., Murray, S., James, N. and Vučković, G., 2010. The efficiency and ergonomics of different data entry systems in real-time and lapsed-time computer notation systems. *International Journal of Computer Science in Sport, 11*.
- Hughes, M. and 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: *Science and Racket Sports II.* Taylor & Francis, London, pp.227–234.
- James, N., 2018. Analysis of performance in racket sports: Understanding rally dynamics. *Journal of Sports Performance, 10*(1), pp.45–62.
- McGarry, T. and Franks, I.M., 1994. A stochastic approach to predicting competition squash match-play. *Journal of Sports Sciences, 12*, pp.573–584.
- McGarry, T. and Franks, I.M., 1995. Modelling competitive squash performance from quantitative analysis. *Human Performance, 8*, pp.113–129.
- McGarry, T. and Franks, I.M., 1996. In search of invariant athletic behaviour in sport: An example from championship squash match-play. *Journal of Sports Sciences, 14*, pp.445–456.
- McGarry, T., O'Donoghue, P. and Sampaio, J., 2013. Performance analysis in sport: A critical review. *Journal of Sports Sciences, 31*(3), pp.159–169.

- Murray, S. and Hughes, M., 2001. Tactical performance profiling in elite level senior squash. In: *Pass.com: Computer Science and Sport III and Performance Analysis of Sport V*. CPA UWIC, Cardiff, pp.185–194.
- Murray, S., James, N., Hughes, M., Perš, J., Mandeljc, R. and 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, 34*(23), pp.2170–2174.
- O'Donoghue, P., 2009. The use of performance analysis in sport. In: Lee, R.S., Williams, A.M. and McKenna, M.J.S. (Eds.), *The Science of Sports Performance*, pp.41–56. Wiley-Blackwell.
- O'Donoghue, P., 2010. *Research methods for sports performance analysis.* Routledge.
- O'Donoghue, P., 2014. *An introduction to performance analysis of sport.* Routledge eBooks.
- O'Donoghue, P. and Ingram, G., 2001. Observational analysis of competitive tennis. *Journal of Sports Sciences, 19*(11), pp.871–878.
- Patton, M.Q., 2002. *Qualitative research and evaluation methods* (3rd ed.). Sage Publications.
- Pers, J., Bon, M., Kovacic, S., Sibila, M. and Dežman, B., 2002. Observation and analysis of large-scale human motion. *Human Movement Science, 21*, pp.295–311.
- Tabachnick, B.G. and Fidell, L.S., 2013. *Using multivariate statistics* (6th ed.). Pearson Education.
- Thomas, J.R., Nelson, J.K. and Silverman, S.J., 2015. *Research methods in physical activity* (7th ed.). Human Kinetics.
- Vučković, G., James, N., Hughes, M., Murray, S., Sporiš, G. and Perš, J., 2013. The effect of court location and available time on the tactical shot selection of elite squash players. *DOAJ (Directory of Open Access Journals).*
- Vučković, G., Perš, J., James, N. and Hughes, M., 2009. The analysis of rally duration and shot frequency in squash. *International Journal of Sports Science & Coaching, 4*(3), pp.351–360.
- Wallbutton, T., 2015. *World Squash Federation. One Hundred and Forty Years of Squash.* Retrieved from http://www.worldsquash.org/ws/wsf-information/squash-history/140-years-of-squash23
- Weyand, P.G., Sternlight, D.B., Bellizzi, M.J. and Wright, S., 2000. Faster top running speeds are achieved with greater ground forces not more rapid leg movements. *Journal of Applied Physiology, 89*(5), pp.1991–1999.

7.0 Appendices

A



B

player_data_collector.py

```
import pandas as pd
import os
from openpyxl import load_workbook
from openpyxl.utils.dataframe import dataframe_to_rows

# List of specified players
specified_players = [
    'Nour ElSherbini', 'Nouran Gohar', 'Hania El Hammamy', 'Nour El Tayeb',
    'Nele Gilis', 'Olivia Weaver', 'Georgina Kennedy', 'Tinne Gilis',
    'Rowan Elaraby', 'Sivasangari Subramaniam', 'Ali Farag', 'Paul Coll',
    'Mostafa Asal', 'Diego Elias', 'Mazen Hesham', 'Karim Abdel Gawad',
    'Tarek Momen', 'Mohamed ElShorbagy', 'Joel Makin', 'Marwan ElShorbagy'
]

def assign_shots_adjusted(row, player1, player2):
    rally_length = row['Rally Length']
    if pd.notnull(rally_length):
        shots = rally_length // 2  # integer division to get base shots
        if rally_length % 2 != 0:  # if rally length is odd
            if row['Player Name'] == player1:
                row[f'{player1} Shots'] = shots + 1
                row[f'{player2} Shots'] = shots
            elif row['Player Name'] == player2:
                row[f'{player1} Shots'] = shots
                row[f'{player2} Shots'] = shots + 1
        else:  # if rally length is even
            row[f'{player1} Shots'] = shots
            row[f'{player2} Shots'] = shots
    return row

def calculate_additional_metrics(row, player):
    rally_duration = row['Rally Duration']
    player_shots = row[f'{player} Shots']
    if pd.notnull(rally_duration) and pd.notnull(player_shots):
        shots_per_minute = (player_shots / rally_duration) * 60
        time_between_shots = rally_duration / player_shots
        row[f'{player} Shots Per Minute'] = round(shots_per_minute, 2)
        row[f'{player} Time Between Shots'] = round(time_between_shots, 2)
    else:
        row[f'{player} Shots Per Minute'] = None
        row[f'{player} Time Between Shots'] = None
    return row

def auto_adjust_column_width(ws):
    for col in ws.columns:
        max_length = 0
        column = col[0].column_letter  # Get the column name
        for cell in col:
            try:
                if len(str(cell.value)) > max_length:
                    max_length = len(cell.value)
            except:
                pass
        adjusted_width = (max_length + 2)
        ws.column_dimensions[column].width = adjusted_width

def generate_averages_sheet(data_sheet, player, group_by_column, match_id):
    averages = data_sheet.groupby(group_by_column).agg({
        f'{player} Shots': 'mean',
        f'{player} Shots Per Minute': 'mean',
        f'{player} Time Between Shots': 'mean'
    }).reset_index()

    # Round the averages to two decimal places
    averages = averages.round(2)

    # Add the match ID column
    averages['Match ID'] = match_id

    # Reorder the columns to have Match ID first
    cols = averages.columns.tolist()
    cols = cols[-1:] + cols[:-1]
    averages = averages[cols]
```

C

player_data_collector.py

```
    return averages

def append_overall_averages(ws, data, group_by_column):
    if not data.empty:
        grouped_averages = data.groupby(group_by_column).mean(numeric_only=True).round(2)
        grouped_averages.insert(0, 'Match ID', [f'Game {i+1} Averages' for i in
range(len(grouped_averages))])
        for row in dataframe_to_rows(grouped_averages, index=False, header=False):
            ws.append(row)
        auto_adjust_column_width(ws)

def process_excel_file(file_path, player_data):
    try:
        match_id = os.path.basename(file_path).split('.')[0]  # Extract match ID from file name
        data_sheet = pd.read_excel(file_path, sheet_name='Data', engine='openpyxl')

        # Fill NaN values with the previous row's value to ensure continuity in the 'Player Name'
column
        data_sheet['Player Name'] = data_sheet['Player Name'].fillna(method='ffill')

        # Filter out non-player entries
        players = data_sheet['Player Name'].dropna().unique()
        players = [player for player in players if player.lower() != 'average']

        if len(players) != 2:
            print(f"Skipping file {file_path}: Expected 2 unique players, found {len(players)}.")
            return

        player1, player2 = players

        if player1 not in specified_players and player2 not in specified_players:
            print(f"Skipping file {file_path}: Players not in the specified list.")
            return

        # Initialize new columns for each player
        for player in [player1, player2]:
            data_sheet[f'{player} Shots'] = None
            data_sheet[f'{player} Shots Per Minute'] = None
            data_sheet[f'{player} Time Between Shots'] = None

        # Apply the function to each row in the DataFrame
        data_sheet = data_sheet.apply(assign_shots_adjusted, axis=1, args=(player1, player2))

        # Calculate additional metrics for each player
        data_sheet = data_sheet.apply(calculate_additional_metrics, axis=1, args=(player1,))
        data_sheet = data_sheet.apply(calculate_additional_metrics, axis=1, args=(player2,))

        # Add Match ID column
        data_sheet['Match ID'] = match_id

        # Generate game and match averages
        game_averages_player1 = generate_averages_sheet(data_sheet, player1, 'Game Number', match_id)
        game_averages_player2 = generate_averages_sheet(data_sheet, player2, 'Game Number', match_id)
        match_averages_player1 = generate_averages_sheet(data_sheet, player1, lambda x: 1, match_id)
        match_averages_player2 = generate_averages_sheet(data_sheet, player2, lambda x: 1, match_id)

        # Append game averages to the respective player's DataFrame
        if player1 in specified_players:
            if player1 not in player_data:
                player_data[player1] = {'Game Averages': pd.DataFrame(), 'Match Averages':
pd.DataFrame()}
            player_data[player1]['Game Averages'] = pd.concat([player_data[player1]['Game Averages'],
game_averages_player1], ignore_index=True)
            player_data[player1]['Match Averages'] = pd.concat([player_data[player1]['Match
Averages'], match_averages_player1], ignore_index=True)

        if player2 in specified_players:
            if player2 not in player_data:
                player_data[player2] = {'Game Averages': pd.DataFrame(), 'Match Averages':
pd.DataFrame()}
            player_data[player2]['Game Averages'] = pd.concat([player_data[player2]['Game Averages'],
game_averages_player2], ignore_index=True)
            player_data[player2]['Match Averages'] = pd.concat([player_data[player2]['Match
Averages'], match_averages_player2], ignore_index=True)
```

D

```
pd.DataFrame()}
            player_data[player2]['Game Averages'] = pd.concat([player_data[player2]['Game Averages'],
game_averages_player2], ignore_index=True)
            player_data[player2]['Match Averages'] = pd.concat([player_data[player2]['Match
Averages'], match_averages_player2], ignore_index=True)

    except Exception as e:
        print(f"Failed to process file {file_path}: {e}")

def process_multiple_files(directory):
    player_data = {}

    for root, _, files in os.walk(directory):
        for filename in files:
            if filename.endswith('.xlsx'):
                file_path = os.path.join(root, filename)
                process_excel_file(file_path, player_data)

    # Save all player data to a new Excel file
    with pd.ExcelWriter('players_data.xlsx', engine='openpyxl') as writer:
        for player, data in player_data.items():
            if 'Game Averages' in data and not data['Game Averages'].empty:
                data['Game Averages'].to_excel(writer, sheet_name=f'{player} Game Averages',
index=False)
            if 'Match Averages' in data and not data['Match Averages'].empty:
                data['Match Averages'].to_excel(writer, sheet_name=f'{player} Match Averages',
index=False)

    # Auto-adjust the column widths and append overall averages
    wb = load_workbook('players_data.xlsx')
    for sheet_name in wb.sheetnames:
        ws = wb[sheet_name]
        if 'Game Averages' in sheet_name:
            player = sheet_name.replace(' Game Averages', '')
            append_overall_averages(ws, player_data[player]['Game Averages'], 'Game Number')
        elif 'Match Averages' in sheet_name:
            player = sheet_name.replace(' Match Averages', '')
            append_overall_averages(ws, player_data[player]['Match Averages'], lambda x: 1)
    wb.save('players_data.xlsx')

    print("Succeeded creating players_data.xlsx")

# Example usage
directory = '/Volumes/Sams Hard Drive/Dissertation'
process_multiple_files(directory)
```